

# An Adaptive Multiantenna Transceiver for Slowly Flat Fading Channels

Ada S. Y. Poon, *Student Member, IEEE*, David N. C. Tse, *Member, IEEE*, and Robert W. Brodersen, *Fellow, IEEE*

**Abstract**—This paper proposes an adaptive multiantenna transceiver for narrowband reception. Blind channel tracking algorithms are developed to track the eigen directions of the channel directly instead of the channel itself. Two algorithms are proposed to track the column space of the channel at the receiver, based on the received data. One of the algorithms is free of any division operation, which is more favorable in practice. For the row space of the channel, two approaches are proposed as well. The first approach requires periodic feedback of the demodulated signal from the receiver back to the transmitter where it can make use of its knowledge on the prior transmitted symbols to estimate the row space. In the second approach, the estimation is done at the receiver based on the detected symbols, and the estimated row space is sending back to the transmitter. Adaptive resource allocation is also incorporated into the design.

**Index Terms**—Adaptive array processing, blind channel tracking, multielement arrays, transceiver architecture.

## I. INTRODUCTION

THE demand for high-quality wireless communication services is increasing at an explosive rate. However, the inherent limited supply of bandwidth and the unpredictability of the propagation channel lead to the investigation of more spectrally efficient and reliable designs. The use of antenna arrays is expected to play an important role in fulfilling these requirements. In the past, antenna arrays were used to increase receive or transmit diversity against multipath fading [1]–[3] or spatially separate mobile devices [4]. Recently, it was shown that in a richly scattering channel, signals even from closely spaced transmit antennas can be separated using adaptive array combining techniques at the receiver. The number of transmit signals that can be separated grows with the number of receive antennas [5], [6]. Consequently, the use of multiple antennas at both transmitter and receiver could increase the degrees of freedom (DOFs) in information transmission, with capacity scaling linearly with the number of transmit or receive antennas, whichever is smaller.

In many systems such as V-BLAST [7], no knowledge of the current channel state is assumed at the transmitter. In a fast fading channel, this lack of channel information at the transmitter can be coped with by interleaving the data in such a

Paper approved by M. Fitz, the Editor for Modulation and Diversity of the IEEE Communications Society. Manuscript received June 17, 2000; revised April 15, 2002. This work was supported by DARPA and the industrial members of the Berkeley Wireless Research Center.

The authors are with the Berkeley Wireless Research Center, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94704 USA (e-mail: sypoon@eecs.berkeley.edu; dtse@eecs.berkeley.edu; rb@eecs.berkeley.edu).

Digital Object Identifier 10.1109/TCOMM.2003.819206

way that the channel effect is averaged out. However, the use of interleaving becomes impossible in a slowly fading channel, such as in an indoor environment. In this case, data have to be transmitted at a conservative rate to keep the outage probability small. If, on the other hand, channel information is known at the transmitter, performance can be improved by adapting the transmission rate according to the quality of the channel. Through adaptation, a higher average throughput is attained. Furthermore, it is more robust and reliable as compared with schemes relying on channel statistics alone [7].

In this paper, we focus on the slow, flat fading scenario and propose an adaptive multiantenna transceiver system. The transceiver architecture is inspired by the result from information theory on the capacity of the multiantenna Gaussian channel [8]. Under flat fading, this capacity is achieved by decomposing the system into parallel subchannels through singular value decomposition (SVD), with the transmitter sending independent data streams across these subchannels. This idea of decomposition carries over naturally to our suggested transceiver architecture, and the problem of tracking the temporal variation of the propagation channel is reduced to tracking the singular values and eigenvectors at the receiver and the transmitter *collaboratively*.

We develop new efficient blind algorithms to perform such channel tracking. At the receiver, since the transmitted signal is transformed by the channel, the received signal contains information about the channel effect. Two algorithms are proposed to extract this partial channel information at the receiver without the knowledge of the transmitted signal. One is division-free with complexity of  $\mathcal{O}(r^2d)$ , while the other needs division in each iteration but of complexity  $\mathcal{O}(rd)$ , where  $r$  is the number of receive antennas and  $d$  is the number of trackable subchannels (which will be defined later.) The proposition toward a division-free algorithm is due to the fact that in practice, dividers are more unstable and suffer longer delay than multipliers [9]. The residual channel information is then estimated at the transmitter, based on the periodic feedback of the demodulated signal from the receiver and knowledge of its prior transmitted signals. A least-mean square (LMS) type of algorithm is proposed to fulfill this with complexity of  $\mathcal{O}(td)$  where  $t$  is the number of transmit antennas. However, continually communicating back the demodulated signal to the transmitter is inefficient. This could be remedied by switching the receiver to a decision-directed mode of operation. The ratio of feedback rate to data rate in both nondecision-directed and decision-directed approaches is compared through Monte-Carlo simulations.

Raleigh *et al.* [10] have suggested a similar architecture for approaching channel capacity of a multi-input multi-output (MIMO) system. In their transceiver architecture, the channel

estimation and tracking are achieved by sending training sequences to the receiver, decomposing the tracked system transfer function into its singular components at the receiver in batch, and communicating back the necessary singular components to the transmitter. This approach is similar to the bulk of research in multiantenna channel estimation and tracking algorithms, which favors using training sequences. In contrast, our work emphasizes blind adaptive SVD tracking algorithms which can be implemented efficiently.

The rest of the paper is organized as follows. In Section II, we introduce the system model and review results from information theory. In Section III, we propose the transceiver architecture, while in Section IV, we present the signal subspace interpretation and the blind tracking algorithms. Then, we include the simulation results in Section V.

In this paper, the following notations will be used. We will use boldfaced capital letters to indicate matrices and boldfaced small letters for vectors.  $\mathbf{I}$  is the identity matrix.  $(\cdot)^\dagger$ ,  $(\cdot)^+$ , and  $E(\cdot)$  denote complex conjugate transpose, taking the positive part and taking the expectation, respectively.  $\mathcal{C}^n$  denotes the set of  $n$ -dimensional complex numbers, whereas  $\mathcal{C}^{n \times m}$  denotes the set of  $n \times m$  complex matrices.

## II. SYSTEM MODEL AND CAPACITIES

Consider a system with  $t$  transmit and  $r$  receive antennas in a frequency nonselective, slowly fading channel. The sampled baseband-equivalent channel model is given by

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

where  $\mathbf{H} \in \mathcal{C}^{r \times t}$  is the complex channel matrix with the  $(i, j)$ th element being the random fading between the  $i$ th receive and  $j$ th transmit antennas.  $\mathbf{z} \in \mathcal{C}^r$  is the additive noise source and is modeled as a zero-mean, circularly symmetric, complex Gaussian random vector with statistically independent elements, that is,  $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_z^2 \mathbf{I}_r)$ . The  $i$ th element of  $\mathbf{x} \in \mathcal{C}^t$  is the symbol transmitted at the  $i$ th transmit antenna, and that of  $\mathbf{y} \in \mathcal{C}^r$  is the symbol received at the  $i$ th received antenna. In the following exposition, we will constrain the total average transmit power to a constant  $P$ .

Substituting the SVD of the matrix  $\mathbf{H}$ , (1) becomes

$$\mathbf{y} = \mathbf{U}\Sigma\mathbf{V}^\dagger\mathbf{x} + \mathbf{z} \quad (2)$$

where  $\mathbf{U}$  is an  $r \times r$  unitary matrix,  $\mathbf{V}$  is an  $t \times t$  unitary matrix, and  $\Sigma$  is a  $r \times t$  matrix with only nonzero main diagonal entries being the nonnegative square root of the eigenvalues of  $\mathbf{H}\mathbf{H}^\dagger$ . Multiplying  $\mathbf{U}^\dagger$  on both sides, (1) can be rewritten as

$$\mathbf{y}' = \Sigma\mathbf{x}' + \mathbf{z}' \quad (3)$$

where  $\mathbf{x}' = \mathbf{V}^\dagger\mathbf{x}$ ,  $\mathbf{y}' = \mathbf{U}^\dagger\mathbf{y}$ , and  $\mathbf{z}' = \mathbf{U}^\dagger\mathbf{z}$ . Noticing that the distribution of  $\mathbf{z}'$  is invariant under unitary transformation, the multiantenna channel is equivalent to  $\min(r, t)$  independent parallel Gaussian subchannels. Each subchannel has gain being the singular value of the fading matrix  $\mathbf{H}$ .

For each realization of  $\mathbf{H}$ , the mutual information is given by

$$I(\mathbf{x}; \mathbf{y}) = \log_2 \det \left( \mathbf{I}_r + \frac{1}{\sigma_z^2} \mathbf{H}\mathbf{K}_x\mathbf{H}^\dagger \right) \quad (4)$$

where  $\mathbf{K}_x$  is the autocorrelation matrix of  $\mathbf{x}$ . Since the mutual information is maximized when the transmitted signals are uncorrelated, so we will assume that the signals sent to the transmitting array are uncorrelated, and let  $P_i$  denote the power allocated to the  $i$ th transmit antenna. Then (4) can be rewritten as

$$I(\mathbf{x}; \mathbf{y}) = \sum_{i=1}^{\min(t, r)} \log_2 \left( 1 + \frac{P_i}{\sigma_i^2} \right) \quad (5)$$

where  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{\min(t, r)})$ .

When both transmitter and receiver know the fading matrix  $\mathbf{H}$ , the transmitter can adapt the rate of information transmission and power allocated across the subchannels according to the quality of the channel. The optimal power allocation strategy is the well-known waterfilling solution and is given by

$$P_i = \left( v - \frac{\sigma_z^2}{\sigma_i^2} \right)^+ \quad (6)$$

where  $v$  satisfies

$$\sum_i \left( v - \frac{\sigma_z^2}{\sigma_i^2} \right)^+ = P. \quad (7)$$

A suboptimal solution could be allocating equal power across the subchannels, that is

$$P_i = \frac{P}{d} \quad (8)$$

where  $d$  denotes the number of nonzero channel eigenmodes. At high signal-to-noise ratios (SNRs), that is, the DOF-limited regime, the mutual information is insensitive to the power allocation strategies used. It is the number of usable subchannels and the identification of them that matter. For simplicity, we will adopt equal power allocation scheme, and focus exclusively on the tracking and identification of the usable subchannels.

## III. TRANSCEIVER ARCHITECTURE

Motivated by the improved performance through adaptive resource allocation in a multiantenna channel, we design an adaptive multiantenna transceiver with periodic feedback from the receiver to the transmitter. Since the performance enhancement of a multiantenna system depends heavily on the accuracy of the channel measurements, we would like to emphasize the acquisition of the current channel conditions in our design.

The SVD in (2) suggests that the transmitter should prefilter the modulated symbols by the right singular matrix  $\mathbf{V}$  to match the channel, and at the same time, the receiver should filter the received signals by the left singular matrix  $\mathbf{U}$  before demodulation and decoding. Through these filtering operations, the channel is naturally decomposed into  $\min(r, t)$  uncorrelated subchannels. To see this, refer to the diagram in Fig. 1. The channel acquisition problem is reduced to the estimation of  $\mathbf{V}$  at the transmitter, and the estimation of  $\Sigma$  and  $\mathbf{U}$  at the receiver.

Under a noiseless condition, the received signals are in the space spanned by the columns of the left singular matrix  $\mathbf{U}$ . As long as there is a sizable number of received symbols within a fraction of the coherence time (period of time in which the temporal variation of the fading coefficients  $H_{ij}$ 's is insignificant), the receiver is able to estimate  $\mathbf{U}$  and  $\Sigma$  from the received data alone. This reduces the frequency in sending pilot symbols, and

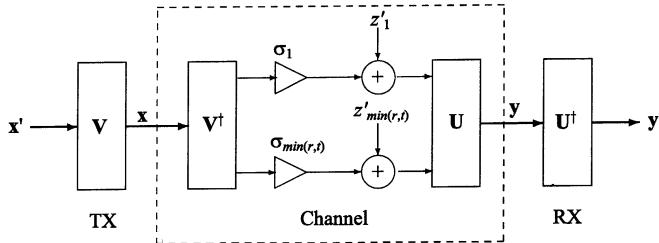


Fig. 1. Diagram illustrating the decoupling of the channel.

is more spectrally efficient. In an indoor environment with pedestrian speed of 2 m/s and transmitting at 5.8 GHz, the maximum Doppler frequency is around 40 Hz, and the coherence time is approximately equal to 0.01 s [11]. Furthermore, the delay spread is usually on the order of 10–100 ns, and the coherence bandwidth [12] is roughly on the order of 2–20 MHz. In order to experience flat fading, the transmitted signals should have symbol rate per transmit antenna of no greater than the coherence bandwidth. Therefore, for a symbol rate of 1 MSymbol/s, the number of symbols in a period of coherence time is approximately  $10^4$ , which is sufficient to get a good estimate on  $\mathbf{U}$  and  $\Sigma$ .

With the presence of additive noise, the space spanned by the received signals can be decomposed into the signal subspace and the noise subspace. The signal subspace contains information about the channel effect on the transmitted symbol that could be tracked at the receiver. The noise subspace is the space orthogonal to the signal subspace and is not very useful in the tracking process, except in the estimation of the additive noise power.

The receiver first exploits the structure of the received signal and partitions the received signal space into the signal and noise subspaces. This partitioning reduces the dimension of the tracking space and saves computational resources. Afterwards, it will track the temporal variation of  $\mathbf{U}$  and  $\Sigma$  corresponding to the signal space by the blind algorithms detailed in Section IV. Simultaneously, it will project the received signal on the estimated  $\mathbf{U}$  and demodulate the projected symbols. Based on the estimated  $\Sigma$ , it will decide the rate of information transmission on each of the useful subchannels and inform the transmitter from time to time. In order to assist the transmitter to track the temporal variation of the right singular matrix  $\mathbf{V}$ , the receiver will send back some of the demodulated symbols to the transmitter periodically, as well.

The transmitter, based on knowledge of the prior transmitted symbols and the feedback information from the receiver, tracks the temporal variation of  $\mathbf{V}$  by an adaptive minimum mean-square error (MMSE) algorithm that will be elaborated in Section IV. Then it will allocate the transmitted data across the subchannels, according to the rate allocation strategy decided by the receiver. Afterwards, it will project the transmitted data on the estimated  $\mathbf{V}$  and send it to the radio frequency (RF) section.

In addition, we address the problem of initialization by sending training sequences at the very beginning and use the proposed tracking algorithms for channel estimation as well with no extra cost. Furthermore, we would reduce the frequency of feedback by switching the receiver to a decision-directed mode of operation, that is, the receiver will use the detected symbols

and the received signals after projection to track  $\mathbf{V}$ . Then it will send the estimated  $\mathbf{V}$  back to the transmitter periodically. Fig. 2 shows the receiver and the transmitter architectures.

#### IV. SIGNAL SUBSPACES PARTITION AND BLIND TRACKING ALGORITHMS

##### A. Signal Subspaces Interpretation

Recall that the system model is

$$\mathbf{y} = \mathbf{U}\Sigma\mathbf{V}^\dagger\mathbf{x} + \mathbf{z} = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \Sigma_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^\dagger \\ \mathbf{V}_n^\dagger \end{bmatrix} \mathbf{x} + \mathbf{z} \quad (9)$$

where the columns of  $\mathbf{U}_s \in \mathcal{C}^{r \times d}$  and  $\mathbf{V}_s \in \mathcal{C}^{t \times d}$  correspond to the direction of the useful eigen modes of the channel. The autocorrelation matrix of the received signal  $\mathbf{y}$  is given by

$$\mathbf{K}_y = E[\mathbf{y}\mathbf{y}^\dagger] = \mathbf{H}\mathbf{K}_x\mathbf{H}^\dagger + \sigma_z^2\mathbf{I}. \quad (10)$$

Given  $\mathbf{V}_s^\dagger\mathbf{K}_x\mathbf{V}_s = P/d\mathbf{I}$ , (10) can be rewritten as

$$\mathbf{K}_y = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \frac{P}{d}\Sigma_s^2 + \sigma_z^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_z^2\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^\dagger \\ \mathbf{U}_n^\dagger \end{bmatrix}. \quad (11)$$

Hence, the space spanned by the columns of  $\mathbf{U}_s$  is termed as the signal subspace, and that by  $\mathbf{U}_n$  is the noise subspace.

The question now is how to estimate  $\mathbf{U}_s$ ,  $\Sigma_s$ , and  $\mathbf{V}_s$  efficiently. Before going into the details of the tracking algorithms used, we first present our mathematical setup and several useful results.

At time  $nT$  (where  $T$  is the sample period), the received signal vector is

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{z}(n).$$

Collect the first  $n$  snapshots of  $\mathbf{y}(i)$ 's,  $\mathbf{x}(i)$ 's, and  $\mathbf{z}(i)$ 's, and form the data matrix as follows:

$$\mathbf{Y} = \frac{1}{\sqrt{n}} [\mathbf{y}(1) \quad \cdots \quad \mathbf{y}(n)] \quad (12a)$$

$$\mathbf{X} = \frac{1}{\sqrt{n}} [\mathbf{x}(1) \quad \cdots \quad \mathbf{x}(n)] \quad (12b)$$

$$\mathbf{Z} = \frac{1}{\sqrt{n}} [\mathbf{z}(1) \quad \cdots \quad \mathbf{z}(n)]. \quad (12c)$$

They are related by

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z}$$

where  $\mathbf{Y}$  is the data matrix and  $\mathbf{Y}\mathbf{Y}^\dagger$  is the sample autocorrelation matrix, denoted by  $\hat{\mathbf{K}}_y$ . Thus, the SVD of  $\mathbf{Y}$  is

$$\mathbf{Y} = \hat{\mathbf{U}}\hat{\Sigma}\hat{\mathbf{V}}^\dagger = [\hat{\mathbf{U}}_s \quad \hat{\mathbf{U}}_n] \begin{bmatrix} \hat{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \hat{\Sigma}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{V}}_s^\dagger \\ \hat{\mathbf{V}}_n^\dagger \end{bmatrix} \quad (13)$$

whereas the eigenvalue decomposition of  $\hat{\mathbf{K}}_y$  is

$$\hat{\mathbf{K}}_y = \hat{\mathbf{U}}\hat{\Sigma}^2\hat{\mathbf{U}}^\dagger = [\hat{\mathbf{U}}_s \quad \hat{\mathbf{U}}_n] \begin{bmatrix} \hat{\Sigma}_s^2 & \mathbf{0} \\ \mathbf{0} & \hat{\Sigma}_n^2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^\dagger \\ \hat{\mathbf{U}}_n^\dagger \end{bmatrix}. \quad (14)$$

By the theory on the perturbation of eigenvalues [1]

$$\hat{\Sigma}_s^2 - \sigma_z^2\mathbf{I} \longrightarrow \frac{P}{d}\Sigma_s^2 \quad \text{a.s.}$$

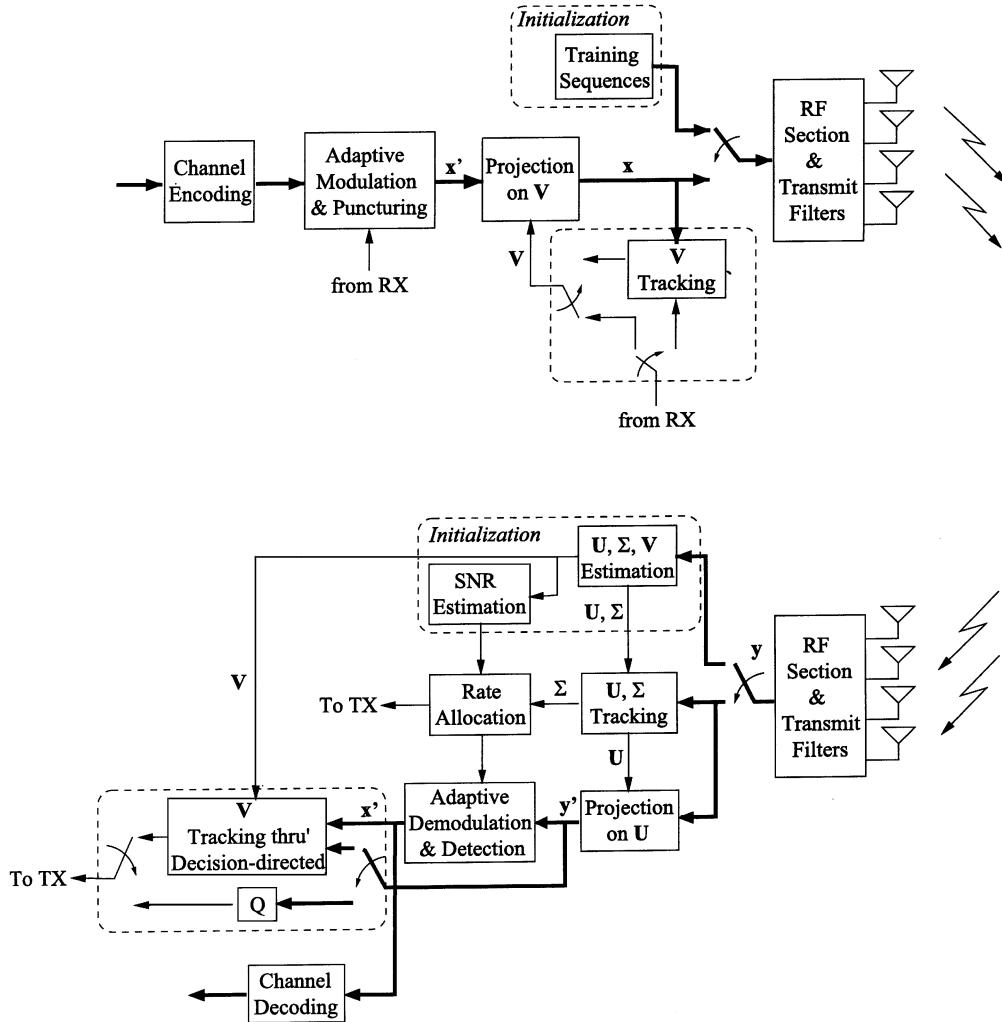


Fig. 2. Block diagram of the multiantenna transceiver.

and by the perturbation theory for invariant subspaces [14], [15]

$$\mathcal{R}(\hat{\mathbf{U}}_s) \longrightarrow \mathcal{R}(\mathbf{U}_s) \quad \text{a.s.}$$

where  $\mathcal{R}(\cdot)$  is the range space.

Consequently,  $\mathbf{U}_s$  and  $\Sigma_s$  can be estimated either by direct eigenvalue decomposition of the sample autocorrelation matrix, or by direct singular value decomposition of the data matrix. But notice that from sample to sample, the autocorrelation and data matrices vary only by a rank-one modification. Instead of applying batch decomposition in every shot, the computation can be considerably reduced by iteratively updating  $\mathbf{U}_s$  and  $\Sigma_s$ .

The sample autocorrelation matrix derived from (12) averages over all available samples. However, in practice, the fading matrix  $\mathbf{H}$  as well as  $\mathbf{U}$  and  $\Sigma$  are slowly time varying. Thus, a moving average estimate from the  $N$  most recent received signal vectors is a better alternative

$$\begin{aligned} \hat{\mathbf{K}}_y(n) &= \frac{1}{N} \sum_{i=n-N+1}^n \mathbf{y}(i)\mathbf{y}(i)^\dagger \\ &= \hat{\mathbf{K}}_y(n-1) \\ &\quad + \frac{1}{N} [\mathbf{y}(n)\mathbf{y}(n)^\dagger - \mathbf{y}(n-N)\mathbf{y}(n-N)^\dagger]. \end{aligned} \quad (15)$$

If  $\mathbf{H}$  is stationary, the moving average estimate is the maximum-likelihood estimate of  $\mathbf{K}_y$ , and if we replace  $N$  by  $N-1$ , then it is the unbiased estimate. To reduce the memory requirement for the old received signals, weighted average estimate is another possibility, which yields

$$\hat{\mathbf{K}}_y(n) = (1-\beta)\hat{\mathbf{K}}_y(n-1) + \beta\mathbf{y}(n)\mathbf{y}(n)^\dagger \quad (16)$$

where  $\beta$  is the forgetting factor and its choice depends on the degree of stationarity of the channel. Similarly, the data matrix shown in (12) may not be practical as well, since its dimension grows with the data length. Same as before, we can fix the number of samples taken into account by deleting a column and appending a new one in every update.

Unlike its counterparts, there is no intimate connection between  $\mathbf{V}_s$  and  $\hat{\mathbf{V}}_s$ . As seen from (9) and (11), it is impossible to estimate  $\mathbf{V}_s$  based on the first- and second-order statistics of the received signals alone. As a result, in tracking the variation of  $\mathbf{V}_s$ , we rely on the knowledge of prior transmitted symbols at the transmitter and propose an LMS type of algorithm, detailed in the following.

### B. Tracking Algorithms for $\mathbf{U}$ and $\Sigma$

In this paper, we develop two algorithms to track the temporal variation of  $\mathbf{U}_s$  and  $\Sigma_s$ . The first one is free of division and requires  $O(r^2d)$  operations every update, while the second one requires  $O(rd)$  operations, where  $r$  is the number of receive antennas and  $d$  is the number of useful subchannels. Basically, we introduce an unconstrained optimization problem with the eigencomponents being the stable stationary points. Then we apply gradient-descent methods and deflation techniques (sequential estimation of the eigencomponents) to arrive at the desired solution. The following theorem shows the objective function that we are using in the algorithms.

*Theorem 1:* Let  $\mathbf{A} \in \mathcal{C}^{n \times n}$  be a semipositive definite matrix. Define

$$J(\mathbf{w}) = 2\mathbf{w}^\dagger \mathbf{A} \mathbf{w} - (\mathbf{w}^\dagger \mathbf{w})^2. \quad (17)$$

All the stationary points of  $J(\mathbf{w})$  are eigenvectors of  $\mathbf{A}$  with magnitude being the square root of the corresponding eigenvalue of  $\mathbf{A}$ . Further, if the dominant eigen pair is of multiplicity one, then the Hessian of  $J(\mathbf{w})$  at the dominant eigenvector is negative definite, while that at the remaining eigenvectors are neither positive definite nor negative definite, that is, the dominant eigen pair is the global maximum point of  $J(\mathbf{w})$ .

*Proof:* First, take the gradient on  $J(\mathbf{w})$ .

$$\nabla_{\mathbf{w}} J(\mathbf{w}) = 4\mathbf{A}\mathbf{w} - 4(\mathbf{w}^\dagger \mathbf{w})\mathbf{w}. \quad (18)$$

Putting to zero, gives

$$\mathbf{A}\mathbf{w} = (\mathbf{w}^\dagger \mathbf{w})\mathbf{w}.$$

Therefore, the stationary points of  $J(\mathbf{w})$  are eigenvectors of  $\mathbf{A}$  with magnitude being the square root of the corresponding eigenvalue of  $\mathbf{A}$ .

Next, we prove that the dominant eigenvector is the only stable stationary point. Let  $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^\dagger$  be the eigenvalue decomposition of  $\mathbf{A}$ . Then the Hessian of  $J(\mathbf{w})$  at  $\mathbf{w}$  is

$$\nabla_{\mathbf{w}}^2 J(\mathbf{w}) = 4\mathbf{A} - 4(\mathbf{w}^\dagger \mathbf{w})\mathbf{I} - 8\mathbf{w}\mathbf{w}^\dagger. \quad (19)$$

Substitute the dominant eigen pair into (19), and we get

$$\begin{aligned} \nabla_{\mathbf{w}}^2 J(\sqrt{\lambda_1} \mathbf{u}_1) &= 4\mathbf{U}\Lambda\mathbf{U}^\dagger - 4\lambda_1 \mathbf{U}\mathbf{U}^\dagger - 8\lambda_1 \mathbf{u}_1 \mathbf{u}_1^\dagger \\ &= 4\mathbf{U} \left( \Lambda - \lambda_1 \mathbf{I} - 2 \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \right) \mathbf{U}^\dagger \\ &\quad \times \mathbf{U}^\dagger \\ &= -4\mathbf{U} \begin{bmatrix} 2\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_1 - \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_1 - \lambda_n \end{bmatrix} \mathbf{U}^\dagger \\ &\prec 0 \end{aligned}$$

TABLE I  
DIVISION-FREE TRACKING ALGORITHM FOR  $\mathbf{U}$  AND  $\Sigma$

---

Receive $\mathbf{y}(n+1)$
Update covariance matrix
$\hat{\mathbf{K}}_y(n+1) = (1 - \beta)\hat{\mathbf{K}}_y(n) + \beta \mathbf{y}(n+1) \mathbf{y}(n+1)^\dagger$
$\mathbf{R} = \hat{\mathbf{K}}_y(n+1)$
for $i = 1 : d$ ,
Update the $i$ -th eigen-pairs
$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu_i(\mathbf{R} - \lambda_i(n) \mathbf{I}) \mathbf{w}_i(n)$
$\lambda_i(n+1) = \mathbf{w}_i(n+1)^\dagger \mathbf{w}_i(n+1)$
Apply deflation
$\mathbf{R} = \mathbf{R} - \mathbf{w}_i(n+1) \mathbf{w}_i(n+1)^\dagger$
end

---

which is negative definite. Substituting the other eigen pairs into (19), we have

$$\begin{aligned} \nabla_{\mathbf{w}}^2 J(\sqrt{\lambda_i} \mathbf{u}_i) &= -4\mathbf{U} \begin{bmatrix} \lambda_i - \lambda_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \lambda_i - \lambda_{i-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \lambda_i - \lambda_n \end{bmatrix} \mathbf{U}^\dagger \end{aligned}$$

which is neither positive definite nor negative definite  $\forall i \neq 1$ . ■

Most of the objective functions in the literature require either normalization or division. One such objective function is the *Rayleigh-Ritz ratio* [16, p. 176]. From the perspective of implementation, both normalization and division are costly and slow. On the other hand, the objective function introduced in (17) is free of any division and normalization.

To find the optimal solution to  $\max_{\mathbf{w}} J(\mathbf{w})$ , it is straightforward to apply gradient descent techniques. The gradient of  $J(\mathbf{w})$  is given in (18), which results in the update formula as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [\mathbf{A} - \mathbf{w}(n) * \mathbf{w}(n)^\dagger] \mathbf{w}(n) \quad (20)$$

where  $\mu$  is the step size.

If  $\mathbf{w}^*$  is the optimal solution to  $\max_{\mathbf{w}} J(\mathbf{w})$ , then the corresponding eigenvalue is given by  $\mathbf{w}^{*\dagger} \mathbf{w}^*$ . Now consider

$$\begin{aligned} \mathbf{A}_\alpha &= \mathbf{A} - \mathbf{w}^* \mathbf{w}^{*\dagger} \\ J_\alpha(\mathbf{w}) &= 2\mathbf{w}^\dagger \mathbf{A}_\alpha \mathbf{w} - (\mathbf{w}^\dagger \mathbf{w})^2. \end{aligned}$$

It follows that the second dominant eigen pair of  $\mathbf{A}$  is the global maximum of  $J_\alpha(\mathbf{w})$ . As the cancellation process proceeds, all the eigen pairs corresponding to nonzero eigenvalues can be obtained. This sequential cancellation technique is called *deflation*.

Recasting to our problem, we are looking for the eigen pairs of the autocorrelation matrix  $\mathbf{K}_y$ , which is estimated by  $\hat{\mathbf{K}}_y$ . Since  $\hat{\mathbf{K}}_y$  is semipositive definite, we can simply replace matrix  $\mathbf{A}$  in the update formula shown in (20) by matrix  $\hat{\mathbf{K}}_y$ . Table I details the tracking algorithm for  $\mathbf{U}_s$  and  $\Sigma_s$ .

TABLE II  
DIVISION-ADHERED TRACKING ALGORITHM FOR  $\mathbf{U}$  AND  $\Sigma$

---

Receive $\mathbf{y}(n+1)$
$\mathbf{r} = \mathbf{y}(n+1)$
for $i = 1 : d$ ,
Update the $i$ -th eigen-pairs
$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu_i(\mathbf{r} \mathbf{r}^\dagger \mathbf{w}_i(n) - \lambda_i(n) \mathbf{w}_i(n))$
$\lambda_i(n+1) = \mathbf{w}_i(n+1)^\dagger \mathbf{w}_i(n+1)$
Apply deflation
$\mathbf{r} = \mathbf{r} - (\mathbf{w}_i(n+1)^\dagger \mathbf{r}) \mathbf{w}_i(n+1) / \lambda_i(n+1)$
end

---

The update formula can be further simplified by approximating the sample autocorrelation matrix using the instantaneous received signals only, that is, setting the forgetting factor  $\beta$  to one. This is the LMS algorithm in adaptive filtering. In tracking the dominant eigen pair at time  $n$ , we apply the same algorithm as above on the received signal  $\mathbf{y}(n)$  only. Then removing the projection of  $\mathbf{y}(n)$  on the direction of the tracked dominant eigenvector, we apply the algorithm on the residue again and track the second dominant eigen pair. The cancellation and tracking processes proceed and all the signal eigen pairs can then be tracked sequentially. Table II details the simplified tracking algorithm for  $\mathbf{U}_s$  and  $\Sigma_s$ . In this algorithm, simplicity is achieved at the expense of introducing division operation in each iteration.

The choice of the step size in (20) directly impacts the convergence speed, stability, and accuracy of the algorithms. Since the objective function introduced  $J(\mathbf{w})$  is a fourth-order function in  $\mathbf{w}$ , analysis on the choice of step size is difficult. In particular, the use of deflation techniques complicates the analysis a lot. Instead, we will give loose bounds on the step size by approximating  $J(\mathbf{w})$  by a quadratic function around the optimal point. We believe that these bounds give useful guidelines in the tracking period, where the estimates are close to the optimal.

Applying a Taylor series expansion to  $J(\mathbf{w})$ , we get

$$\begin{aligned} J(\mathbf{w}) &= J(\sqrt{\lambda_1} \mathbf{u}_1) \\ &+ \frac{1}{2} (\mathbf{w} - \sqrt{\lambda_1} \mathbf{u}_1)^\dagger \nabla_{\mathbf{w}}^2 J(\sqrt{\lambda_1} \mathbf{u}_1) (\mathbf{w} - \sqrt{\lambda_1} \mathbf{u}_1) + \dots \end{aligned}$$

For  $\mathbf{w}$  around the optimal point  $\sqrt{\lambda_1} \mathbf{u}_1$ ,  $J(\mathbf{w})$  can be approximated by

$$\begin{aligned} J(\mathbf{w}) &\approx J(\sqrt{\lambda_1} \mathbf{u}_1) \\ &+ \frac{1}{2} (\mathbf{w} - \sqrt{\lambda_1} \mathbf{u}_1)^\dagger \nabla_{\mathbf{w}}^2 J(\sqrt{\lambda_1} \mathbf{u}_1) (\mathbf{w} - \sqrt{\lambda_1} \mathbf{u}_1). \end{aligned}$$

Therefore, the step size is loosely bounded by

$$0 < \mu < \frac{2}{\frac{1}{2} \lambda_{\max} (-\nabla_{\mathbf{w}}^2 J(\sqrt{\lambda_1} \mathbf{u}_1))} = \frac{1}{2\lambda_1}. \quad (21)$$

TABLE III  
TRACKING ALGORITHM FOR  $\mathbf{V}$

---

Receive $\mathbf{y}(n+1)$
for $i = 1 : d$ ,
Update the $i$ -th eigenvector
$\mathbf{v}_i(n+1) = \mathbf{v}_i(n) + \mu_i (\mathbf{y}_i(n+1) - \mathbf{v}_i(n)^\dagger \mathbf{x}(n+1))^\dagger \mathbf{x}(n+1)$
end

---

### C. Tracking Algorithm for $\mathbf{V}$

First, we rewrite (2) as

$$\Sigma_s^{-1} \mathbf{U}_s^\dagger \mathbf{y} = \mathbf{V}_s^\dagger \mathbf{x} + \Sigma_s^{-1} \mathbf{U}_s^\dagger \mathbf{z}. \quad (22)$$

Now, the receiver can send the signal

$$\mathbf{y}_f = \hat{\Sigma}_s^{-1} \hat{\mathbf{U}}_s^\dagger \mathbf{y} \quad (23)$$

back to the transmitter where it can make use of its knowledge on the prior transmitted symbols. Assuming error free in the feedback channel, we choose to minimize the MSE of

$$\mathbf{y}_f - \mathbf{V}_s^\dagger \mathbf{x}$$

to track the temporal variation of  $\mathbf{V}_s$ . The tracking algorithm is summarized in Table III, in which the application of LMS algorithm in the MMSE criterion is straightforward. Furthermore, the algorithm has complexity of  $\mathcal{O}(td)$ .

In practice, we can reduce the frequency of feedback by applying the tracking algorithm on the detected symbols at the receiver instead of on the prior transmitted symbols at the transmitter.

### D. Unequal Power Allocation

Up to this point, the transceiver allocates equal power to the usable subchannels but varies the information rate across them. However, the tracking algorithms introduced can be adapted to unequal power allocation. Suppose the power allocated to the  $i$ -th subchannel is  $P_i$ , for  $i = 1, \dots, d$ . Then the autocorrelation matrix of the received signal  $\mathbf{y}$  in (11) becomes

$$\mathbf{K}_y = \frac{P}{d} = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \Sigma_x^2 \Sigma_s^2 + \sigma_z^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_z^2 \mathbf{I} \end{bmatrix} [\mathbf{U}_s^\dagger \quad \mathbf{U}_n^\dagger] \quad (24)$$

where  $\Sigma_x^2 = \text{diag}(P_1, \dots, P_d)$ . The algorithms presented in Tables I and II will track the variation of the signal space  $\mathbf{U}_s$  as before. For the singular values, in contrast, they will track  $\Sigma_x^2 \Sigma_s^2$  instead of  $P/d \Sigma_s^2$ . As a result, one more step is required to get the channel gains  $\Sigma_s^2$  from  $\Sigma_x^2 \Sigma_s^2$ . But this computation can be carried out with ease.

## V. SIMULATION RESULTS

Since a half-wavelength separation is sufficient for decorrelation among antenna elements in an indoor channel [17], a carrier frequency of 5.8 GHz (the ISM band) requires only a minimum

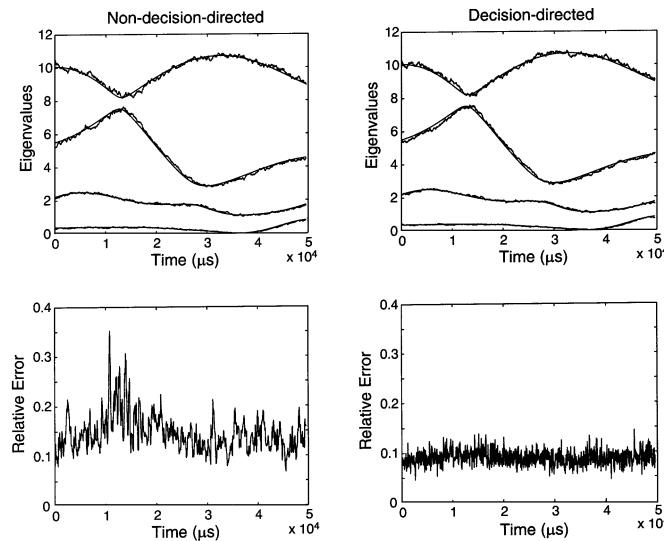


Fig. 3. (Upper) Tracked and actual eigenvalues under SNR of 10 dB (the first curve from the top shows the variation of the dominant eigenvalue, the second one shows that of the second dominant eigenvalue, and so on). (Lower) Relative error in the tracked  $H$ .

separation of 2.6 cm. Hence, for a mobile device operating in the ISM band, four antennas may be a reasonable choice. On the other hand, with the advent of more sophisticated complementary metal–oxide–semiconductor (CMOS) technology, it is possible to integrate 10–15 analog front ends on a single chip together with the digital back ends, though the antenna elements may be correlated. In the following simulations, we will use a four transmit, four receive antennas system under independent Rayleigh fading. The symbol rate is 1 Msymbols/s and the Doppler frequency is 10 Hz. The constellation scheme used is  $M$ -ary phase-shift keying (M-PSK), ranging from binary phase-shift keying (BPSK) to 16-PSK. To increase the resolution in the bit allocation, repetition codes are employed with coding gains ranging from 1/3 to 1/11. The channel model used is the popular Clarke and Jakes' model [17], [18]. First, we will report on the tracking performance of the division-free algorithm detailed in the Section IV. Then, we will show the probability of bit error versus the ratio of feedback rate to data rate in both nondecision-directed and decision-directed modes of operation. For the division-adhered algorithm, similar results were obtained so we will not repeat it.

Fig. 3 shows the tracking of the eigenvalues (square of the singular values) and the relative error (in the sense of Frobenius norm) in the tracked fading matrix. As seen from the graphs, both approaches can track the variation of the eigenvalues closely. Moreover, the relative error in the tracked fading matrix has a smaller range of fluctuation in the decision-directed approach than in the nondecision-directed approach.

Fig. 4 shows the angle between the tracked and actual dominant left and right singular vectors in the decision-directed approach. There are relatively large angles of discrepancies at around  $1.5 \times 10^4 \mu$ s, but the relative error in the tracked  $H$  matrix is not very significant. This is because at around  $1.5 \times 10^4 \mu$ s, the two most dominant eigenvalues are very close to each other, and hence, the main interest is the tracking of the subspaces spanned by the corresponding singular vectors rather

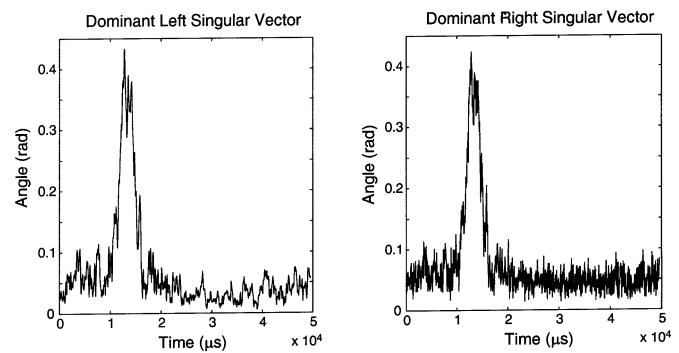


Fig. 4. Angle between dominant tracked and actual singular vectors under SNR 10 dB using the decision-directed approach.

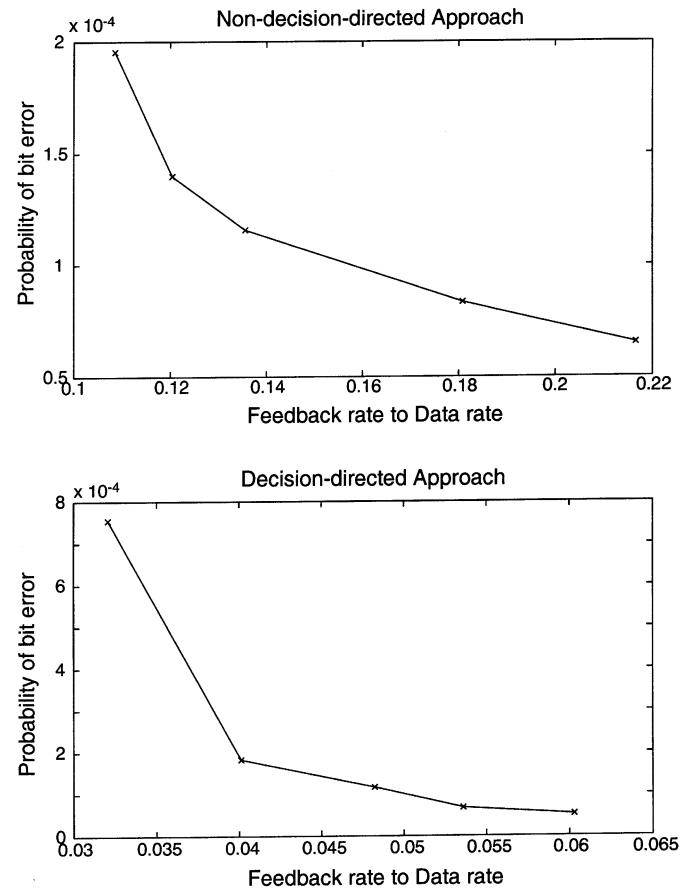


Fig. 5. Probability of bit error versus ratio of feedback rate to data rate.

than any particular basis. Similar performance was found in the nondecision-directed approach.

In order to study the spectral efficiency and the extent of feedback information of the system, we performed Monte-Carlo simulations over 500 000 transmitted symbols. Following the definition in [11], the coherence time is approximately 0.0423 s. Hence, the simulations ran over more than 10 cycles of coherence time. The feedback information was quantized to give fair comparisons. Fig. 5 shows the probability of bit error versus different ratio of feedback rate to data rate at 10 dB SNR. At a bit-error rate of  $10^{-4}$ , the spectral efficiency in the forward link is around 3.32 b/s/Hz for both approaches, compared with

3.34 b/s/Hz obtained with perfect channel knowledge. However, in contrast to the channel capacity of 11.51 b/s/Hz, the discrepancy is due to the use of simple modulation and coding schemes in the simulated system. With the use of more sophisticated schemes, it could approach the channel capacity. Moreover, from the graphs, we see that the feedback ratio for the decision-directed approach is an order of magnitude less than its counterpart.

## VI. CONCLUSIONS

Division-free and division-adhered algorithms are proposed to track the  $\mathbf{U}$  matrix. Both of them have similar performance. However, the former is more practical to implement, while the latter requires fewer operations. Consequently, this would be a tradeoff between the complexity of hardware implementation and the complexity of the algorithms ( $O(r^2d)$  versus  $O(rd)$ ). For the tracking of the  $\mathbf{V}$  matrix, the decision-directed approach is spectrally more efficient and shows a smaller range of fluctuation. Furthermore, extension of the algorithms to include unequal power allocation is proved to be feasible as well.

## REFERENCES

- [1] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," in *Proc. IEEE Vehicular Technology Conf.*, May 1993, pp. 508–511.
- [2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [3] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
- [4] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, pp. 1740–1751, Feb.-Apr. 1994.
- [5] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multielement antennas," *Bell Labs Tech. J.*, pp. 41–59, Autumn 1996.
- [6] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.
- [7] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE*, 1998, pp. 295–300.
- [8] E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T-Bell Labs Internal Tech. Memo, 1995.
- [9] B. Parhami, *Computer Arithmetic: Algorithms and Hardware Designs*. Oxford, U.K.: Oxford Univ. Press, 2000.
- [10] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.
- [11] T. S. Rappaport, *Wireless Communications: Principle and Practice*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [12] W. C. Y. Lee, *Mobile Cellular Telecommunications Systems*. New York: McGraw-Hill, 1989.
- [13] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford, U.K.: Oxford Univ. Press, 1965.
- [14] G. W. Stewart, "Error and perturbation bounds for subspaces associated with certain eigenvalue problems," *SIAM Rev.*, vol. 15, pp. 727–64, 1973.
- [15] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [16] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [17] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.
- [18] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, 1968.



**Ada S. Y. Poon** (S'98) received the B.Eng. and M.Phil. degrees in electrical and electronic engineering from the University of Hong Kong in 1996 and 1997, respectively. She received the M.S. degree in electrical engineering from the University of California at Berkeley in 1999, where she is currently working toward the Ph.D. degree.

Her research interests are wireless communications and digital signal processing.



**David N. C. Tse** (S'91-M'96) received the B.A.Sc. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 1989 and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1991 and 1994, respectively.

From 1994 to 1995, he was a Postdoctoral Member of Technical Staff with AT&T Bell Laboratories. Since 1995, he has been with the Department of Electrical Engineering and Computer Sciences, at Berkeley, where he is currently a Professor. His research interests are in information theory, wireless communications and networking.

Dr. Tse received a 1967 NSERC four-year graduate fellowship from the government of Canada in 1989, a NSF CAREER award in 1998, the Best Paper Awards at the Infocom 1998 and Infocom 2001 conferences, the Erlang Prize in 2000 from the INFORMS Applied Probability Society, and the IEEE Communications and Information Theory Society Joint Paper Award in 2001. He is currently an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY.



**Robert W. Brodersen** (M'76-SM'81-F'82) received the B.S. degrees in electrical engineering and mathematics from California State Polytechnic University, Pomona, in 1966, the Engineers and M.S. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1968, and the Ph.D. degree in engineering from MIT in 1972. He received the award Honor Doctor of Technology (Technologie Doctor Honoris Causa) from the University of Lund, Lund, Sweden in 1999.

From 1972–1976, he was a member of the Technical Staff, Central Research Laboratory, Texas Instruments, Dallas. In 1976, he joined the Department of Electrical Engineering and Computer Sciences (EECS) faculty of the University of California at Berkeley, where he is now the John R. Whinnery Distinguished Professor. His research is focused on new applications of integrated circuits in the area of low power design and wireless communications; and the algorithms, architectures, and CAD tools necessary to support these activities.

Dr. Brodersen is a member of the National Academy of Engineering. He has received a number of awards including the IEEE Morris N. Liebmann Award in 1983; the Technical Achievement Award from the IEEE Circuits and Systems Society in 1986; the Technical Achievement Award from the Signal and Processing Society in 1991; the IEEE Solid-State Circuits Award in 1997; the Mobicom Award from ACM Sigmobile in 1998; the IEEE CAS Golden Jubilee Medal in 2000; and the IEEE Millennium Medal in 2000 for exceptional contributions toward advancing in various forms the Society's goals during the first 50 years of its history. He received the Best Paper conference awards at Eascon in 1973, the International Solid-State Circuits Conferences in 1975, and the European Solid-State Circuits in 1978. In 1979, he received the W.G. Baker Award for the outstanding paper in IEEE Journals and Transactions. In 1985, he received the Best Paper award from the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS. In 1992, he received the Best Tutorial Paper award of the IEEE Communications Society. In 1993, he was awarded the IEEE Communications Journal Best Paper Award. In 2001, he was awarded the Lewis Winner Award for outstanding paper at the IEEE International Solid-State Circuits Conference (ISSCC). In 2003, he was given an award from ISSCC for being one of the top ten contributors over the 50 years of that society.