Space–Time Transmit Precoding With Imperfect Feedback

Eugene Visotsky, *Member, IEEE,* and
Upamanyu Madhow, *Senior Member, IEEE*

***Abstract*—The use of channel feedback from receiver to transmitter is standard in wireline communications. While knowledge of the channel at the transmitter would produce similar benefits for wireless communications as well, the generation of reliable channel feedback is complicated by the rapid time variations of the channel for mobile applications. The purpose of this correspondence is to provide an information–theoretic perspective on optimum transmitter strategies, and the gains obtained by employing them, for systems with transmit antenna arrays and imperfect channel feedback. The spatial channel, given the feedback, is modeled as a complex Gaussian random vector. Two extreme cases are considered: *mean* feedback, in which the channel side information resides in the mean of the distribution, with the covariance modeled as white, and *covariance* feedback, in which the channel is assumed to be varying too rapidly to track its mean, so that the mean is set to zero, and the information regarding the relative geometry of the propagation paths is captured by a nonwhite covariance matrix. In both cases, the optimum transmission strategies, maximizing the information transfer rate, are determined as a solution to simple numerical optimization problems. For both feedback models, our numerical results indicate that, when there is a moderate disparity between the strengths of different paths from the transmitter to the receiver, it is nearly optimal to employ the simple *beamforming* strategy of transmitting all available power in the direction which the feedback indicates is the strongest.**

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

I. Introduction

Antenna arrays, at the receiver or at the transmitter, are widely recognized as an effective means of improving the capacity and reliability of a wireless communication link. In a typical cellular or personal communication system, size and complexity limitations preclude deployment of an antenna array at the mobile, usually a small, hand-held unit. On the other hand, it is reasonable to assume that the base station is equipped with an antenna array. In such a setting, the use of transmit antenna arrays provides a powerful method for increasing downlink (base-to-mobile) capacity. There are two key techniques that have been proposed in the literature for exploiting transmit antenna arrays.

1) Space–time coding [1]–[3], which provides diversity in a fading environment. This does not require any knowledge of the spatial channel on the part of the transmitter.
2) Transmit beamforming, which assumes that the transmitter has accurate knowledge of the channel through feedback from the receiver, and can, therefore, perform spatial matched filtering or (in a multiuser context) interference suppression [4], [5].

These two strategies are based on two different, and extreme, assumptions regarding the channel feedback available at the transmitter: space–time coding requires no feedback, whereas beamforming requires accurate feedback. Clearly, there are situations where neither of these assumptions is valid, and one would expect that the transmitter strategy in such situations would be some blend of space–time coding and beamforming. Our purpose in this correspondence is to make this intuition precise by providing information–theoretic insights into the appropriate transmitter strategies when the channel feedback available to the transmitter is imperfect.

Characterization of the information–theoretic capacity of channels with imperfect feedback is the subject of several recent publications. The approach in this correspondence is motivated by the results obtained in [6], [7], which provide forward and converse coding theorems for certain channels with imperfect feedback. A similar feedback model is adopted in [8], where optimum transmission strategies with perfect and imperfect feedback are examined and classified. In [8], a given transmission strategy is classified according to the rank of its input spatial covariance matrix. For instance, a beamforming strategy corresponds to a rank-one matrix while a covariance matrix with full rank indicates a diversity strategy. Such a classification is also adopted in this correspondence. With perfect channel feedback, it is shown in [8] that the optimal strategy entails transmission in a single direction specified by the feedback (beamforming strategy). Conversely, with no channel feedback, it is shown in [9] and [10] that the optimum strategy is to transmit equal power in orthogonal independent directions (diversity strategy). Many practical diversity transmission strategies are analyzed in [9]. Optimum power control and variable-rate transmission strategies when the same side information is available to transmitter and receiver, as well as when the side information is available only to receiver, are analyzed in [11]. The remainder of this correspondence is organized as follows. Section II contains the system model and formal problem statement. Section III contains our main results. Numerical results are presented in Section IV. Section V contains a discussion of these results and of possible directions for future work.

II. Modeling and Overview

It is assumed that the transmit antenna has M elements, and that the receive antenna has a single element. The channel coefficients from the M transmit elements to the receive element are denoted by the $M \times 1$ complex vector $\boldsymbol{h}$. We consider the following abstraction to model partial knowledge of the channel at the transmitter. The corresponding system model is depicted in Fig. 1.

Problem Setup: The input to the channel is given by the $M \times 1$ complex vector $\boldsymbol{x}$. The receiver knows $\boldsymbol{h}$, and receives

$$y = \boldsymbol{x}^H \boldsymbol{h} + n$$

where $n \sim \mathcal{N}(0, \sigma^2)$ is circularly symmetric complex Gaussian noise with variance $\sigma^2/2$ per dimension, and where $\boldsymbol{a}^H$ denotes conjugate transpose of vector $\boldsymbol{a}$. The transmitter receives channel feedback $\boldsymbol{f}$ from the receiver. Given $\boldsymbol{f}$, the transmitter knows that the channel $\boldsymbol{h}$ is distributed according to a complex Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ denote the mean and covariance of $\boldsymbol{h}$. Note that both $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be functions of $\boldsymbol{f}$.

Problem: For $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, what is the input distribution $p(\boldsymbol{x})$ that maximizes the mutual information $I(\boldsymbol{x}; y\,|\boldsymbol{h})$, subject to $E\{\|\boldsymbol{x}\|^2\} \le P$.

Manuscript received April 17, 2000; revised February 2, 2001. This work was supported by Motorola under the University Partnerships in Research Program and by the National Science Foundation under Grants NCR96-24008 (CAREER) and CCR99-79381. The material in the correspondence was presented in part at the IEEE International Symposium on Information Theory, Sorrento, Italy, June 2000 and at the 34th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, October 2000.

E. Visotsky is with the Communication Systems and Technologies Labs, Motorola Labs, Schaumburg, IL 60196 USA (e-mail: visotsky@labs.mot.com).

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

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

Publisher Item Identifier S 0018-9448(01)05458-X.

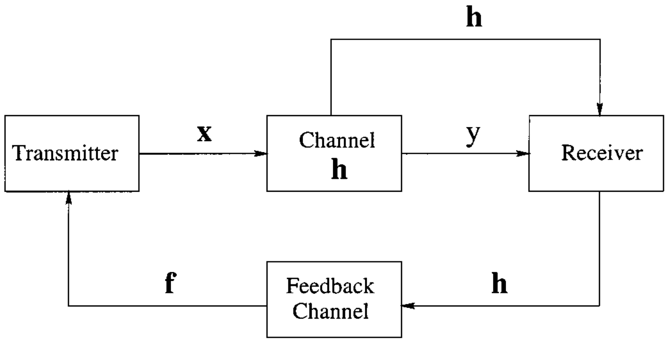

Fig. 1. System model.

Note: For notational simplicity, throughout this correspondence $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is assumed to be a *proper*[1] complex Gaussian random vector, so that its distribution is completely specified by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. If $\boldsymbol{h}$ is not a *proper* complex Gaussian vector, then the problem needs to be reformulated in terms of real-valued variables. Our main results, Theorems 3.1 and 3.2, apply to real as well as complex-valued channels.

The preceding maximization problem can be simplified as follows. Let $p^o(\boldsymbol{x})$ be the maximizing input distribution with covariance matrix $\boldsymbol{Q}$ and power $E\{\|\boldsymbol{x}\|\}^2 = \mathit{trace}\{\boldsymbol{Q}\} = P$. For fixed $\boldsymbol{h}$, the channel is an additive Gaussian noise channel, with input distribution constrained to have covariance $\boldsymbol{Q}$. As shown in [10], the maximum mutual information of this channel is $\log(\boldsymbol{h}^H\boldsymbol{Q}\boldsymbol{h}/\sigma^2 + 1)$ and the maximizing input distribution is proper complex Gaussian, $p^o(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$. Since $p^o(\boldsymbol{x})$ is not a function of $\boldsymbol{h}$, it also maximizes $I(\boldsymbol{x}; y\,|\boldsymbol{h})$. The optimization problem is now one of finding the optimum choice of the covariance matrix $\boldsymbol{Q}^o$ maximizing the mutual information $I(\boldsymbol{x}; y\,|\boldsymbol{h})$ for power constraint P. The problem can be stated as follows:

$$\max_{\boldsymbol{Q}} E_{\boldsymbol{h}} \left\{ \log\left(\frac{\boldsymbol{h}^H\boldsymbol{Q}\boldsymbol{h}}{\sigma^2} + 1 \right) \right\} \tag{1}$$

subject to the power constraint $\mathit{trace}\{\boldsymbol{Q}\} = P$, where σ^2 is variance of the additive circularly symmetric complex Gaussian noise. Note that the expectation in (1) is computed using distribution $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

A given transmission strategy is completely characterized by its covariance matrix $\boldsymbol{Q}$. The strategy consists of transmitting independent complex circular Gaussian symbols along the corresponding directions specified by the eigenvectors of $\boldsymbol{Q}$, with the corresponding eigenvalues specifying the powers allocated in each direction. Adapting definitions in [8], a transmission strategy is defined as beamforming if the rank of $\boldsymbol{Q}$ is one, and as q-fold diversity if $\mathit{rank}\{\boldsymbol{Q}\} = q$. In other words, beamforming is a strategy where transmission is performed only in a single direction, while q-fold diversity utilizes q transmit directions. Under this definition, the space–time coding techniques recently proposed in [1]–[3] can be viewed as attempting to provide full (M-fold) diversity.

A. Discussion

When the channel feedback $\boldsymbol{f}$ is a deterministic function of the channel realizations $\boldsymbol{h}$, recent results by Caire and Shamai [7] and by Viswanathan [6] imply that an achievable information transfer rate between the input and the output in the preceding system model can be computed in the following two steps.

1) At time t, based on the "instantaneous" information provided by the current channel feedback vector $\boldsymbol{f}(t)$, find the optimum input distribution which maximizes the mutual information, subject to an instantaneous power constraint $E\{\|\boldsymbol{x}(t)\|^2\} \le P(t)$.

2) Based on the solution to Step 1), and on the joint distribution of the channel and the feedback, find the optimum power profile $\{P(t), 0 \le t < \infty\}$.

While the results reported in this correspondence may be viewed as solving Step 1) of the preceding approach in a specific setting, such an interpretation would not hold in most practical situations, where the feedback may be noisy. In this case, maximization of the mutual information based on the instantaneous feedback as outlined in Step 1) need not be optimal, since knowledge of the feedback values at multiple times can help combat the randomness in the feedback. However, the framework of this correspondence can form the basis for a practical, albeit suboptimal, design in such situations. For example, the transmitter may derive a (possibly suboptimal) estimate of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ based on more than one value of the feedback, and may then (again suboptimally) use the problem setup in (1) to find a strategy that maximizes the long-term information transfer rate.

Presently, the general solution to the optimization problem in (1) for the general form of $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is not known. The following special cases have been previously considered in the literature.

1) The optimum covariance for the special case of no feedback, $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \alpha\boldsymbol{I})$, is derived by Telatar in [10]. In that case, the diversity strategy with power distributed equally in orthogonal independent transmit directions is optimum.

2) The optimum transmission strategy in the low signal-to-noise ratio (SNR) regime $(\sigma \to \infty)$ for the special case $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$ is found in [8]. It is shown that beamforming in the direction corresponding to the largest eigenvalue of the channel covariance matrix $\boldsymbol{\Sigma}$ is asymptotically optimum as the SNR tends to zero.

B. Summary of Results

In this correspondence, the optimum distribution is characterized in the following two cases.

1) Mean Feedback: In this case, the channel distribution is modeled at the transmitter as $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha\boldsymbol{I})$, where the mean $\boldsymbol{\mu}$ may be interpreted as an estimate of the channel based on the feedback, and α may be interpreted as the variance of the estimation error. In practice, these quantities could be computed at the transmitter based on knowledge of the joint statistics of $\boldsymbol{h}$ and $\boldsymbol{f}$. An example of such a computation is given in Section IV as part of the numerical results for mean feedback.

   The quality of the mean feedback depends on the *feedback SNR* $\|\boldsymbol{\mu}\|^2/\alpha$, and this ratio is zero when no feedback is available. Our results show that the optimum solution is to use beamforming along $\boldsymbol{\mu}$ ($\boldsymbol{Q}$ is unit rank) when the feedback SNR is larger than a threshold, and to use M-fold diversity ($\boldsymbol{Q}$ is full rank) otherwise. In the latter case, the power is distributed according to a water pouring strategy between the direction $\boldsymbol{\mu}$ and the remaining $M - 1$ orthogonal directions, which receive equal powers.

2) Covariance Feedback: The channel distribution known to the transmitter is $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. This models a situation in which the channel may be varying too rapidly for the feedback to give an accurate estimate of the current channel value. However, the relative geometry of the propagation paths changes more slowly, and is reflected in the covariance matrix $\boldsymbol{\Sigma}$. See [13] for a physical channel model justifying these assumptions. In practice, the covariance matrix could be computed at the receiver via long-term time averaging of the channel realizations

[1]As defined in [12], a complex Gaussian vector $\boldsymbol{y}$ with mean $\boldsymbol{\mu}_{\boldsymbol{y}}$ is proper if $E\{(\boldsymbol{y} - \boldsymbol{\mu}_{\boldsymbol{y}})(\boldsymbol{y} - \boldsymbol{\mu}_{\boldsymbol{y}})^T\} = \boldsymbol{0}$.

and reliably transmitted to the transmitter through a low data rate feedback channel.

For covariance feedback, our results show that the optimum solution consists of independent Gaussian inputs along the M eigenvectors of $\boldsymbol{\Sigma}$. The solution resembles water pouring, in that eigenvectors corresponding to larger eigenvalues receive more power (the power along some of the eigenvectors may be zero, so that the optimal diversity order may be less than M).

III. RESULTS

Notation: Let $\boldsymbol{U_A \Lambda_A U_A^H}$ denote the spectral decomposition of a nonnegative definite matrix $\boldsymbol{A}$, where $\boldsymbol{\Lambda_A}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{A}$, and $\boldsymbol{U_A}$ is a unitary matrix containing as columns the eigenvectors of $\boldsymbol{A}$. Likewise, $\boldsymbol{U}^o \boldsymbol{\Lambda}^o (\boldsymbol{U}^o)^H$ is the spectral decomposition of the optimum covariance matrix $\boldsymbol{Q}^o$ and λ_i^o, $i = 1 \cdots M$, denote the optimum eigenvalues. Let $\boldsymbol{A}[i]$ denote the ith column of matrix $\boldsymbol{A}$, and $\boldsymbol{a}[i]$ denote the ith component of column or row vector $\boldsymbol{a}$.

Theorems 3.1 and 3.2 characterize the optimum strategies for mean and covariance feedback, respectively. Proofs of both theorems follows the same general outline: first, we "guess" the optimum transmission directions; next, we project the channel realizations onto these directions; finally, convex optimization arguments are invoked to prove that the guessed directions are indeed optimum.

Theorem 3.1 (Mean Feedback): Let $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha \boldsymbol{I})$. Then, the maximizing covariance matrix in (1) is given by $\boldsymbol{U}^o[1] = \boldsymbol{\mu}/\|\boldsymbol{\mu}\|$, and $\boldsymbol{U}^o[2], \ldots, \boldsymbol{U}^o[M]$ are arbitrarily chosen, except for the restriction that $\boldsymbol{U}^o[1], \ldots, \boldsymbol{U}^o[M]$ are orthonormal. Furthermore,

$$\lambda_2^o = \cdots = \lambda_{\min}^o \equiv \lambda^o$$

where $\lambda^o = (P - \lambda_1^o)/(M-1)$.

Proof: Consider the optimization problem of (1), which can be equivalently expressed as

$$\max_{\boldsymbol{Q}} E_{\boldsymbol{h}} \left\{ \log \left(\boldsymbol{h}^H \boldsymbol{Q} \boldsymbol{h} + \sigma^2 \right) \right\} \tag{2}$$

subject to the power constraint *trace*$\{\boldsymbol{Q}\} = P$. Let $\boldsymbol{U_\mu}$ denote an orthonormal matrix such that $\boldsymbol{U_\mu}[1] = \boldsymbol{\mu}/\|\boldsymbol{\mu}\|$ and $\boldsymbol{U_\mu}[1], \ldots, \boldsymbol{U_\mu}[M]$ are orthonormal and arbitrary. Matrix $\boldsymbol{U_\mu}$ contains as columns our choice for the transmitting directions of the optimum strategy. Below we establish that this choice is globally optimum, that is, $\boldsymbol{U_\mu} = \boldsymbol{U}^o$.

First, consider a restricted optimization problem where the optimization is restricted to only nonnegative definite matrices $\boldsymbol{W}$ satisfying the power constraint, such that $\boldsymbol{U_W} = \boldsymbol{U_\mu}$. Using the spectral decomposition $\boldsymbol{W} = \boldsymbol{U_\mu \Lambda_W U_\mu^H}$ we have that *trace*$\{\boldsymbol{\Lambda_W}\} = P$ as our constraint. The constraint set of the restricted problem is convex and compact, while the cost function is strictly concave in $\boldsymbol{W}$. Hence, there exists a unique matrix $\boldsymbol{W}^o$ (with $\boldsymbol{U_{W^o}} = \boldsymbol{U_\mu}$) which solves the restricted optimization problem. The matrix $\boldsymbol{W}^o$ must satisfy the necessary optimality condition $\delta f(\boldsymbol{W}^o; \boldsymbol{W} - \boldsymbol{W}^o) \le 0$ for all matrices $\boldsymbol{W}$ in the restricted constraint set, where $\delta f(\boldsymbol{W}^o; \boldsymbol{W} - \boldsymbol{W}^o)$ is a Frechet differential [14, p. 178] of the cost function f in the direction $\boldsymbol{W} - \boldsymbol{W}^o$, evaluated at $\boldsymbol{W}^o$. Differentiating, the necessary condition can be expressed as

$$E_{\boldsymbol{h}} \left\{ \frac{\boldsymbol{h}^H (\boldsymbol{W} - \boldsymbol{W}^o) \boldsymbol{h}}{\boldsymbol{h}^H \boldsymbol{W}^o \boldsymbol{h} + \sigma^2} \right\} \le 0. \tag{3}$$

Consider a projection of the channel $\boldsymbol{h}$ onto the "guessed" directions, given by $\boldsymbol{v} = \boldsymbol{U_\mu^H h}$. The components of $\boldsymbol{v}$ are independent with proper complex Gaussian distributions $\boldsymbol{v}[i] \sim \mathcal{N}(0, \alpha)$, for $i = 2 \cdots M$, and $\boldsymbol{v}[1] \sim \mathcal{N}(\|\boldsymbol{\mu}\|, \alpha)$. Performing this change of variables in (3), we obtain an equivalent necessary condition

$$E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}^H (\boldsymbol{\Lambda_W} - \boldsymbol{\Lambda_{W^o}}) \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{\Lambda_{W^o}} \boldsymbol{v} + \sigma^2} \right\} \le 0 \tag{4}$$

where $\boldsymbol{\Lambda_W}$ is any diagonal matrix with nonnegative entries satisfying the power constraint.

It is now shown that $\boldsymbol{W}^o$ is also optimum for the unrestricted optimization problem, i.e., $\boldsymbol{Q}^o = \boldsymbol{W}^o$. In view of the strict concavity of the cost function and the convexity of the constraint region, a sufficient condition for the overall optimality of W^o is $\delta f(\boldsymbol{W}^o; \boldsymbol{Q} - \boldsymbol{W}^o) \le 0$ for all nonnegative definite matrices $\boldsymbol{Q}$ satisfying the power constraint. Equivalently, the sufficient condition can be expressed as

$$E_{\boldsymbol{h}} \left\{ \frac{\boldsymbol{h}^H (\boldsymbol{Q} - \boldsymbol{W}^o) \boldsymbol{h}}{\boldsymbol{h}^H \boldsymbol{W}^o \boldsymbol{h} + \sigma^2} \right\} \le 0. \tag{5}$$

Again, making the substitution $\boldsymbol{v} = \boldsymbol{U_\mu^H h}$, we obtain as the sufficient condition

$$E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}^H \left(\tilde{\boldsymbol{Q}} - \boldsymbol{\Lambda_{W^o}} \right) \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{\Lambda_{W^o}} \boldsymbol{v} + \sigma^2} \right\} \le 0 \tag{6}$$

where $\tilde{\boldsymbol{Q}} = \boldsymbol{U_\mu^H Q U_\mu}$ is related to $\boldsymbol{Q}$ by an orthonormal transformation, and hence is nonnegative definite with the same trace, so that it also satisfies the power constraint. Decomposing $\tilde{\boldsymbol{Q}}$ into a sum of matrices $\tilde{\boldsymbol{D}}$ and $\hat{\boldsymbol{Q}}$ containing diagonal and off-diagonal terms, respectively, expression (6) is rewritten as

$$E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}^H \left(\tilde{\boldsymbol{D}} - \boldsymbol{\Lambda_{W^o}} \right) \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{\Lambda_{W^o}} \boldsymbol{v} + \sigma^2} \right\} + E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}^H \hat{\boldsymbol{Q}} \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{\Lambda_{W^o}} \boldsymbol{v} + \sigma^2} \right\} \le 0. \tag{7}$$

The first summand in (7) is less than or equal to zero by the necessary condition for optimality given in (4), since matrix $\tilde{\boldsymbol{D}}$ satisfies the power constraint (*trace*$\{\tilde{\boldsymbol{D}}\}$ =*trace*$\{\tilde{\boldsymbol{Q}}\} = P$), and is diagonal and nonnegative. The second sum can be decomposed into a weighted sum of the terms of the form

$$E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}[i]^* \boldsymbol{v}[j]}{\boldsymbol{v}^H \boldsymbol{\Lambda_{W^o}} \boldsymbol{v} + \sigma^2} \right\} \le 0 \;\; i \ne j. \tag{8}$$

Note that the M-dimensional probability distribution of $\boldsymbol{v}$ is symmetric with respect to all axes, excluding the $\boldsymbol{v}[1]$ axis (due to the nonzero mean of $\boldsymbol{v}[1]$). Furthermore, the function under the expectation in (8) is antisymmetric. We conclude that the expectation in (8) is zero, so that the sufficient condition in (7) is indeed satisfied by $\boldsymbol{W}^o$.

It remains to be shown that the eigenvalues λ_2^o through λ_M^o are equal. This easily follows from substituting $\boldsymbol{W}^o$ into the cost function and noticing that the cost function is symmetric with respect to the eigenvalues of interest. An application of the Jensen's inequality yields the desired result. □

To complete the solution, it remains only to specify the value of λ_1^o.

Computation of λ_1^o in Theorem 3.1: The power constraint implies that λ_1^o lies in the range $0 \le \lambda_1^o \le P$. Although a closed-form analytic solution for this quantity does not appear to be available, λ_1^o can be determined numerically by a one-dimensional search over the range. Another possibility is to use the observation that, as a function of the two parameters λ_1^o and λ^o, the cost function is concave, while the power

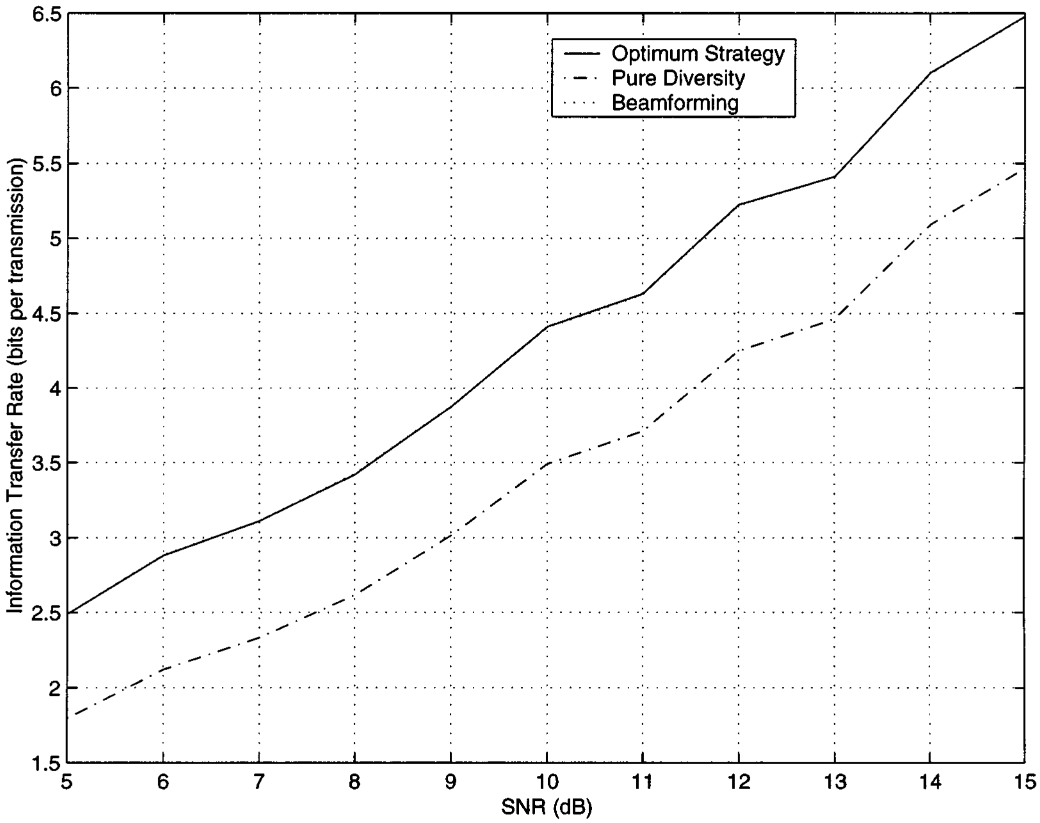

Fig. 2. Information transfer rate achievable with the optimum, pure diversity and beamforming strategies, $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha \boldsymbol{I})$, $M = 2$, $a^d = 0.9$. The curves for the optimum and beamforming strategies coincide.

constraint is convex. Hence, these parameters can be numerically determined by the projected gradient descent algorithm [15], which in this case is guaranteed to converge to the global minimum.

The mean feedback model was previously considered in [8], where a similar computation is performed for two antenna elements $(M = 2)$. Specifically, the optimum transmit directions established analytically in Theorem 3.1 are determined in [8] by numerical simulations, and a sufficient condition on feedback SNR $\|\boldsymbol{\mu}\|^2/\alpha$ for the beamforming strategy to be optimum is given.

Theorem 3.2 (Covariance Feedback): Let $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. Then, the eigenvectors of the maximizing covariance matrix in (1) satisfy $\boldsymbol{U}^o = \boldsymbol{U}_{\boldsymbol{\Sigma}}$. That is, the optimal strategy is to employ independent complex circular Gaussian inputs along the eigenvectors of $\boldsymbol{\Sigma}$.

Proof: For conciseness, only an outline of the argument is presented, since, with slight modifications, arguments used to prove Theorem 3.1 also apply here. In this case, our "guess" for the optimum transmit directions is given by the columns of $\boldsymbol{U}_{\boldsymbol{\Sigma}}$. After projecting the channel onto the "guessed" directions as $\boldsymbol{v} = \boldsymbol{U}_{\boldsymbol{\Sigma}}^H \boldsymbol{h}$, the optimization problem in (1) can be equivalently expressed as

$$\max_{\boldsymbol{Q}} E_{\boldsymbol{v}} \left\{ \log \left(\boldsymbol{v}^H \boldsymbol{Q} \boldsymbol{v} + \sigma^2 \right) \right\} \tag{9}$$

subject to the power constraint *trace*$\{\boldsymbol{Q}\} = P$, where $\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Lambda}_{\boldsymbol{\Sigma}})$. To establish the theorem, it suffices to show that the maximizing covariance matrix in (9) is diagonal.

As in the proof of Theorem 3.1, consider a restricted optimization problem, where the search is restricted to only diagonal nonnegative definite matrices satisfying the power constraint in (9). By the convexity of the restricted optimization problem, there exists a unique diagonal matrix $\boldsymbol{D}^o$, with real nonnegative entries, which solves the restricted optimization problem. Furthermore, $\boldsymbol{D}^o$ satisfies the necessary condition of optimality on the restricted constraint set which is used to show its overall optimality.

Similarly to (7), it is sufficient to establish that $\boldsymbol{D}^o$ satisfies the following inequality for all nonnegative complex matrices $\boldsymbol{Q}$ satisfying the power constraint

$$E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}^H \left(\tilde{\boldsymbol{D}} - \boldsymbol{D}^o \right) \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{D}^o \boldsymbol{v} + \sigma^2} \right\} + E_{\boldsymbol{v}} \left\{ \frac{\boldsymbol{v}^H \hat{\boldsymbol{Q}} \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{D}^o \boldsymbol{v} + \sigma^2} \right\} \leq 0 \tag{10}$$

where $\tilde{\boldsymbol{D}}$ and $\hat{\boldsymbol{Q}}$ contain the diagonal and off-diagonal terms of $\boldsymbol{Q}$, respectively. Following arguments identical to those used to show that (7) holds, the proof of the theorem is concluded. □

Computation of the Optimum Distribution in Theorem 3.2: It follows from Theorem 3.2 that, for $\boldsymbol{\mu} = \boldsymbol{0}$, the transmit directions are precisely the eigenvectors of the channel covariance matrix $\boldsymbol{\Sigma}$. To completely specify the transmission strategy, one needs to solve for the M transmit powers given by the diagonal entries of $\boldsymbol{D}^o$. This is an M-dimensional water pouring problem which can be solved by a number of numerical algorithms, such as the projected gradient descent algorithm [15]. Intuitively, higher power should be transmitted in the directions of the larger eigenvalues of $\boldsymbol{\Sigma}$, since large eigenvalues correspond to stronger, and hence more reliable, channels.

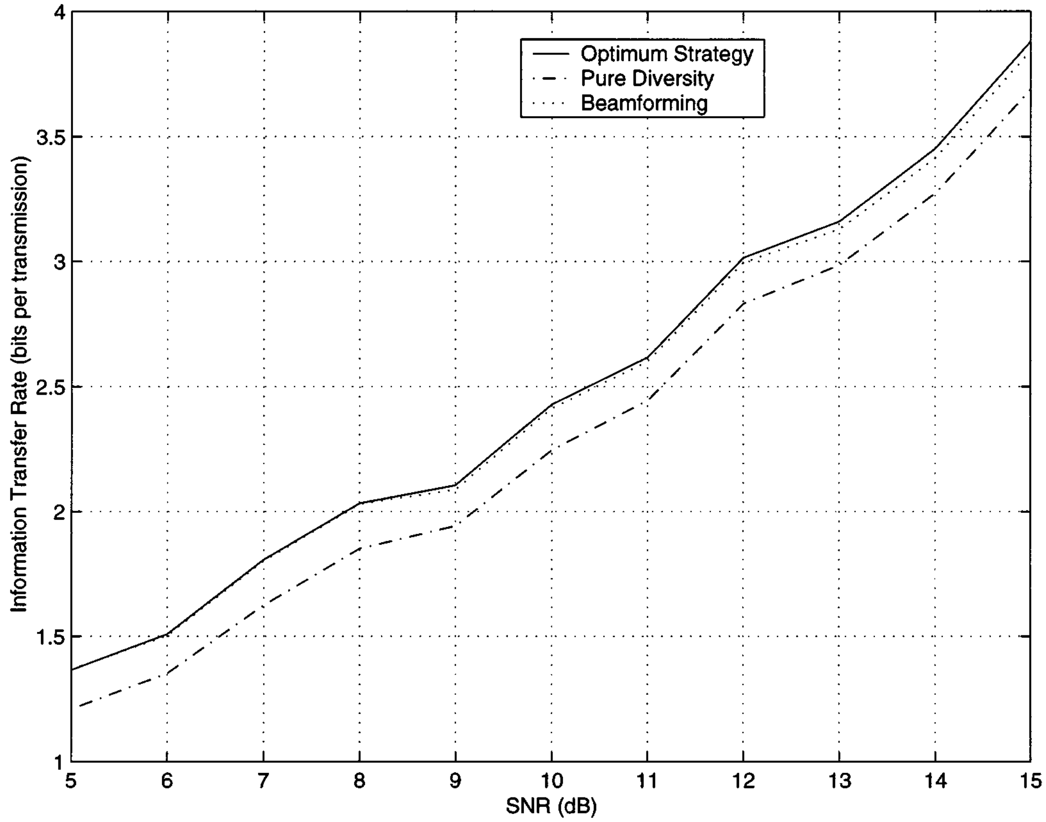

Fig. 3. Information transfer rate achievable with the optimum, pure diversity and beamforming strategies, $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha \boldsymbol{I})$, $M = 2$, $a^d = 0.6$.

IV. Numerical Results

Mean Feedback: $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha \boldsymbol{I})$*:* To obtain numerical results for this case, the following Rayleigh fading channel model is simulated. Let $\boldsymbol{h}(t)$ be an $AR(1)$ random process with forgetting factor a, $\boldsymbol{h}(t) = a\boldsymbol{h}(t-1) + \boldsymbol{w}(t)$, where $\boldsymbol{w}(t)$ is an $M \times 1$ vector of independent and identically distributed (i.i.d.) circularly symmetric Gaussians, each of variance σ_w^2. The feedback channel is modeled as a lossless channel with delay d: at time t, the transmitter observes $\boldsymbol{f}(t) = \boldsymbol{h}(t-d)$ at the output of the feedback channel. It is straightforward to show that, conditioned on $\boldsymbol{f}(t)$, $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha \boldsymbol{I})$, where $\boldsymbol{\mu} = a^d \boldsymbol{h}(t-d)$ and $\alpha = \sigma_w^2(1-a^{2d})/(1-a^2)$. The distribution of $\boldsymbol{f}(t) = \boldsymbol{h}(t-d)$ is the same as that of $\boldsymbol{h}(t)$ and is given by $\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{0}, \sigma_w^2/(1-a^2)\boldsymbol{I})$. Hence, α can be rewritten as

$$\alpha = E\{\|\boldsymbol{h}\|^2\}(1-a^{2d})/M.$$

Figs. 2–4 display information transfer rate (in bits per transmission) achievable with beamforming, diversity and the optimum transmission strategy for a two-element antenna array. The information transfer rate achieved by a given transmission strategy is obtained by substituting an appropriate covariance matrix $\boldsymbol{Q}$ into (1) and removing conditioning on the feedback. The optimum transmission strategy uses the covariance matrix $\boldsymbol{Q}^o$ computed according to Theorem 3.1, where λ_1^o and λ^o are numerically optimized using the projected gradient descent algorithm, as described in the note following Theorem 3.1. Since optimization of the instantaneous input power, given by *trace*$\{\boldsymbol{Q}\}$, is not undertaken, *trace*$\{\boldsymbol{Q}\}$ is set equal to one for all transmission strategies without loss of generality. The input covariance matrix for the beamforming strategy, where beamforming is performed in the direction specified by the conditional mean, is then given by $\boldsymbol{Q}_B = \boldsymbol{\mu}\boldsymbol{\mu}^H/\|\boldsymbol{\mu}\|^2$, while $\boldsymbol{Q}_D = (1/M)\boldsymbol{I}$ specifies the input covariance matrix for diversity transmission. In Figs. 2–4, the information transfer rate is plotted versus SNR (in decibels), where

$$\text{SNR} = E\{\|\boldsymbol{h}\|^2\}/\sigma^2, \qquad E\{\|\boldsymbol{h}\|^2\} = 1$$

and $a^d = 0.9$, 0.6, and 0.3, respectively. Parameter a^d indicates the feedback quality and, more precisely, can be related to the average feedback SNR per antenna element as

$$E\{\|\boldsymbol{\mu}\|^2\}/M\alpha = a^{2d}/(1-a^{2d}).$$

When $a^d = 0.9$, the quality of the feedback is relatively high, which means that the channel estimate specified by the conditional mean is close to the true channel. This results in the beamforming strategy performing identical to the optimum strategy. The diversity strategy in this case is too conservative, losing about 2 dB in performance over a wide range of SNRs as compared to the beamforming strategy. With $a^d = 0.6$, the conditional mean conveys less accurate information about the state of the channel. Nevertheless, the performance of the beamforming and optimum strategies is almost identical. This result is somewhat surprising, and highlights the robustness of the beamforming strategy to imperfections in the feedback. Of course, the gap between the performance of the diversity and optimum strategies is smaller than in Fig. 2. Finally, the channel feedback quality is poor when $a^d = 0.3$. This case turns out roughly equivalent to the case of no feedback, in that the diversity strategy performs close to the optimum strategy, while the beamforming strategy performs about 1 dB worse than the optimum at high SNRs.

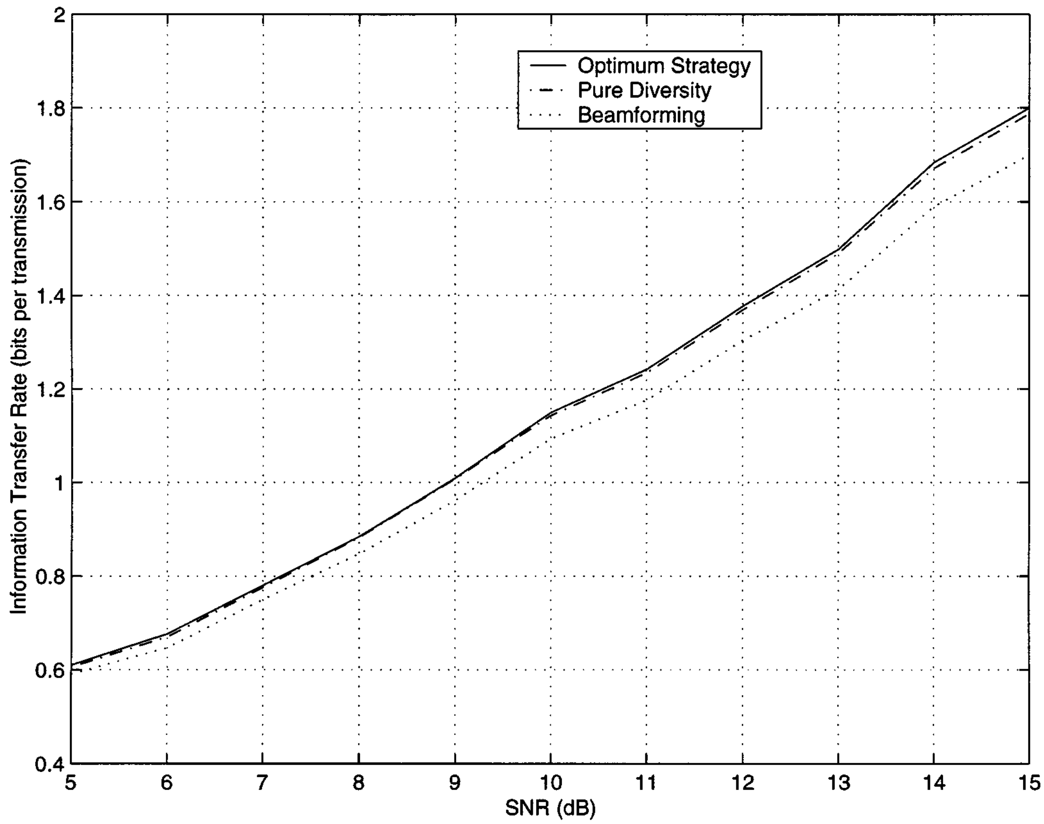

Fig. 4. Information transfer rate achievable with the optimum, pure diversity and beamforming strategies, $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{\mu}, \alpha\boldsymbol{I})$, $M = 2$, $a^d = 0.3$.

Covariance Feedback: $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$*:* In this case, the transmitter has no information about the mean, but has long-term knowledge of the spatial correlation matrix Σ. The achievable long-term information transfer rate for a transmit strategy with input covariance matrix $\boldsymbol{Q}$ is given by evaluating (1) with an appropriate covariance matrix $\boldsymbol{Q}$. In this case, it is not necessary to remove the conditioning on the feedback since the channel feedback is modeled as time-invariant. The input power *trace*$\{\boldsymbol{Q}\}$ and the channel power $E\{\|\boldsymbol{h}\|^2\}$ =*trace*$\{\boldsymbol{\Sigma}\}$ are set to one, and the information transfer rates are obtained for the beamforming, diversity, and optimum transmission strategies as a function of the channel SNR, $E\{\|\boldsymbol{h}\|^2\}/\sigma^2$. Since the eigenvectors of the optimum covariance matrix $\boldsymbol{Q}^o$, by Theorem 3.2, coincide with the eigenvectors of $\boldsymbol{\Sigma}$, without loss of generality, both matrices are taken to be diagonal for the simulations. The diagonal entries of $\boldsymbol{Q}^o$ are determined by the projected gradient descent algorithm. The diversity strategy is implemented by the input covariance matrix $(1/M)\boldsymbol{I}$, and the beamforming strategy is implemented by transmitting all of the available power in the direction of the unit vector corresponding to the largest diagonal entry of $\boldsymbol{\Sigma}$.

Let ν_i denote the ith eigenvalue (in this case the ith diagonal entry) of $\boldsymbol{\Sigma}$. Fig. 5 displays the information transfer rate (bits per transmission) achieved by the transmission strategies with $M = 3, \nu_1 = \nu_2 = \nu_3$. In this case, the performance of the diversity strategy is optimum, while the transmit beamforming strategy (equivalent here to using a single transmit antenna element) performs less than 1 dB worse than the optimum over a wide range of SNR values. Fig. 6 displays the performance of the transmission strategies for $M = 3, \nu_1/\nu_2 = \nu_1/\nu_3 = 2$. In this case, there is a single direction which is 3 dB stronger than the other directions specified by $\boldsymbol{\Sigma}$. Somewhat surprisingly, the beamforming strategy transmitting in the stronger direction achieves the performance of the optimum strategy for all displayed values of SNR. Hence, no penalty is incurred for not utilizing the lower power directions, leading us to conjecture that, in general, such a beamforming strategy is close to being optimal for small M, when $\boldsymbol{\Sigma}$ has at least a moderate eigenvalue spread.

V. Conclusion and Future Work

For a single transmit element, prior work [6], [11], [7] has shown that, in terms of information–theoretic limits, there is little to be gained by exploiting knowledge of the channel at the transmitter for a single transmit antenna element. As our numerical results indicate, for transmit antenna arrays, the gain through even partial knowledge of the channel can be substantial. For mean feedback, the beamforming strategy performs close to the optimal strategy when the feedback is of reasonable quality. The beamforming strategy performs close to the optimal strategy for covariance feedback when there is a stronger path present which can be exploited by the beamforming. Overall, the beamforming strategy appears to be a viable transmission strategy when meaningful channel feedback is present. Furthermore, the use of the beamforming strategy simplifies operation at the mobile, since only a single data stream needs to be decoded. The optimum transmission strategies presented in this work are based on random coding arguments. Hence, future work in this area includes design of practical coding strategies for exploiting partial knowledge of the spatial channel.

A possible extension of the information–theoretic approach taken in this correspondence is to consider optimum transmission strategies in the context of the multiuser system model. The problem can be formulated as a search for the optimum transmission strategies in a broadcast channel as a function of the quality of channel feedback at the transmitter. Such optimum strategies, if found, would provide a valuable benchmark for the performance of the practical transmit beamforming strategies, such as those proposed in [4], [5].

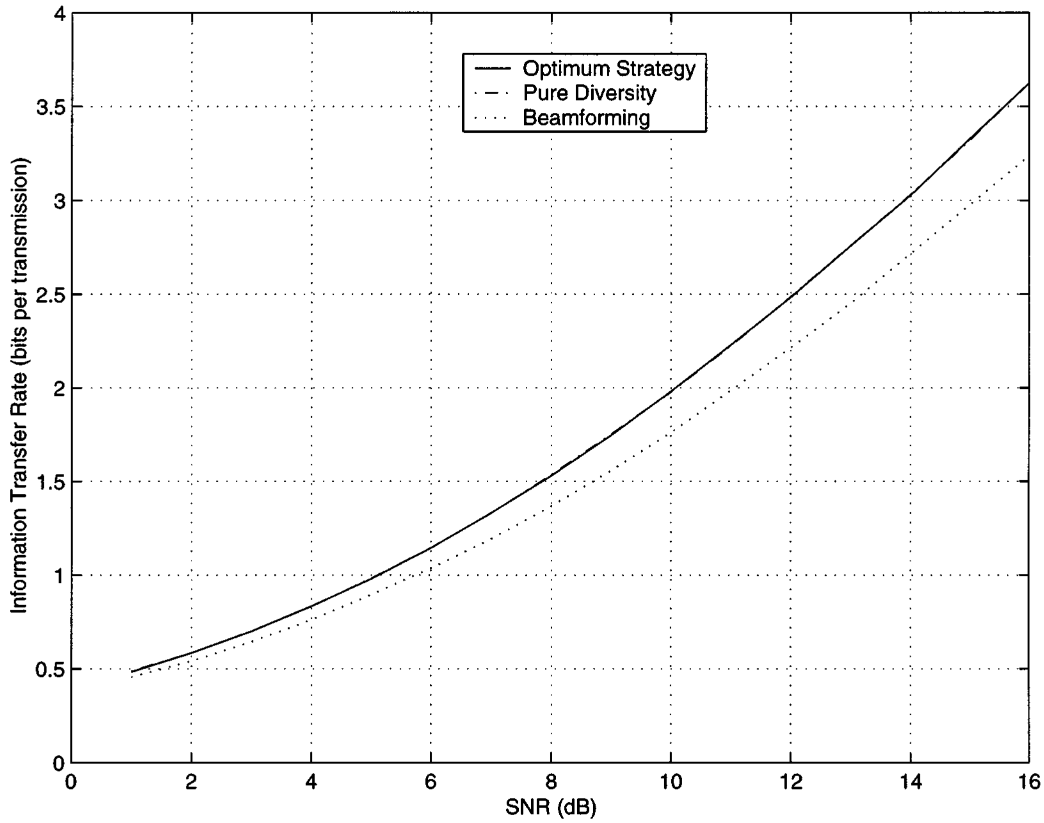

Fig. 5. Information transfer rate achievable with the optimum, pure diversity and beamforming strategies, $\boldsymbol{h} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, $M = 3$, $\nu_1 = \nu_2 = \nu_3$. The curves for the optimum and diversity strategies coincide.

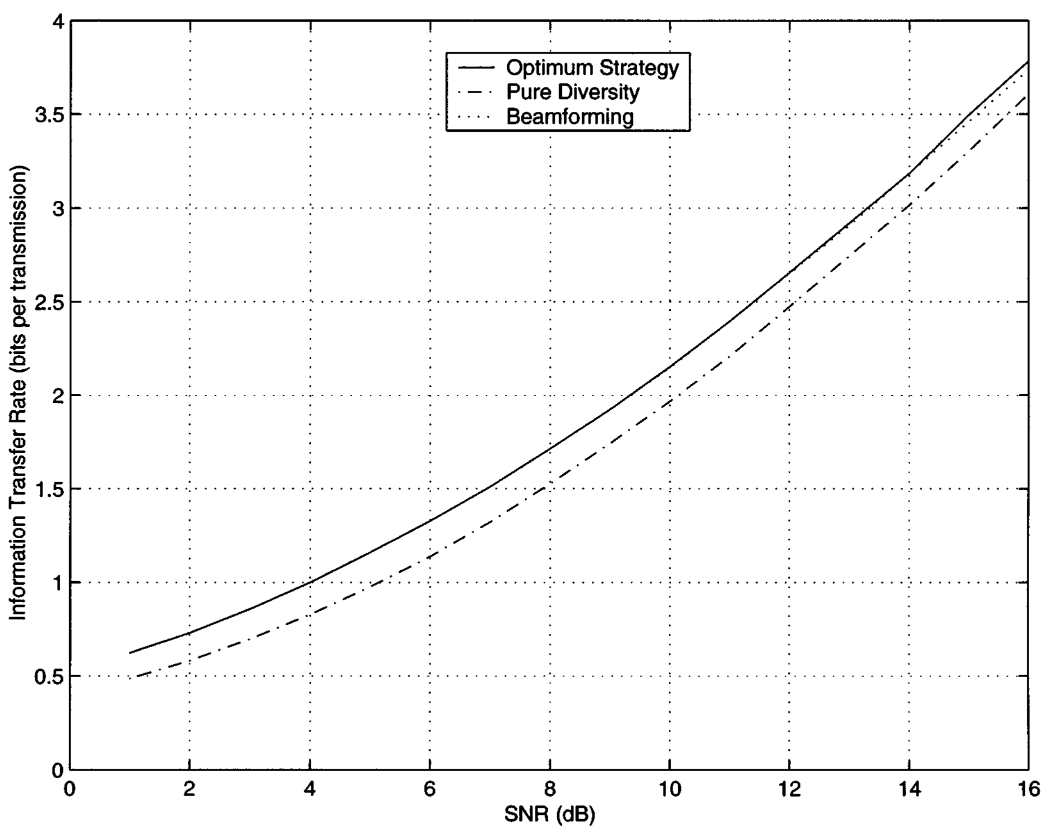

Fig. 6. Information transfer rate achievable with the optimum, pure diversity and beamforming strategies, $\boldsymbol{h} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, $M = 3$, $\nu_1/\nu_2 = \nu_1/\nu_3 = 2$. The curves for the optimum and beamforming strategies coincide.

ACKNOWLEDGMENT

The authors wish to thank B. Hajek and M. Medard for helpful discussions.

REFERENCES

[1] V. Tarokh, N. Seshadri, and A. Calderbank, "Space–time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
[2] B. Hochwald and T. Marzetta, "Unitary space–time modulation for multiple-antenna communications in Rayleigh flat fading," in *36th Annu. Allerton Conf. Communication, Control and Computing*, Sept. 1998.
[3] D. Warrier and U. Madhow, "Noncoherent communication in space and time," in *Proc. Conf. Information Sciences and Systems*, Mar. 1999.
[4] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.
[5] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE 49th Vehicular Technology Conf., VTC'99*, May 1999, pp. 851–856.
[6] H. Viswanathan, "Capacity of Markov channels with receiver CSI and delayed feedback," *IEEE Trans. Inform. Theory*, vol. 45, pp. 744–765, Mar. 1999.
[7] G. Caire and S. Shamai (Shitz), "On the capacity of some channels with channel state information," *IEEE Trans. Information Theory*, vol. 45, pp. 2007–2019, Sept. 1999.
[8] A. Narula, M. Lopez, M. Trott, and G. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.
[9] A. Narula, M. Trott, and G. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418–2433, Nov. 1999.
[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Labs, Tech. Rep. BL0112170-950615-07TM, 1995.
[11] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
[12] F. Nesser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1293–1302, July 1993.
[13] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna methods for multipath environments," in *Proc. IEEE GLOBECOM'94*, Mar. 1994, pp. 425–429.
[14] D. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley.
[15] D. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

On the Capacity of a Pulse-Position-Hopped CDMA System

Ola Wintzell, *Student Member, IEEE*, Dmitri K. Zigangirov, and Kamil Sh. Zigangirov, *Fellow, IEEE*

***Abstract*—Pulse-Position-Hopped (PPH) code division multiple access (CDMA) is a new promising multiple-access technique which is very well suited for short-range multipath communications and has several benefits, such as coherent reception and low transmit power density. In this correspondence, we analyze the error-correcting capability of a system employing PPH-CDMA. The results show that the system capacity is proportional to the bandwidth, in a similar fashion as for carrier-based transmission techniques.**

***Index Terms*—Impulse radio, time hopping, ultra-wide bandwidth (UWB).**

I. Introduction

The current emphasis on constant-envelope spread-spectrum modulations has caused engineers to ignore one design, which has considerable potential, namely pulse-position hopping (PPH), also known as time hopping. Under the names impulse radio multiple access (IRMA) and ultra-wide bandwidth (UWB) transmission, this modulation is proposed and analyzed in [1]–[5]. PPH transmission has several benefits, such as coherent reception and low transmit power density.

In [1], the basics of the technology for generation of the narrow pulses of duration less than 1 ns and the very low spectral density is thoroughly described. The study of the capacity of a binary pulse position modulation (PPM) IRMA system [3] shows that it can reach an order of several thousands of active users per cell. In [6], an experimental design is described and measurements of the multipath channel are presented. There are several patents covering receiver structures, see for instance [4], and systems, see for instance [5]. We will study a slightly different modulation method in comparison to [3], namely, binary on–off modulation. In this correspondence, we will present an information-theoretical analysis of a PPH code-division multiple-access (CDMA) system and will present a lower bound to the overall effective capacity of the system in the downlink and the uplink directions. Our approach follows classical information-theoretical analysis methods [7]; similar methods were applied for analysis of direct-sequence CDMA (DS-CDMA) and frequency-hopping CDMA (FH-CDMA) systems in [8].

The remaining part of this correspondence is organized as follows. In Section II, the system model is described, in Sections III and IV, the effective capacity of the uplink and downlink system is estimated, and in Section V conclusions and future work are discussed.

Manuscript received October 30, 1999; revised November 16, 2000. This work was supported in part by the Swedish Foundation for Strategic Research—Personal Computing and Communication (PCC), Ericsson Mobile Communications, and by the Royal Swedish Academy of Science in cooperation with the Russian Academy of Science. The material in this correspondence was presented in part at the International Symposium on Information Theory, Sorrento, Italy, June 2000.

O. Wintzell and K. Sh. Zigangirov are with the Department of Information Technology, Lund University, 221 00 Lund, Sweden (e-mail: ola.wintzell@it.lth.se; kamil.zigangirov@it.lth.se).

D. K. Zigangirov is with the Institute for Problems of Information Transmission, Moscow, 101447 Russia (e-mail: zig@iitp.ru).

Communicated by M. L. Honig, Associate Editor for Communications.
Publisher Item Identifier S 0018-9448(01)05454-2.