# Differential MMSE: A Framework for Robust Adaptive Interference Suppression for DS-CDMA Over Fading Channels

Upamanyu Madhow, *Fellow, IEEE*, Kristoffer Bruvold, *Student Member, IEEE*, and Liping Julia Zhu

***Abstract*—The linear minimum mean-squared error (MMSE) criterion is known to provide adaptive algorithms for interference suppression in direct-sequence (DS) code-division multiple-access (CDMA) systems. However, standard MMSE adaptation is not robust to fast fading, being unable to compensate for rapid channel variations. In this paper, we provide a framework for deriving robust adaptive algorithms in this setting based on a new *differential* MMSE (DMMSE) criterion, which is a constrained optimization problem in which the quantity to be tracked is the *ratio* of the data appearing in two successive observation intervals. When applied to a DS-CDMA system with short spreading waveforms (i.e., with period equal to the symbol interval) operating over a flat-fading channel, the DMMSE criterion avoids tracking the fades, exploiting the negligible variation of the fading gain over two consecutive symbols. For frequency-selective fading, the DMMSE criterion is extended to provide a new *eigenrake* receiver which provides interference suppression and diversity combining without requiring explicit information regarding the desired user's propagation channel.**

***Index Terms*—Adaptive equalization, code-division multiple access (CDMA), differential minimum mean-squared error (DMMSE), fading channels, interference suppression, multiuser detection, near–far problem.**

## I. INTRODUCTION

IT HAS BEEN known for some time now that for direct-sequence (DS) code-division multiple-access (CDMA) systems with short spreading waveforms (i.e., in which the period of the spreading waveform equals the inverse of the symbol rate), the multiple-access interference (MAI) has a cyclostationary structure (at the symbol rate) which can be learned and exploited by an adaptive receiver. The resulting adaptive multiuser-detection schemes provide large potential gains over conventional matched-filter receivers without requiring explicit estimates of the MAI parameters. In particular, the linear minimum mean-squared error (MMSE) receiver for a desired user can be implemented adaptively either using a training sequence for that user [1]–[4], or (semi) blindly by using knowledge of the desired user's spreading waveform and propagation channel [5], [7], [20]. However, standard training-based adaptation is known to break down in the presence of time-varying channels, typical of wireless environments [6]. On the other hand, blind adaptation, as in [5], exhibits a higher misadjustment than training-based adaptation under ideal conditions, and is vulnerable to mismatch due to errors in the receiver's estimate of the desired user's propagation channel.

In this paper, we present a new approach to adaptive interference suppression over rapidly time-varying channels based on the *differential* MMSE (DMMSE) criterion. This is a reformulation of the classical linear MMSE criterion, wherein the quantity being tracked can be interpreted as the ratio of two successive elements of the desired data sequence, rather than the raw data sequence. It is shown that the DMMSE criterion leads to adaptive interference-suppression techniques that are robust to channel time variations. The key idea behind the DMMSE criterion is the avoidance of the problem of channel compensation by exploiting instead the observation that even for rapidly varying channels, the channel fading gains in two consecutive observation intervals are approximately the same.

For flat-fading channels, the DMMSE criterion yields a number of adaptive algorithms robust to channel time variations, with complexities comparable to (but slightly larger than) that of analogous algorithms based on the MMSE criterion. Under standard assumptions, the DMMSE correlator is shown to be a scalar multiple of the MMSE correlator. Thus, it inherits the well-known [2], [19] interference-suppression properties of the MMSE correlator, including its immunity to the near–far problem. For frequency-selective fading, the DMMSE criterion provides the starting point for obtaining the *eigenrake* receiver, which provides diversity as well as interference suppression, implicitly acquiring the timing of the significant multipath components for the desired user. As with standard MMSE adaptation, the proposed adaptive algorithms require an initial training period, in which the symbols transmitted by the desired user are known to the receiver, and can subsequently operate in decision-directed mode. The receiver does not require explicit knowledge of the spreading waveforms and propagation channels for either the desired user or the interfering users. Since DMMSE-based algorithms do not explicitly track the channel, they must either be used with a noncoherent demodulation technique (e.g., differential demodulation), or the channel information required for coherent demodulation must be obtained by some other means (e.g., by using pilot symbols).

We have previously reported preliminary results on DMMSE-based reception in conference publications [6], [8], [9]. In this

Paper approved by A. Zanella, the Editor for Wireless Systems of the IEEE Communications Society. Manuscript received May 22, 2004; revised October 25, 2004; January 12, 2005; and February 15, 2005. This work was supported in part by the Office of Naval Research under Grants N00014-95-1-0647 and N00014-03-1-0090, in part by the National Science Foundation under Grants NCR96-24008 (CAREER), EIA-0080134, and ANI-0220118 (ITR), and in part by the U.S. Army Research Office under Grants DAAH04-95-1-0246 and DAAD19-00-1-0567.

The authors are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA (e-mail: madhow@ece.ucsb.edu; bruvold@ece.ucsb.edu).

Digital Object Identifier 10.1109/TCOMM.2005.852851

paper, we provide a comprehensive treatment that includes a detailed development of the theoretical properties of the DMMSE solution for CDMA systems with short spreading sequences, the introduction of the eigenrake receiver for frequency-selective channels, and simulation results that explore various aspects of DMMSE-based reception.

There are several other recent papers that address the problem of adaptive interference suppression over time-varying channels. An algorithm similar to DMMSE, but without a crucial constraint required to obtain the linear MMSE receiver, was proposed in [10]. This algorithm is not robust in itself, and periodically needs to switch to a fallback mode in which the blind algorithm of [5] is used. It was shown in [6] that this problem can be avoided by appropriately scaling the correlator updates. As we point out here, this fix actually corresponds to one possible approximate implementation of the DMMSE criterion. In [11], decision-directed adaptation based on reconstruction of the transmitted symbols after differential demodulation is explored. When channel estimates are available, typically via the use of pilots, an alternative approach, which we term *channel-compensated* MMSE, is to incorporate these estimates into the data sequence being tracked by standard MMSE adaptation. A number of variants of this basic idea have appeared in the literature [12]–[15]. Both DMMSE and channel-compensated MMSE relieve the adaptive mechanism of the burden of channel tracking, but DMMSE does so without requiring explicit channel estimation. A detailed comparison of all of these different approaches for adaptation over time-varying channels is not undertaken here, since our objective is to provide an initial exposition of the DMMSE criterion.

There has also been substantial work in recent years on nonadaptive, noncoherent multiuser detection (e.g., see [16]–[18], and the references therein). For these schemes, the complexity increases with the number of users. The linear decorrelation techniques in [16] have linear complexity, and the optimal detection techniques in [17] and [18] have exponential complexity, while a suboptimal decision-feedback scheme in [17] has quadratic complexity. Furthermore, these techniques require knowledge of the signaling waveforms of all users, even though the channel gains are unknown (additionally, knowledge of the individual users' signal strengths is required in [18]). In contrast, the DMMSE-based implementations proposed here require knowledge only of a short sequence of training symbols for the desired user, with complexity independent of the number of users.

The remainder of this paper is organized as follows. In Section II, the DMMSE criterion is discussed in the context of a complex baseband system model. Section III presents algorithms for adaptively obtaining the DMMSE correlator in the presence of flat fading. DMMSE reception for frequency-selective channels is presented in Section IV, where a new *eigenrake* receiver is proposed. Numerical results are given in Section V, and Section VI contains our conclusions.

## II. DMMSE Reception Over Flat-Fading Channels

We consider a discrete-time, complex baseband, synchronous CDMA system with flat fading. There are $K$ users, with the desired user labeled user 1. Let $\mathbf{r}[n]$ denote the $N$-dimensional vector of samples obtained from the $n$th observation interval, given by

$$\mathbf{r}[n] = F_1[n]b_1[n]\mathbf{u}_1 + \sum_{k=2}^{K} F_k[n]b_k[n]\mathbf{u}_k + \mathbf{w}[n] \quad (1)$$

where, for $1 \le k \le K$, $\mathbf{u}_k$ is the signal vector for user $k$, $\{b_k[n]\}$ is the stream of symbols transmitted by user $k$, and $\{F_k[n]\}$ is the sequence of complex fading gains seen by user $k$. The complex vector $\mathbf{w}[n]$ is discrete-time, additive white Gaussian noise (AWGN) with variance $\sigma^2$ per dimension (the DMMSE formulation applies to colored noise as well, but white noise is considered here for simplicity). For Rayleigh fading, the gains $\{F_k[n]\}$ are modeled as a wide-sense stationary, zero mean, circular Gaussian random process. It is convenient to introduce the *faded symbol* sequence $B_k[n] = F_k[n]b_k[n]$, $1 \le k \le K$, for each user. We make the following assumptions.

1) Symbols are zero mean, and independent across time and users: $E[b_k[n]] = 0$ for all $k, n$, and $b_k[n]$ is independent of $b_j[m]$ for $k \neq j$ or $n \neq m$.
2) Symbols are independent of fading gains: $b_k[n]$ is independent of $F_j[m]$ for any $k, j, n, m$.
3) For each user $k$, the sequence of fading gains $\{F_k[n]\}$ is wide-sense stationary.
4) The fading gains for different users are independent: for $j \neq k$, $F_j[m]$ is independent of $F_k[n]$ for all $m, n$.

While the preceding conditions are stronger than necessary for proving the basic properties of the DMMSE solution, they simplify the proofs of these properties, while preserving the insight into why the DMMSE correlator suppresses interference.

As in much of the multiuser detection literature (see [19] and the references therein), we restrict attention to a synchronous CDMA system for developing the basic properties of the proposed methods. It is known (e.g., see [20] and [19, Ch. 2]) that interference-suppression algorithms based on the synchronous model apply to an asynchronous CDMA system by reducing the latter to an "equivalent synchronous discrete-time model" which depends on the receive filter, sampler, and the length of the observation interval used for each symbol decision (see [20] for a tutorial description of how this is done). Simulation results on the application of DMMSE to asynchronous systems can be found in [6] and [8]. However, the theorems in Section II-B would need slight modifications for their hypotheses, as well as conclusions for an asynchronous system. For simplicity of development, we omit such modifications from this paper.

In general, a linear receiver computes a decision statistic of the form $\langle \mathbf{c}, \mathbf{r}[n] \rangle = \mathbf{c}^H \mathbf{r}[n]$, where $\mathbf{x}^H$ denotes the complex conjugate transposed for a vector $\mathbf{x}$. The standard linear MMSE receiver minimizes the mean squared error (MSE) between the desired user's symbol sequence $b_1[n]$ and the receiver output, given by $E[|b_1[n] - \langle \mathbf{c}, \mathbf{r}[n] \rangle|^2]$. The MMSE correlator is given by the formula

$$\mathbf{c}_{\text{mmse}} = \mathbf{R}^{-1}\mathbf{p} \quad (2)$$

where $\mathbf{R} = E[\mathbf{r}[n](\mathbf{r}[n])^H]$ and $\mathbf{p} = E[b_1^*[n]\mathbf{r}[n]]$.

For Rayleigh fading in the model (1), assuming that $b_1[n]$ is uncorrelated with $F_1[n]$ and $B_k[n]$ for $k \neq 1$, we obtain that

$$\mathbf{p} = E\left[|b_1[n]|^2\right] E\left[F_1[n]\right] \mathbf{u}_1 = \mathbf{0}. \tag{3}$$

This implies that $\mathbf{c}_{\text{mmse}} = \mathbf{0}$ when averaged over the desired user's Rayleigh fading coefficient. Adaptive implementations [21] of the MMSE correlator may be viewed as replacing statistical expectations with empirical averages; for example, replacement of the statistical expectation in (2) by a block-based empirical average leads to the block least-squares implementation, while replacement by an exponentially weighted average corresponds to the recursive least squares (RLS) algorithm. Thus, if the averaging time constant used by an MMSE-based adaptive algorithm is comparable to or larger than the coherence time of the fading, as is the case in many outdoor mobile wireless environments, we should expect poor performance by virtue of (3). For example, a normalized Doppler spread[1] of 0.01 corresponds to a coherence time of 100 symbols, and could result from operating at a symbol rate of 20 Ksymbols/s, a carrier frequency of 2 GHz, and a relative velocity between transmitter and receiver of approximately 100 km/h. In this setting, RLS adaptation employing an exponential forget factor of 0.99, which effectively averages over hundreds of symbols (i.e., an interval of the order of the coherence time), fails (see the simulation results in Section V). Decreasing the averaging time in the adaptive algorithm would alleviate this problem, but would then provide insufficient averaging to overcome the effect of noise and interference.

Most commercial systems employ known pilot codes or pilot symbols in order to track the channel and perform coherent demodulation. In this case, the receiver may be able to estimate the fading gain $F_1[n]$,[2] and the MMSE criterion can be modified so as to track the faded symbol $B_1[n]$ using the cost function $E[|F_1[n]b_1[n] - \langle \mathbf{c}, \mathbf{r}[n]\rangle|^2] = E[|B_1[n] - \langle \mathbf{c}, \mathbf{r}[n]\rangle|^2]$. This eliminates the need to compensate for the fading gain. Assuming now that $B_1[n]$ is uncorrelated with $B_k[n]$ for $k \neq 1$, the channel-compensated solution is given by

$$\mathbf{c}_{\text{cc}} = \mathbf{R}^{-1}\mathbf{p}_{\text{cc}} \tag{4}$$

where

$$\mathbf{p}_{\text{cc}} = E\left[B_1^*[n]\mathbf{r}[n]\right] = E\left[|B_1[n]|^2\right]\mathbf{u}_1 \sim \mathbf{u}_1 \tag{5}$$

so that the overall solution is proportional to $\mathbf{R}^{-1}\mathbf{u}_1$.

The channel-compensated MMSE solution can be interpreted as a standard MMSE solution in a time-invariant setting, except that the data being tracked is the faded symbol $B_1[n]$, rather than the symbol $b_1[n]$. From well-known properties of the MMSE solution [2], we can infer that channel-compensated MMSE is effective in interference suppression. This approach to dealing with channel time variations has been considered in several recent publications [12]–[14], [22], [23].

[1]The normalized Doppler spread is the product $f_d T_s$ of the maximum Doppler frequency $f_d$ and the symbol period $T_s$.

[2]Of course, accurate channel estimation prior to interference suppression may not be easy, especially when there is a near–far problem.

### A. DMMSE Criterion

In contrast to the channel-compensated MMSE approach in (4) and (5), the DMMSE criterion does not require explicit estimation of the fading gains for the desired user. It relies instead on the assumption that $F_1[n] \approx F_1[n-1]$, even for "fast" fading environments, to obtain a correlator equivalent to the channel-compensated MMSE correlator. The formal statement of the DMMSE criterion is as follows.

*The DMMSE Criterion*: Choose a correlator $\mathbf{c}$ that solves the following problem.

*Problem P1*: Minimize over $\mathbf{c}$

$$J(\mathbf{c}) = E\left[\left|b_1[n]\langle \mathbf{c}, \mathbf{r}[n-1]\rangle - b_1[n-1]\langle \mathbf{c}, \mathbf{r}[n]\rangle\right|^2\right] \tag{6}$$

subject to

$$E\left[|\langle \mathbf{c}, \mathbf{r}[n]\rangle|^2\right] = \mathbf{c}^H\mathbf{R}\mathbf{c} = 1. \tag{7}$$

The intuition behind the preceding optimization problem is as follows. Given the difficulty in tracking $F_1[n]$, we aim to design an adaptive receiver that achieves a more modest goal, that of suppressing the interference and recovering the faded sequence $B_1[n] = F_1[n]b_1[n]$ *up to an arbitrary complex multiple*, $\alpha$. A correlator $\mathbf{c}$ that achieves this goal will satisfy

$$\langle \mathbf{c}, \mathbf{r}[n-1]\rangle \approx \alpha B_1[n-1] \approx \alpha F_1[n-1]b_1[n-1]$$
$$\langle \mathbf{c}, \mathbf{r}[n]\rangle \approx \alpha B_1[n] \approx \alpha F_1[n]b_1[n]$$

so that, assuming $F_1[n] \approx F_1[n-1]$, we have

$$b_1[n]\langle \mathbf{c}, \mathbf{r}[n-1]\rangle - b_1[n-1]\langle \mathbf{c}, \mathbf{r}[n]\rangle \approx 0.$$

This implies that (6) is the natural cost function to minimize. However, the solution to an unconstrained minimization of (6) is the zero correlator, the avoidance of which requires the imposition of a suitable constraint. As stated in *Theorem 1* in the next section, the specific constraint (7) on the average output energy leads to an optimization problem whose solution, under mild assumptions, is the linear MMSE solution.

### B. Basic Properties

We state two theorems below. *Theorem 1* supplies the basis for adaptive algorithms based on the DMMSE criterion. *Theorem 2* states that under natural uncorrelatedness conditions, the DMMSE solution is a scalar multiple of the channel-compensated linear MMSE solution in (4).

*Theorem 1 (DMMSE Solution):* Assuming that the desired user employs a constant modulus alphabet, the general solution to problem P1 is the eigenvector corresponding to the largest eigenvalue of the following generalized eigenvalue problem:

$$\mathbf{A}\mathbf{c} = \lambda\mathbf{R}\mathbf{c} \tag{8}$$

where

$$\mathbf{R} = E\left[\mathbf{r}[n]\left(\mathbf{r}[n]\right)^H\right] \tag{9a}$$

$$\mathbf{A} = E\Big[b_1[n]b_1^*[n-1]\mathbf{r}[n-1]\left(\mathbf{r}[n]\right)^H + b_1^*[n]b_1[n-1]\mathbf{r}[n]\left(\mathbf{r}[n-1]\right)^H\Big]. \tag{9b}$$

*Proof:* Expanding the cost function in (6), we obtain that $J(\mathbf{c}) = \mathbf{c}^H \mathbf{K} \mathbf{c}$, where

$$\begin{aligned}\mathbf{K} = & E\left[|b_1[n]|^2 \mathbf{r}[n-1] (\mathbf{r}[n-1])^H\right] \\ & + E\left[|b_1[n-1]|^2 \mathbf{r}[n] (\mathbf{r}[n])^H\right] \\ & - E\left[b_1[n] b_1^*[n-1] \mathbf{r}[n-1] (\mathbf{r}[n])^H\right] \\ & - E\left[b_1^*[n] b_1[n-1] \mathbf{r}[n] (\mathbf{r}[n-1])^H\right]. \end{aligned} \tag{10}$$

Under our assumptions, $b_1[n-1]$ is independent of $\mathbf{r}[n]$, and $b_1[n]$ is independent of $\mathbf{r}[n-1]$. Normalizing the symbol energy $E[|b_1[n]|^2] = 1$ without loss of generality, we see that the first two terms in (10) are each equal to $\mathbf{R}$. The third and fourth terms correspond to the matrix $\mathbf{A}$ defined in (9b), so that

$$\mathbf{K} = 2\mathbf{R} - \mathbf{A}.$$

The cost function (6) therefore reduces to

$$J(\mathbf{c}) = 2\mathbf{c}^H \mathbf{R} \mathbf{c} - \mathbf{c}^H \mathbf{A} \mathbf{c} = 2 - \mathbf{c}^H \mathbf{A} \mathbf{c} \tag{11}$$

where we have used the constraint (7) to obtain the second equality. It is now clear that the DMMSE problem of minimizing $J(\mathbf{c})$ under the constraint (7) is equivalent to the following problem:

$$\text{Maximize } \mathbf{c}^H \mathbf{A} \mathbf{c} \text{ subject to } \mathbf{c}^H \mathbf{R} \mathbf{c} = 1. \tag{12}$$

The Lagrangian for this problem is

$$\Lambda(\mathbf{c}) = \mathbf{c}^H \mathbf{A} \mathbf{c} + \lambda \mathbf{c}^H \mathbf{R} \mathbf{c}.$$

Setting the gradient of $\Lambda$ with respect to the complex conjugate of $\mathbf{c}$ to zero yields

$$\mathbf{A}\mathbf{c} = \lambda \mathbf{R} \mathbf{c}.$$

Multiplying both sides of this equation by $\mathbf{c}^H$, and using the constraint $\mathbf{c}^H \mathbf{R} \mathbf{c} = 1$, we obtain

$$\lambda = \mathbf{c}^H \mathbf{A} \mathbf{c}.$$

Since the right-hand side above is the quantity to be maximized, the optimal solution is the eigenvector corresponding to the largest eigenvalue. This completes the proof. □

*Remark 1 (General Structure of DMMSE Solution): Theorem 1* does not depend on the flat-fading model in (1). Rather, it is a general characterization of the structure of the DMMSE solution, analogous to the well-known formula (2) for the MMSE solution. Note that $\mathbf{R}$ is Hermitian nonnegative definite and $\mathbf{A}$ is Hermitian, so that there are a number of well-known algorithms [24] that can be brought to bear on the "symmetric-definite" generalized eigenvalue problem in (8).

Next, we invoke the specific features of the CDMA model (1), and show that the DMMSE correlator is equivalent to the channel-compensated MMSE solution in (4), and hence, suppresses interference.

*Theorem 2 (DMMSE Interference Suppression):* Suppose that the fading gains in consecutive intervals for the desired user are positively correlated; that is, defining $\rho = E[F_1[n] F_1^*[n-1]]$, we require that $\Re\{\rho\} > 0$.

Then the solution to optimization problem P1 is a scalar multiple of $\mathbf{R}^{-1}\mathbf{u}_1$ and is, therefore, equivalent to the channel-compensated MMSE solution. That is, it is a scalar multiple of the MMSE solution for demodulating the desired user's faded symbol, $B_1[n] = F_1[n] b_1[n]$. The DMMSE solution therefore inherits the interference-suppression properties of the MMSE solution.

*Proof: Theorem 1* applies, since the conditions imposed here are a subset of those of *Theorem 1*. We now compute the matrix $\mathbf{A}$ in *Theorem 1* for the model (1). Letting $\mathbf{T}_1 = E[b_1^*[n] b_1[n-1] \mathbf{r}[n] (\mathbf{r}[n-1])^H]$ and plugging in the model (1), we obtain

$$\mathbf{T}_1 = E\left[b_1^*[n] b_1[n-1] \sum_{k=1}^{K} F_k[n] b_k[n] \mathbf{u}_k \times \sum_{j=1}^{K} F_j^*[n-1] b_j^*[n-1] \mathbf{u}_j^H\right].$$

Under the independence assumptions of our model, it is easy to verify that

$$E\left[b_1^*[n] b_1[n-1] \left(F_k[n] b_k[n]\right) \left(F_j^*[n-1] b_j^*[n-1]\right)\right] = 0$$

*unless* $j = k = 1$. Thus, we obtain that

$$\begin{aligned}\mathbf{T}_1 &= E\left[\,|b_1[n-1]|^2\,\right] E\left[\,|b_1[n]|^2\,\right] E\left[F_1[n] F_1^*[n-1]\right] \mathbf{u}_1 \mathbf{u}_1^H \\ &= \rho \mathbf{u}_1 \mathbf{u}_1^H.\end{aligned}$$

Similarly, $\mathbf{T}_1^H = E\left[b_1[n] b_1^*[n-1] \mathbf{r}[n-1] (\mathbf{r}[n])^H\right] = \rho^* \mathbf{u}_1 \mathbf{u}_1^H$. We therefore obtain from (9b) that

$$\mathbf{A} = \mathbf{T}_1^H + \mathbf{T}_1 = 2\Re\{\rho\} \mathbf{u}_1 \mathbf{u}_1^H. \tag{13}$$

Since this is a rank-one matrix, there is a unique nonzero eigenvalue for the generalized eigenvalue problem (8), which can be rewritten as

$$2\Re\{\rho\} \mathbf{u}_1 \mathbf{u}_1^H \mathbf{c} = \lambda \mathbf{R} \mathbf{c}.$$

Multiplying each side by $\mathbf{R}^{-1}$, it is clear that the generalized eigenvector corresponding to the nonzero eigenvalue is a scalar multiple of $\mathbf{R}^{-1}\mathbf{u}_1$. The condition $\Re\{\rho\} > 0$ ensures that the unique nonzero eigenvalue is positive. If this were not the case, then the largest eigenvalue for the solution to the DMMSE problem would be zero. The latter corresponds to the zero correlator, which is useless for demodulation. The necessity of $\Re\{\rho\} > 0$ for obtaining a useful DMMSE correlator is, of course, not surprising, since the formulation of the DMMSE criterion is based on the assumption that the fading gains in successive intervals are approximately equal. This concludes the proof. □

*Remark 2 (Tracking Fading Unnecessary With DMMSE):* Note that the signal vector $\mathbf{u}_1$ is independent of the fading gains $\{F_1[n]\}$, and does not vary over time. Furthermore

$$\mathbf{R} = \sum_{k=1}^{K} E\left[|B_k[n]|^2\right] \mathbf{u}_k \mathbf{u}_k^H + 2\sigma^2 \mathbf{I}$$

depends only on the average power of the faded symbol sequences $\{B_k[n]\}$, so that its computation does not require tracking of (any unknown gains embedded in) these sequences. However, when the interfering users see multipath channels, each independent multipath component will appear as a separate *virtual* interferer, causing performance degradation. A centralized multiuser detector that tracks the gains of each multipath component could potentially avoid this, at the cost of higher implementation complexity. Another scenario in which it may be beneficial to track instantaneous interference gains is when the distribution of the fading power can vary more widely around its mean than is the case for Rayleigh fading. In this case, a receiver suppressing interference vectors based on their average powers may perform significantly worse than one employing instantaneous powers.

*Remark 3 (Alternative Interpretation of DMMSE):* For the flat-fading model (1), under the conditions of *Theorems 1* and *2*, the DMMSE criterion may be alternatively interpreted as *constrained maximization of the desired output energy*, as follows [25]:

$$\text{Maximize } |\langle \mathbf{c}, \mathbf{u}_1 \rangle|^2 \text{ subject to } \mathbf{c}^H \mathbf{R} \mathbf{c} = 1.$$

To see this, insert the formula (13) for $\mathbf{A}$ from the proof of *Theorem 2* into the alternative formulation (12) of the DMMSE criterion in the proof of *Theorem 1*.

This amounts to maximizing the desired signal power at the output, subject to a constraint on the output energy, which includes the energy due to the desired signal, the interference, and the noise. There is an interesting duality between this and the blind constrained minimum output energy detector in [5], which resulted from minimizing the output energy $\mathbf{c}^H \mathbf{R} \mathbf{c}$, subject to $\langle \mathbf{c}, \mathbf{u}_1 \rangle = 1$ [25].

*Remark 4 (Whitening Interpretation):* The MMSE solution $\mathbf{R}^{-1}\mathbf{u}_1$ can be shown [2] (or by a simple application of the matrix-inversion lemma; see the Sherman–Morrison–Woodbury formula in [24]) to be a scalar multiple of $\hat{\mathbf{c}} = \mathbf{R}_I^{-1}\mathbf{u}_1$, where

$$\begin{aligned} \mathbf{R}_I &= \sum_{k=2}^{K} E\left[|B_k[n]|^2\right] \mathbf{u}_k \mathbf{u}_k^H + 2\sigma^2 \mathbf{I} \\ &= \mathbf{R} - E\left[|B_1[n]|^2\right] \mathbf{u}_1 \mathbf{u}_1^H \end{aligned} \tag{14}$$

is the correlation matrix for the interference and noise. The decision statistic $Z = \hat{\mathbf{c}}^H \mathbf{r}[n] = (\mathbf{R}_I^{-1}\mathbf{u}_1)^H \mathbf{r}[n]$ can be rewritten as $(\mathbf{R}_I^{-\frac{1}{2}}\mathbf{u}_1)^H \mathbf{R}_I^{-\frac{1}{2}} \mathbf{r}[n]$. Since the transformation $\mathbf{R}_I^{-\frac{1}{2}}$ whitens the sum of the interference and noise, the MMSE solution is simply the matched filter in the whitened domain. Since a direct estimate of $\mathbf{R}_I$ is not available, the equivalence (up to scalar multiple) of $\mathbf{R}^{-1}\mathbf{u}_1$ and $\mathbf{R}_I^{-1}\mathbf{u}_1$ implies that $\mathbf{R}^{-\frac{1}{2}}$, based on the correlation matrix for the signal plus interference, can be used as the whitening transformation. This interpretation is useful when discussing the effects of the constraint in the next section, and in extending the DMMSE criterion to multipath fading channels in Section IV.

### C. Properties of Unconstrained DMMSE

Having explored the basic properties of the DMMSE criterion, we can now comment in more detail on the role of the constraint $\mathbf{c}^H \mathbf{R} \mathbf{c} = 1$. It is most convenient to discuss this under the assumption that the received signal has been prewhitened. Using $\tilde{\mathbf{x}}$ to denote the whitened version of $\mathbf{x}$, we have

$$\tilde{\mathbf{r}}[n] = \mathbf{R}^{-\frac{1}{2}} \mathbf{r}[n] \tag{15}$$

$$\tilde{\mathbf{u}}_1 = \mathbf{R}^{-\frac{1}{2}} \mathbf{u}_1 \tag{16}$$

$$\tilde{\mathbf{R}} = \mathbf{I} \tag{17}$$

$$\tilde{\mathbf{R}}_I = \mathbf{R}^{-\frac{1}{2}} \mathbf{R}_I \mathbf{R}^{-\frac{1}{2}}. \tag{18}$$

From (14), we now have that $\tilde{\mathbf{R}} = \mathbf{I} = \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H + \tilde{\mathbf{R}}_I$. Thus, as long as $\tilde{\mathbf{R}}_I$ (the noise and interference correlation matrix in the whitened domain) is positive definite, we have that $\mathbf{I} - \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H$ is positive definite. This can be used to show that $\|\tilde{\mathbf{u}}_1\|^2 < 1$.

Now, consider the DMMSE cost function (without the constraint) in the whitened domain, specializing (11) and (13)

$$J(\tilde{\mathbf{c}}) = 2\tilde{\mathbf{c}}^H \tilde{\mathbf{c}} - 2\Re\{\rho\} \tilde{\mathbf{c}}^H \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H \tilde{\mathbf{c}}. \tag{19}$$

The gradient of the preceding cost function (with respect to $\tilde{\mathbf{c}}^*$) is given by $2\tilde{\mathbf{c}} - 2\Re\{\rho\} \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H \tilde{\mathbf{c}}$, so that a gradient-descent update is of the form

$$\begin{aligned} \tilde{\mathbf{c}}[n+1] &= \tilde{\mathbf{c}}[n] - \alpha \left(\tilde{\mathbf{c}}[n] - \Re\{\rho\} \tilde{\mathbf{u}}_1 \tilde{\mathbf{u}}_1^H \tilde{\mathbf{c}}[n]\right) \\ &= (1-\alpha)\tilde{\mathbf{c}}[n] + \alpha \left(\Re\{\rho\} \tilde{\mathbf{u}}_1^H \tilde{\mathbf{c}}[n]\right) \tilde{\mathbf{u}}_1. \end{aligned} \tag{20}$$

It is now easy to see what happens in the absence of the constraint. Recall that in the whitened domain, the desired MMSE solution is simply $\tilde{\mathbf{u}}_1$. From (20), the component of $\tilde{\mathbf{c}}[0]$ orthogonal to $\tilde{\mathbf{u}}_1$ gets attenuated exponentially as $(1-\alpha)^n$, which is exactly the desired behavior. However, the component of $\tilde{\mathbf{c}}$ along $\tilde{\mathbf{u}}_1$ is also shrinking (but more slowly). This happens because the desired signal vector $\tilde{\mathbf{u}}_1$ has norm strictly less than one in the whitened domain. To see this, suppose that $\tilde{\mathbf{c}}[0] = \beta \tilde{\mathbf{u}}_1$. Then

$$\tilde{\mathbf{c}}[n] = \beta \gamma^n \tilde{\mathbf{u}}_1$$

where $\gamma = 1 - \alpha + \alpha \Re\{\rho\} \|\tilde{\mathbf{u}}_1\|^2 < 1$, since $\|\tilde{\mathbf{u}}_1\|^2 < 1$. Thus, while $\tilde{\mathbf{c}}[n]$ points in the right direction, its norm is shrinking to zero (although this shrinkage can be slow if $\gamma$ is close to one). We will see the consequence of this on adaptive implementations when we discuss the scaled RLS algorithm in Section III-A.2.

### D. Demodulation Based on the DMMSE Output

In our performance evaluations for the flat-fading model (1), we will consider differential phase-shift keying (DPSK) transmission where the transmitted symbols satisfy $b_k[n] = a_k[n] b_k[n-1]$, with $\{a_k[n]\}$ denoting the sequence of information symbols for the $k$th user. Assuming that the $\{a_k[n]\}$ are zero mean and independent across users and time (i.e., both $k$ and $n$), the transmitted symbols $\{b_k[n]\}$ are also, and assumption 1) stated at the beginning of this section holds. DPSK is naturally matched to the DMMSE criterion, since the latter gives the MMSE solution up to complex scaling. We will assume that symbol decisions for the information sequence $a_1[n]$ are based on the standard differentially coherent decision statistic $\langle \mathbf{c}, \mathbf{r}[n] \rangle \langle \mathbf{c}, \mathbf{r}[n-1] \rangle^*$. Specifically, we focus on binary DPSK and employ the decision rule $\hat{a}_1[n] = \text{sign}(\Re\{\langle \mathbf{c}, \mathbf{r}[n] \rangle \langle \mathbf{c}, \mathbf{r}[n-1] \rangle^*\})$. For larger PSK alphabets, significant performance gains can be obtained using block differential demodulation [26], [27], if the channel is approximately constant over a larger block of symbols, but we do not investigate such issues here. Another option, not explored in detail here, is to use the DMMSE algorithm for robust adaptation (the robustness is because the correlator taps do not have to rotate to track fades) and to perform coherent demodulation after resolving the scalar ambiguity at the output of the correlator via the use of pilot symbols.

## III. Adaptive Algorithms

We now provide a number of adaptive algorithms for implementing the DMMSE criterion (see Section IV-D for a note on their complexity). We consider a DPSK system as in Section II-D, and assume that the sequence $a_1[n] = b_1[n]b_1^*[n-1]$ is known in training mode and can be estimated in decision-directed mode.

### A. Gradient-Based Adaptive Algorithms

Two recursive algorithms, termed scaled RLS and scaled normalized least mean squares (NLMS), respectively, obtained based on heuristic reasoning in our previous work [6], [28], can be interpreted as approximate implementations of the DMMSE criterion. Proceeding as when deriving the conventional RLS or NLMS algorithms, the gradient of the unconstrained cost function

$$J(\mathbf{c}) = E\left[\left|b_1[n]\left\langle \mathbf{c}, \mathbf{r}[n-1]\right\rangle - b_1[n-1]\left\langle \mathbf{c}, \mathbf{r}[n]\right\rangle\right|^2\right]$$

is given by $\nabla J = 2E\left[(b_1[n]\mathbf{r}[n-1] - b_1[n-1]\mathbf{r}[n])e^*[n]\right]$, where the error $e[n] = b_1[n]\langle \mathbf{c}, \mathbf{r}[n-1]\rangle - b_1[n-1]\langle \mathbf{c}, \mathbf{r}[n]\rangle$. Thus, given that the receiver is on the constraint surface, one method of iterative minimization is to perform an RLS or NLMS update based on a stochastic version of the preceding gradient, as usual. However, after the update, the correlator obtained may lie outside the constraint surface. It is then scaled back to the surface, using the "stochastic constraint" $|\langle \mathbf{c}, \mathbf{r}[n]\rangle| = 1$. Other techniques are also applied to optimizing the performance further (see [28] and [6] for details). However, the key issue is that the scaling places successive updates roughly on the constraint surface, thereby avoiding the zero solution (some practical issues regarding the implementation are discussed in Section III-A.2). Since RLS can be viewed as a *whitened stochastic gradient* algorithm, the scaled RLS and scaled NLMS algorithms thus obtained can both be viewed as *projection gradient* algorithms. For completeness, we provide the description of the scaled RLS algorithm (which is used to generate some of the numerical results reported here) below. The scaled NLMS algorithm is found not to perform well in a time-varying environment, and is not described here.

*1) Scaled RLS Algorithm:* The scaled RLS adaptation [28] is as follows.

First, compute

$$\begin{aligned}
\mathbf{k}[n] &= \frac{b_1[n-1]\mathbf{P}[n-1]\mathbf{r}[n]}{\beta + |b_1[n-1]|^2 (\mathbf{r}[n])^H \mathbf{P}[n-1]\mathbf{r}[n]} \\
\mathbf{P}[n] &= \beta^{-1}\mathbf{P}[n-1] - \beta^{-1} b_1^*[n-1]\mathbf{k}[n] (\mathbf{r}[n])^H \mathbf{P}[n-1] \\
\xi[n] &= b_1[n] (\mathbf{c}_{sc}[n-1])^H \mathbf{r}[n-1] \\
&\quad - b_1[n-1] (\mathbf{c}_{sc}[n-1])^H \mathbf{r}[n] \qquad (21)
\end{aligned}$$

then update the scaled RLS receiver $\mathbf{c}_{sc}$

$$\mathbf{c}_{sc}[n] = \frac{\mathbf{c}_{sc}[n-1] + \beta_a \mathbf{k}[n]\xi^*[n]}{\left|(\mathbf{c}_{sc}[n-1])^H \mathbf{r}[n-1]\right|}. \qquad (22)$$

The factor $\beta_a \leq 1$ represents a slowing down of the RLS update, which appears to help alleviate the effect of unreliable estimates during deep fades of the desired signal.

*2) Effect of Scaling:* We can now illustrate the discussion in Section II-C of gradient descent on the unconstrained DMMSE cost function via a concrete adaptive implementation. To this end, consider an unscaled RLS algorithm, which is identical to (21) and (22) except that the scaling in the denominator is eliminated. An algorithm equivalent to the latter was reported in [10], where it was called differential least squares. Since RLS can be viewed as stochastic gradient descent in the whitened domain, we expect from the discussion of gradient descent in Section II-C that the unscaled RLS solution will "point in the right direction," but will ultimately converge to the zero correlator. To see this, consider the explicit unscaled RLS iteration, removing the scaling from (22)

$$\begin{aligned}
\mathbf{c}_{us}[n] &= \mathbf{c}_{us}[n-1] + \beta_a \mathbf{k}[n]\xi^*[n] \\
&= \mathbf{Q}[n]\mathbf{c}_{us}[n-1] \qquad (23)
\end{aligned}$$

with

$$\mathbf{Q}[n] = \mathbf{I} + \beta_a \mathbf{k}[n] (b_1[n]\mathbf{r}[n-1] - b_1[n-1]\mathbf{r}[n])^H .$$

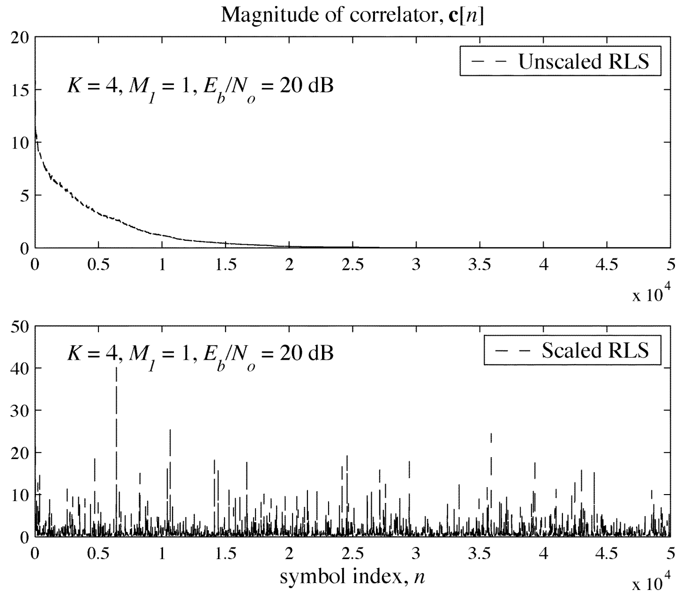


Fig. 1. Magnitude of correlator $\mathbf{c}[n]$ for unscaled (top) and scaled (bottom) RLS adaptations with $E_b/N_0 = 20$ dB.

As shown in Fig. 1, the unscaled RLS correlator does converge to zero, while the scaled RLS correlator does not. However, the correlator magnitude for the latter fluctuates wildly, due to the stochastic scaling we employ (as discussed later, an averaged scale factor may be more appropriate for practical implementations).

In order to further explore the issue of scaling, we iterate (23) to obtain

$$\mathbf{c}_{us}[n] = \mathbf{V}[n]\mathbf{c}_{us}[0] \qquad (24)$$

where $\mathbf{V}[n] = \prod_{i=1}^{n} \mathbf{Q}[i] = \mathbf{V}[n-1]\mathbf{Q}[n]$. In Fig. 2, we plot the largest eigenvalue of $\mathbf{V}[n]$; its decrease with $n$ implies, from (24), that $\mathbf{c}_{us}[n]$ converges to the zero correlator.

Moreover, it is easy to see that the scaled RLS iteration (22) yields a scalar multiple of the unscaled correlator, as follows:

$$\mathbf{c}_{sc}[n] = \frac{1}{\left|(\mathbf{c}_{us}[n-1])^H \mathbf{r}[n-1]\right|}\mathbf{c}_{us}[n]. \qquad (25)$$

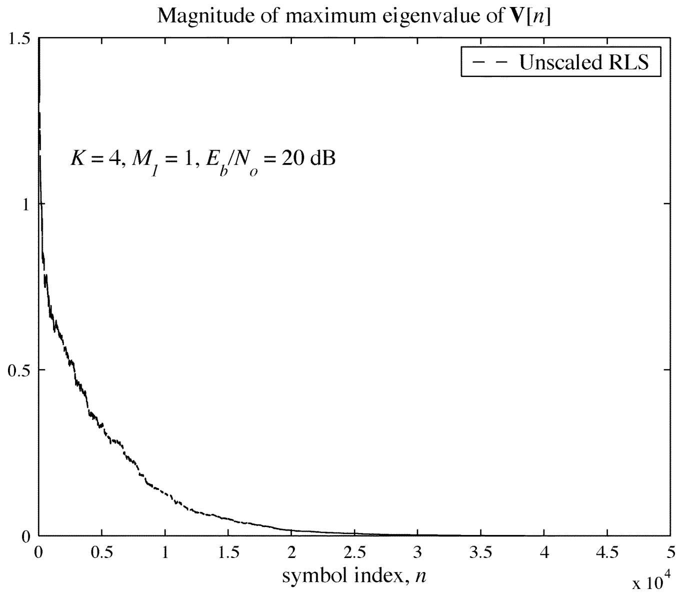


Fig. 2. Magnitude of maximum eigenvalue of $\mathbf{V}[n]$ for the unscaled RLS adaptation with $E_b/N_0 = 20$ dB.

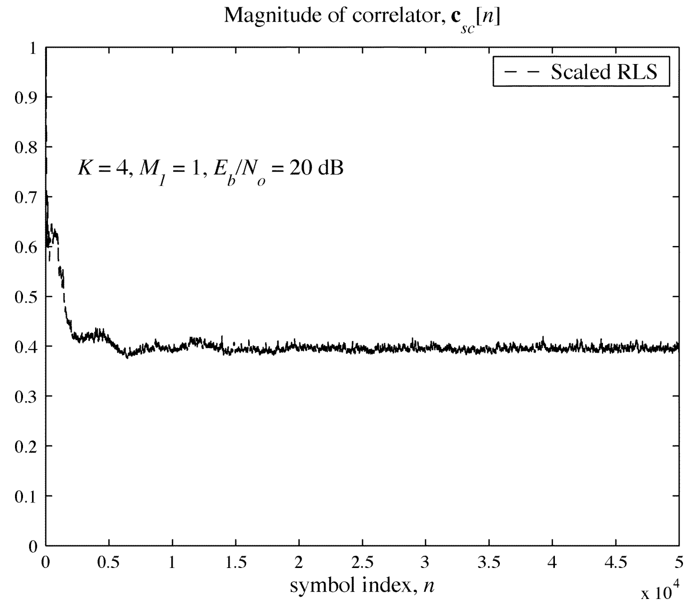


Fig. 3. Magnitude of correlator $\mathbf{c}_{sc}[n]$, using scaling of (26) with $E_b/N_0 = 20$ dB.

Thus, for implementations which represent the correlators with sufficient precision, the bit-error rate (BER) performance of the scaled and unscaled RLS correlators should be precisely the same. We have verified from our double precision floating-point Matlab simulations that this is indeed the case.

There are several practical implications of the preceding observations. First, any scaling that keeps $\mathbf{c}_{us}[n]$ away from zero will work. The particular scaling in the scaled RLS algorithm introduced in [28] and reproduced in (21) is actually quite noisy, and a more appropriate scaling might be based on averaging the power at the output of the correlator, using an empirical estimate of the correlation matrix $\hat{\mathbf{R}}[n] = (\frac{1}{n}) \sum_{i=1}^{n} \mathbf{r}[i](\mathbf{r}[i])^H$, and the constraint in (7), as follows:

$$\mathbf{c}_{sc}[n] = \frac{\mathbf{Q}[n]\mathbf{c}_{sc}[n-1]}{\sqrt{\left|(\mathbf{c}_{sc}[n-1])^H \hat{\mathbf{R}}[n-1]\mathbf{c}_{sc}[n-1]\right|}}. \tag{26}$$

The smoothed response of the correlator's magnitude to this scaling is shown in Fig. 3.

Second, the convergence to zero of the unscaled correlator is often slow enough that, for packetized communication with a small enough number of bits, scaling of the correlator may not be required. For example, in Fig. 1, although the correlator norm decreases by a factor of nine over 10 000 user data symbols, it is still greater than one, which may be acceptable for moderate-sized packets.

For the remainder of the paper, we focus on block rather than recursive implementations of the DMMSE criterion, since these extend more readily to multipath channels.

### *B. Block Power Updates*

Replacing the statistical expectations $\mathbf{R}$ and $\mathbf{A}$ in *Theorem 2* by empirical averages over $T$ observation intervals (i.e., training over $T$ symbols) leads to a block adaptive implementation. This is analogous to the block least-squares algorithm for standard MMSE adaptation. We compute empirical estimates of $\mathbf{R}$ and $\mathbf{A}$ as follows:

$$\hat{\mathbf{R}} = \frac{1}{T}\sum_{n=1}^{T} \mathbf{r}[n]\left(\mathbf{r}[n]\right)^H$$

$$\begin{aligned}\hat{\mathbf{A}} &= \frac{1}{T}\sum_{n=1}^{T} \Big(b_1[n]b_1^*[n-1]\mathbf{r}[n-1]\left(\mathbf{r}[n]\right)^H \\ &\qquad + b_1^*[n]b_1[n-1]\mathbf{r}[n]\left(\mathbf{r}[n-1]\right)^H\Big) \\ &= \frac{1}{T}\sum_{n=1}^{T} a_1[n]\mathbf{r}[n-1]\left(\mathbf{r}[n]\right)^H + a_1^*[n]\mathbf{r}[n]\left(\mathbf{r}[n-1]\right)^H.\end{aligned}$$

The generalized eigenvalue problem for the pair $(\hat{\mathbf{A}}, \hat{\mathbf{R}})$ may be solved by standard techniques such as the QZ method [24]. However, since only the dominant generalized eigenvector is required, an efficient algorithm is to apply the power method [24] to the matrix $\mathbf{M} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{A}}$ until convergence. Given an initial condition $q^{(0)} \in \mathcal{C}^N$, the power method produces a sequence of vectors $q^{(k)}$ as follows:

$$\begin{aligned}&\text{For } k = 1, 2, \ldots \\ &\qquad z^{(k)} = \mathbf{M}q^{(k-1)}; \\ &\qquad q^{(k)} = \frac{z^{(k)}}{\left\|z^{(k)}\right\|}; \\ &\qquad \lambda^{(k)} = \left(q^{(k)}\right)^H \mathbf{M}q^{(k)}; \\ &\text{end.}\end{aligned} \tag{27}$$

The $(\lambda^{(k)}, q^{(k)})$ pair converges to the dominant eigenvalue, eigenvector pair if the initial vector $q^{(0)}$ has a component in the direction of the dominant eigenvector [24].

*Remark 5:* The power algorithm converges to the eigenvalue with the largest magnitude, while the DMMSE criterion requires the eigenvalue with the largest value, accounting for sign. Under the conditions of *Theorem 2*, $\mathbf{A}$, and hence $\mathbf{R}^{-1}\mathbf{A}$, is nonnegative definite so that the largest eigenvalue is positive. In the block adaptive implementation above, the empirical average $\hat{\mathbf{R}}$ is nonnegative definite for any realization, but $\hat{\mathbf{A}}$ may only be approximately nonnegative definite, so that $\hat{\mathbf{R}}^{-1}\hat{\mathbf{A}}$ may not be nonnegative definite. However, the negative eigenvalues are small in *magnitude*, given that the steady-state matrix is nonnegative definite. Hence, the power iterations should still converge to the right solution. This has been the case in all the simulations in this paper.

## IV. Eigenrake Reception for Frequency-Selective Fading

The preceding sections described the fundamentals of the DMMSE criterion in the context of the flat-fading model (1). We now discuss the extension of DMMSE concepts to frequency-selective fading channels. The received vector for the $n$th symbol decision now takes the following form:

$$\mathbf{r}[n] = b_1[n]\mathbf{u}_1[n] + \sum_{k=2}^{K} b_k[n]\mathbf{u}_k[n] + \mathbf{w}[n] \tag{28}$$

where the time-varying, multipath signal vector for the $k$th user, $1 \leq k \leq K$, is given by

$$\mathbf{u}_k[n] = \sum_{i=1}^{M_k} F_{k,i}[n]\mathbf{v}_{k,i}.$$

Here, $M_k$ is the number of resolvable fading paths for the $k$th user, $F_{k,i}[n]$ is the time-varying channel gain for the $i$th resolvable path of the $k$th user, and $\mathbf{v}_{k,i}$ is the effective spreading waveform for the $i$th path of the $k$th user.

For fast-fading channels, uncorrelated multipath components from an interfering user appear as different "users" to a linear MMSE or DMMSE receiver, which therefore tries to separately suppress the interference corresponding to each multipath component. This *interferer multiplication* phenomenon leads to increased noise enhancement, as is well known [6], [22]. This penalty is unavoidable for any linear interference-suppression scheme, unless the multipath components of the interfering users are tracked and combined prior to interference suppression. We do not focus on this issue here, since our interest is in the effect of multipath fading for the desired user (and in the absence of knowledge or estimates of the desired user's channel at the receiver, as in [14]). Thus, it is convenient to rewrite the model (28) as follows, hiding the structure of the interference due to other users and noise in a single vector $\mathbf{i}[n]$:

$$\begin{aligned}\mathbf{r}[n] &= b_1[n]\mathbf{u}_1[n] + \mathbf{i}[n] \\ &= b_1[n]\sum_{i=1}^{M_1} F_{1,i}[n]\mathbf{v}_{1,i} + \mathbf{i}[n].\end{aligned} \tag{29}$$

We have seen in Section II that the DMMSE-based algorithm avoids tracking the fading gain for a single path by the use of differential demodulation. However, if the desired user undergoes multipath fading, in order to automatically combine two paths, the DMMSE-based algorithms must track the time-varying linear combination $F_{1,1}[n]\mathbf{v}_{1,1} + F_{1,2}[n]\mathbf{v}_{1,2}$, which amounts to tracking the relative complex gain $F_{1,2}[n]/F_{1,1}[n]$ with the single-path techniques of Section III. This imposes a limit on the automatic multipath combining capability of the DMMSE-based algorithm as the fading rate increases [6]. It is necessary, therefore, to extend the basic DMMSE criterion to a multipath setting.

The idea is to convert the frequency-selective fading channel into several parallel frequency-nonselective fading subchannels, apply the basic DMMSE algorithm for obtaining an interference-suppressing correlator for each subchannel, and to then noncoherently combine the correlator outputs for each subchannel to obtain the decision statistic. Further, the preceding should be accomplished without knowledge of the subchannels. Therefore, we describe next a DMMSE-based approach, termed the *eigenrake* receiver, which achieves interference suppression and diversity combining without requiring explicit information regarding the multipath fading gains or timing.

### A. The Eigenrake Receiver

Consider the generalized eigenvalue problem (8) in *Theorem 1*, where the matrices $\mathbf{R}$ and $\mathbf{A}$ are defined as in (9). Let $\{\mathbf{c}_i, 1 \leq i \leq M_e\}$ denote the eigenvectors corresponding to the positive eigenvalues. The eigenrake receiver employs a subset of these as correlators to obtain both interference suppression and diversity. Specifically, application of the $i$th correlator yields the decision statistic

$$Z_i[n] = \langle \mathbf{c}_i, \mathbf{r}[n]\rangle = \hat{F}_{1,i}[n]b_1[n] + x_i[n] \tag{30}$$

where

$$\hat{F}_{1,i}[n] = \langle \mathbf{c}_i, \mathbf{u}_1[n]\rangle \tag{31}$$

is the *effective* fading gain on the $i$th subchannel, and $x_i[n]$ is the residual interference plus noise at the output of the $i$th correlator. The outputs of these subchannels are then combined to generate the following decision statistic for differential demodulation:

$$Z[n] = \sum_{i=1}^{M_e} \alpha_i \langle \mathbf{c}_i, \mathbf{r}[n]\rangle \langle \mathbf{c}_i, \mathbf{r}[n-1]\rangle^* \tag{32}$$

where the $\alpha_i \geq 0$ are combining gains. For differentially encoded data, this statistic is fed to a slicer. For example, for binary DPSK, the bit estimates are given by

$$\hat{a}_i[n] = \text{sign}\left(\Re\{Z[n]\}\right). \tag{33}$$

We discuss the structure and properties of the eigenrake receiver in the next section.

### B. Structure of Eigenrake Receiver

For the signal model (29), assuming that the fading gains for different multipath components of the desired user are uncorrelated, we have

$$\mathbf{R} = \sum_{i=1}^{M_1} \gamma_i \mathbf{v}_{1,i}\mathbf{v}_{1,i}^H + \mathbf{R}_I \tag{34}$$

$$\mathbf{A} = 2\sum_{i=1}^{M_1} \Re\{\rho_i\}\mathbf{v}_{1,i}\mathbf{v}_{1,i}^H \tag{35}$$

where, for the $i$th multipath component of the desired user (with $M_1$ total multipath components), $\gamma_i = E[|F_{1,i}[n]|^2]$ denotes the average strength, and $\rho_i = E[F_{1,i}[n]F_{1,i}^*[n-1]]$ denotes the correlation between the fading gains in successive symbol intervals. As was done for the flat-fading environment in *Remark 4* and Section II-C, it is useful to develop the frequency-selective fading notation in the whitened domain for the derivation of the eigenrake receiver and its properties.

*Whitened Domain:* Using (15)–(18), as well as the whitened representation of $\mathbf{A}$ given by

$$\tilde{\mathbf{A}} = \mathbf{R}^{-\frac{1}{2}}\mathbf{A}\mathbf{R}^{-\frac{1}{2}}$$

the corresponding eigenvalue problem (8) becomes

$$\tilde{\mathbf{A}}\tilde{\mathbf{c}} = \lambda\tilde{\mathbf{R}}\tilde{\mathbf{c}} = \lambda\tilde{\mathbf{c}}. \tag{36}$$

Note that there is a one-to-one correspondence between the eigenvectors of (8) and (36): $\mathbf{c}$ is an eigenvector of (8) if and only if $\tilde{\mathbf{c}} = \mathbf{R}^{\frac{1}{2}}\mathbf{c}$ is an eigenvector of (36) with the same eigenvalue. For $1 \le i \le M_e$, let $\tilde{\mathbf{c}}_i = \mathbf{R}^{\frac{1}{2}}\mathbf{c}_i$ denote the eigenvectors corresponding to the positive eigenvalues for the whitened problem (36). Thus, the decision statistics in both domains are identical; that is

$$\langle \mathbf{c}, \mathbf{r}[n] \rangle = \langle \tilde{\mathbf{c}}, \tilde{\mathbf{r}}[n] \rangle .$$

In the whitened domain, the received signal model (29) can be rewritten as

$$\begin{aligned}\tilde{\mathbf{r}}[n] &= b_1[n]\tilde{\mathbf{u}}_1[n] + \tilde{\mathbf{i}}[n] \\ &= b_1[n]\sum_{i=1}^{M_1} F_{1,i}[n]\tilde{\mathbf{v}}_{1,i} + \tilde{\mathbf{i}}[n]\end{aligned} \tag{37}$$

where $\tilde{\mathbf{v}}_{1,i} = \mathbf{R}^{-1/2}\mathbf{v}_{1,i}$, $1 \le i \le M_1$.

We are now ready to formally state the properties of the eigenrake receiver in the form of the following theorem.

*Theorem 3 (Eigenrake Receiver):* The eigenvectors $\{\mathbf{c}_i, 1 \le i \le M_e\}$, satisfy the following properties.

1) The number of branches in the eigenrake receiver is at most equal to the number of multipath components for the desired user; that is, $M_e \le M_1$.
2) For each $1 \le i \le M_e$, the correlator $\mathbf{c}_i$ is an interference-suppressing, near–far resistant linear receiver, providing an estimate of the desired symbol sequence (up to complex scaling).
3) As long as $\Re\{\rho_i\}$ are approximately equal for all $i$, the effective fading gains $\hat{F}_{1,i}[n]$ are approximately uncorrelated for different $i$, $1 \le i \le M_e$. That is, the eigenrake receiver provides $M_e$-fold diversity.

*Proof:* Using (34) and (35), we obtain that

$$\mathbf{R}^{-1}\mathbf{A} = \sum_{i=1}^{M_1} \mathbf{d}_i \mathbf{v}_{1,i}^H$$

where

$$\begin{aligned}\mathbf{d}_i &= 2\Re\{\rho_i\}\mathbf{R}^{-1}\mathbf{v}_{1,i} \\ &= 2\Re\{\rho_i\}\left(\mathbf{R}_I + \sum_{j=1}^{M_1}\gamma_j \mathbf{v}_{1,j}\mathbf{v}_{1,j}^H\right)^{-1}\mathbf{v}_{1,i}.\end{aligned} \tag{38}$$

The form of (38) is that of an MMSE correlator for which the desired signal is the $i$th multipath component of the desired user, and the interference corresponds to the interference due to other users, as well as the other multipath components of the desired user. Thus, $\mathbf{d}_i$ inherits the classical interference-suppression properties of the MMSE correlator (including its near–far resistance).

Now, consider an eigenvector $\mathbf{c}$ of $\mathbf{R}^{-1}\mathbf{A}$ corresponding to a nonzero eigenvalue $\lambda$. Such an eigenvector must satisfy

$$\begin{aligned}\mathbf{c} &= \frac{1}{\lambda}\mathbf{R}^{-1}\mathbf{A}\mathbf{c} \\ &= \frac{1}{\lambda}\sum_{i=1}^{M_1}\left(\mathbf{v}_{1,i}^H\mathbf{c}\right)\mathbf{d}_i\end{aligned}$$

so that $\mathbf{c}$ is a linear combination of the interference-suppressing correlators $\{\mathbf{d}_i, 1 \le i \le M_1\}$. Hence, $\mathbf{c}$ also suppresses interference, is near–far resistant, and produces a scaled version of the desired symbol sequence. Further, since each eigenvector corresponding to a positive eigenvalue must be a linear combination of $\{\mathbf{d}_i, 1 \le i \le M_1\}$, the number of such vectors that can be linearly independent is at most $M_1$. Thus, the number of eigenvectors $M_e$ is at most $M_1$. This completes the proof of properties 1) and 2).

To show diversity, it is easier to work in the whitened domain, where (35) reduces to

$$\tilde{\mathbf{A}} = 2\sum_{i=1}^{M_1}\Re\{\rho_i\}\tilde{\mathbf{v}}_{1,i}\tilde{\mathbf{v}}_{1,i}^H.$$

Comparing this with the covariance matrix for the whitened desired signal vector, given by

$$\begin{aligned}\tilde{\mathbf{R}}_d &= E\left[\tilde{\mathbf{u}}_1[n]\tilde{\mathbf{u}}_1^H[n]\right] \\ &= \sum_{i=1}^{M_1} E\left[|F_{1,i}[n]|^2\right]\tilde{\mathbf{v}}_{1,i}\tilde{\mathbf{v}}_{1,i}^H\end{aligned}$$

we see that as long as $\Re\{\rho_i\}$ is approximately the same for each $i$, $\tilde{\mathbf{A}}$ is approximately a scalar multiple of $\tilde{\mathbf{R}}_d$. Thus, the eigenvectors of $\tilde{\mathbf{A}}$ provide an approximate Karhunen–Loeve (KL) decomposition of the whitened desired signal vector $\tilde{\mathbf{u}}_1[n]$, which implies that the effective fading gains along the directions of these eigenvectors are approximately uncorrelated (and hence, approximately independent, if the fading coefficients are jointly complex Gaussian). This proves property 3), and completes the proof of the theorem. □

*Remark 6:* The condition that $\Re\{\rho_i\}$ be approximately the same for all $i$ is satisfied in practice, since $\Re\{\rho_i\} \approx 1$ for typical fading rates (the fading gains for a given multipath component are roughly equal across successive symbols).

*Remark 7:* The eigenrake receiver achieves implicit timing acquisition, interference suppression, and diversity, providing a KL decomposition of the faded signal vector in the whitened domain, without requiring explicit estimation of the location or strengths of the multipath components for the desired user.

*Remark 8:* In order to get the full performance benefit from the eigenrake receiver, it is important to employ correlators that coincide with the eigenvectors with positive eigenvalues. Simply choosing correlators lying in the subspace spanned by these eigenvectors does not work as well. Thus, an application of subspace-tracking methods, such as [29]–[32] (or the orthogonal iteration of [8] and [24]), will, at best, yield results comparable to the eigenrake receiver.

### C. Combining Rule

It remains to specify the combining gains $\{\alpha_i\}$ in (32). If the number of multipaths of the desired user $M_1$ were known, then by *Theorem 3*, the eigenrake receiver should use at most the $M_1$ eigenvectors corresponding to the $M_1$ largest eigenvalues of (8) for demodulation in (32). Among these correlators may be some that should be deselected, because their output is of poor quality (due to low signal strength or bad interference patterns for the corresponding multipath component). However, the eigenrake receiver has no prior knowledge of the number of multipath components or the quality of the corresponding correlators, and employs instead a selection strategy based on the eigenvalues of (8) to choose which correlators to use. Once this set is selected, we have found by experimentation that equal-gain combining is the most effective approach. Specifically, the combining rule we use is

$$\alpha_i = \begin{cases} 1, & \text{if } \lambda_i \geq t_0 \\ 0, & \text{if } \lambda_i < t_0. \end{cases}$$

To choose the threshold $t_0$, consider the DMMSE constrained cost function $J(\mathbf{c})$ in (6). If a correlator $\mathbf{c}_i$ is working well, then we must have $J(\mathbf{c}_i) \approx 0$. But, from (11), $J(\mathbf{c}_i) = 2\mathbf{c}_i^H \mathbf{R} \mathbf{c}_i - \mathbf{c}_i^H \mathbf{A} \mathbf{c}_i = 2 - \lambda_i$, which means that $\lambda_i \approx 2$ for a correlator that is producing a good reproduction of the desired symbol sequence. From our numerical results, we find that $t_0 = 1$, which balances the tradeoff between false indication (i.e., incorrectly indicating a specific path is present) and failed detection (i.e., not detecting the presence of a particular path), works well.

Alternatively, it is possible to optimize the combining gains $\alpha_i$, based on an estimate of the signal-to-interference ratio (SIR) on each branch. Denote the desired signal power at the output of the $i$th correlator by $P_i$, and interference-plus-noise power by $\sigma_i^2$. At the output of correlator $\mathbf{c}_i$, the average power of the net received signal is $\mathbf{c}_i^H \mathbf{R} \mathbf{c}_i = 1$ (by virtue of the normalization we have imposed). The average power of the desired signal is $P_i = E[|\mathbf{c}_i^H \mathbf{u}_1[n]|^2] \approx (\frac{1}{2})\mathbf{c}_i^H \mathbf{A} \mathbf{c}_i = \lambda_i/2$, assuming that $\Re\{\rho_i\} \approx 1$. Hence, the averaged interference-plus-noise power is $\sigma_i^2 = 1 - (\lambda_i/2)$. It can be shown that the maximal ratio combining coefficients $\alpha_i$ in (32) should be set as $\alpha_i^2 = P_i/\sigma_i^4$ (i.e., the same as classically defined in [33] and recently reissued in [34]), approximating the outputs of the fingers of the eigenrake as independent, differentially coherent Rayleigh fading channels with Gaussian noise. However, based on our simulations, such optimization does not improve upon the simpler equal gain, selective combining strategy described earlier.

### D. On the Implementation Complexity of DMMSE

For the single-correlator DMMSE algorithms given in Section III, the implementation is of the same order of complexity as standard MMSE correlators. The scaling in the denominator of (22) is the only difference between the scaled RLS recursion and standard RLS adaptations [21]. For block power updates, the complexity is dominated by the computation of $\hat{\mathbf{R}}^{-1}$, since the power algorithm in (27) usually converges in a few iterations (details of the convergence analysis can be found in [24]). Thus, block power updates are comparable in complexity to standard least squares solutions for (2). However, obtaining up to $M_e$ eigenvectors for the eigenrake receiver generally requires more complex approaches to solving the eigenvalue problem in (8). One approach [24] applies the QZ algorithm, resulting in $O(N^3)$ computational complexity. In [35], a tight bound on this problem was shown to be $O\big(N^3 + (N \log^2 N) \log b\big)$ when the relative error is bounded by $2^{-b}$. Alternatively, an "online" recursion, such as in [36], may be applied.

## V. Numerical Results

For comparison among the receivers proposed in this paper and other standard approaches, we simulate several MMSE and DMMSE receiver implementations as well as a standard Rake receiver. For DMMSE reception, results are obtained using the following methods: block power update of Section III-B; eigenrake of Section IV-A with knowledge of the number of desired user multipaths present, as well as without multipath knowledge, but using the selective-combining rule of Section IV-C; scaled RLS, as described in Section III-A.1; and the unscaled RLS from Section III-A.2. Both unscaled and scaled RLS DMMSE receivers are simulated using differential decoding where the decoded symbol estimates are given by $\hat{a}_1[n] = \text{sign}\big(\Re\{\langle \mathbf{c}[n], \mathbf{r}[n]\rangle \langle \mathbf{c}[n-1], \mathbf{r}[n-1]\rangle^*\}\big)$. Additionally, to demonstrate the standard MMSE receiver in (2), a standard RLS MMSE algorithm [21] is simulated using direct decoding, where the recursive correlator updates are computed via

$$\begin{aligned} \mathbf{k}[n] &= \frac{\mathbf{P}[n-1]\mathbf{r}[n]}{\beta + \mathbf{r}^H[n]\mathbf{P}[n-1]\mathbf{r}[n]} \\ \mathbf{P}[n] &= \beta^{-1}\mathbf{P}[n-1] - \beta^{-1}\mathbf{k}[n]\mathbf{r}^H[n]\mathbf{P}[n-1] \\ \xi[n] &= b_1[n] - \mathbf{c}^H[n-1]\mathbf{r}[n] \\ \mathbf{c}[n] &= \mathbf{c}[n-1] + \beta_a \mathbf{k}[n]\xi^*[n] \end{aligned}$$

and the decoded symbol estimates are computed as $\hat{b}_1[n] = \text{sign}(\Re\{\langle \mathbf{c}[n-1], \mathbf{r}[n]\rangle\})$. For benchmarks, included are results obtained using instantaneous, ideal versions of the MMSE solution and standard Rake. Both of these techniques assume perfect knowledge of all the users' spreading waveforms and fading gains, and are computed at every symbol time. The instantaneous, ideal MMSE solution is computed as

$$\mathbf{c}[n] = \mathbf{R}_I^{-1}[n] \left( \sum_{i=1}^{M_1} F_{1,i}[n] \mathbf{v}_{1,i} \right) \tag{39}$$

where $\mathbf{R}_I[n]$ is computed as in (14), appropriately modified for multipath and perfect channel knowledge

$$\mathbf{R}_I[n] = \sum_{k=2}^{K} \sum_{i=1}^{M_k} |F_{k,i}[n]|^2 \, \mathbf{v}_{k,i} \mathbf{v}_{k,i}^H + 2\sigma^2 \mathbf{I}.$$

The instantaneous, ideal Rake solution is computed as

$$\mathbf{c}[n] = \sum_{i=1}^{M_1} F_{1,i}[n] \mathbf{v}_{1,i}. \tag{40}$$

In both cases, the symbols are differentially decoded via $\hat{a}_1[n] = \text{sign}\big(\Re\{\langle \mathbf{c}[n], \mathbf{r}[n]\rangle \langle \mathbf{c}[n], \mathbf{r}[n-1]\rangle^*\}\big)$.

For all results reported here, we consider a synchronous CDMA system, given by (1) for flat fading, and (34) for multipath fading. The processing gain $N = 10$, with $K = 1$ and

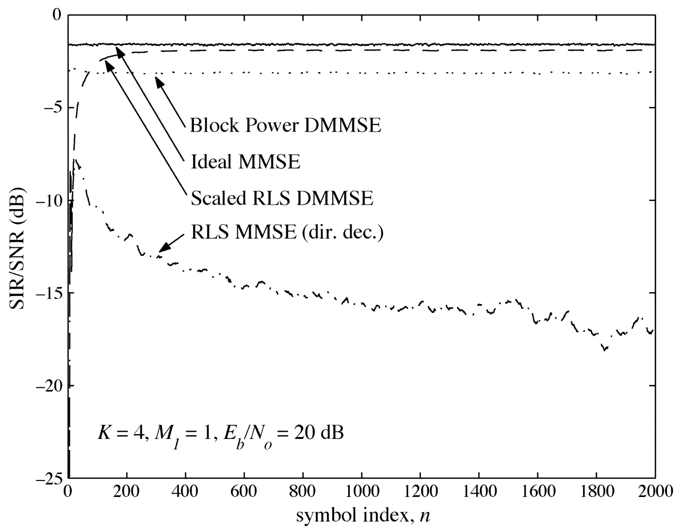


Fig. 4. SIR/SNR averaged over 250 adaptations for $K = 4$ and a flat-fading channel using ideal MMSE, standard RLS MMSE (with direct decoding), block power DMMSE, and scaled RLS DMMSE. The training period is $T = 60$ symbols, and the RLS adaptations switch to decision-directed mode after training.

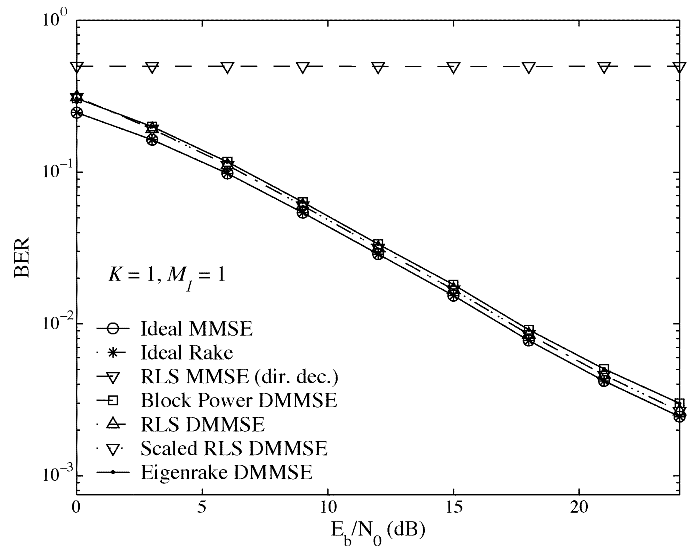


Fig. 5. BER for $K = 1$ user in a flat-fading channel using ideal MMSE and Rake, standard RLS MMSE (with direct decoding), block power DMMSE, unscaled and scaled RLS DMMSE, and eigenrake DMMSE receivers. The training period is $T = 100$ symbols, and the RLS adaptations switch to decision-directed mode after training.

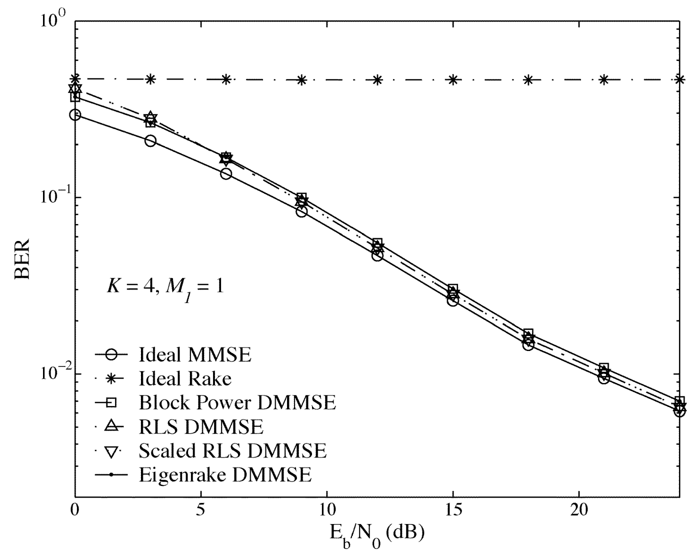


Fig. 6. BER for $K = 4$ users in a flat-fading channel using ideal MMSE and Rake, block power DMMSE, unscaled and scaled RLS DMMSE, and eigenrake DMMSE receivers. The training period is $T = 100$ symbols, and the RLS adaptations switch to decision-directed mode after training.

four users, each with a fixed but randomly chosen spreading sequence. The Rayleigh fading coefficients are generated independently for different paths and users using a modified Jakes simulator [37]–[39], with a normalized Doppler spread of 0.01 (see the discussion in Section II of parameters that might correspond to such a system). When there are multiple users present, each interfering user has average power 20 dB higher than that of the desired user. We thereby demonstrate that DMMSE interference suppression is robust under severe near–far conditions, which is to be expected, given the theoretical results on near–far resistance proven in Section IV-B.

First, we consider frequency-nonselective fading for each user. Since the signal-to-noise ratio (SNR) for the desired transmission varies with time due to fading, the ability of an adaptive algorithm to track the channel and to suppress interference is gauged by the difference between the SIR and SNR, rather than the raw value of SIR [6]. Thus, with the AWGN power set so that the desired user's average $E_b/N_0 = 20$ dB, we plot SIR/SNR (dB) averaged over 250 adaptations in Fig. 4 for four schemes: ideal MMSE, standard RLS MMSE with direct decoding, block power DMMSE, and scaled RLS DMMSE. The algorithms have a training length of $T = 60$ symbols, after which the block power receiver is fixed, and both the standard and scaled RLS receivers switch to decision-directed mode. Both RLS receivers were simulated with $\beta = \beta_a = 0.99$. Clearly, the standard RLS MMSE algorithm with direct decoding cannot track the fading at all, and actually begins to diverge immediately after the training period. The block power DMMSE algorithm is able to track the desired user's signal, but at a level 1.5–2 dB worse than the scaled RLS DMMSE algorithm. In Figs. 5 and 6, we plot the average BER for $K =$ 1 and 4, respectively, over 2000 simulations at each $E_b/N_0$ for six schemes: ideal MMSE and Rake, block power DMMSE, standard RLS MMSE, unscaled and scaled RLS DMMSE, and eigenrake DMMSE (using only one correlator). For the simulations, the training length was $T = 100$ symbols per packet, and the packet length was 2000 symbols. In Fig. 5 with $K = 1$, all of the receivers perform similarly well, except for the standard RLS MMSE with direct decoding. The ideal Rake and MMSE receivers match exactly, since there are no interferers. Further, the block power and eigenrake receivers match exactly, since they are equivalent when there is no multipath or MAI, and are only slightly worse than the ideal MMSE receiver. Both the unscaled and scaled RLS DMMSE receivers perform just in the gap (and the same, as discussed in Section III-A.2) between the eigenrake and ideal MMSE, while the standard RLS MMSE fails completely to track the channel-fading gains. In the presence of other users $(K = 4)$ in Fig. 6, the results are similar to those for the single-path, no-interferer case with one exception:

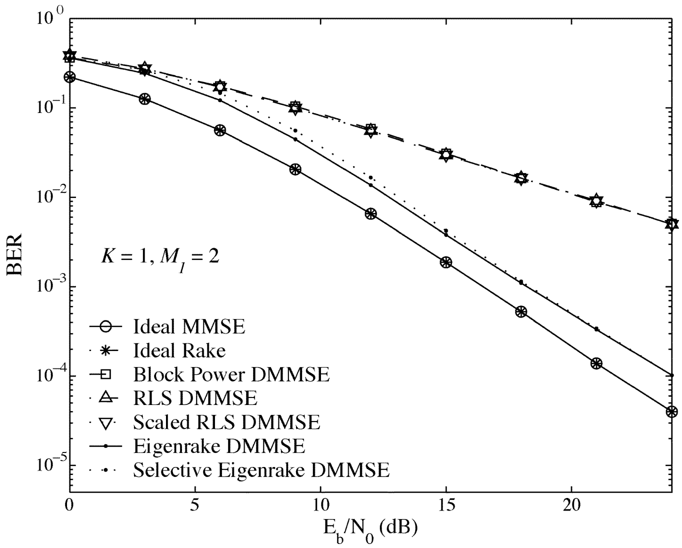


Fig. 7. BER for $K = 1$ user with $M_1 = 2$ multipath components using ideal MMSE and Rake, block power DMMSE, unscaled and scaled RLS DMMSE, and eigenrake and selective eigenrake DMMSE receivers. The training period is $T = 100$ symbols, and the RLS adaptations switch to decision-directed mode after training.

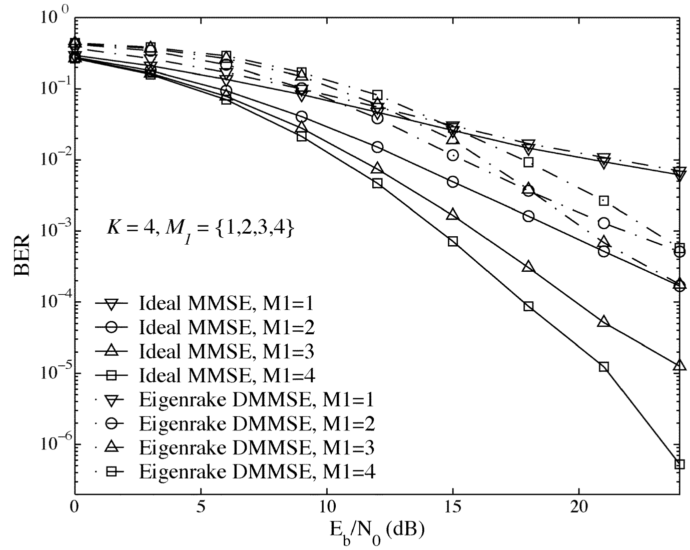


Fig. 9. BER for $K = 4$ users with $M_1 = \{1, 2, 3, 4\}$ multipath components for the desired user using the ideal MMSE and the eigenrake DMMSE receivers, with the number of eigenrake correlators fixed as the number of multipaths for the desired user.

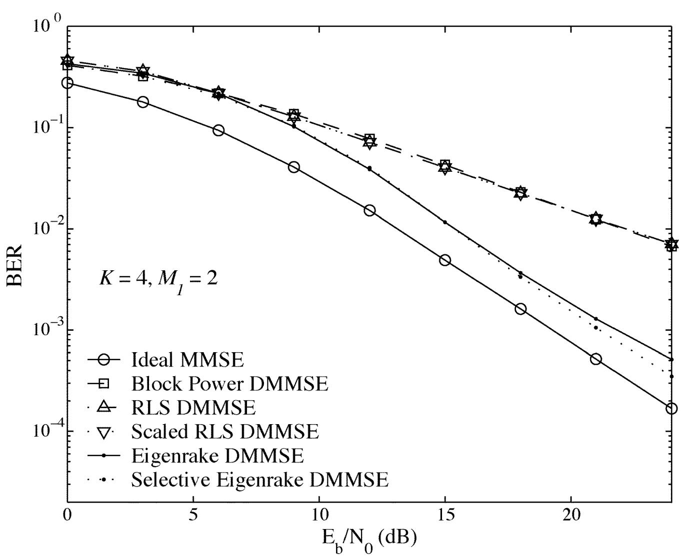


Fig. 8. BER for $K = 4$ users with $M_1 = 2$ multipath components for the desired user using ideal MMSE, block power DMMSE, unscaled and scaled RLS DMMSE, and eigenrake and selective eigenrake DMMSE receivers. The training period is $T = 100$ symbols, and the RLS adaptations switch to decision-directed mode after training.

the ideal Rake fails completely due to MAI. The MAI causes an approximately 3-dB decrease in performance for the group.

Next, we consider a situation in which the desired user has two independently faded multipath components, while each interferer still sees a single-path fading channel. Average BER is plotted in Figs. 7 and 8 for the case of $M_1 = 2$ multipath components for the desired user, and 2000 symbol packets with training length $T = 100$ symbols. An additional curve is shown in these BER plots for the multipath case: the selective eigenrake DMMSE implementation is tested (i.e., the combining rule of Section IV-C is used). For $K = 1$, Fig. 7 shows the diversity gains obtained by the eigenrake and selective eigenrake algorithms, while the single-correlator implementations of the other techniques grow progressively worse in performance as the SNR improves. This is expected, since the single-correlator techniques actually experience the extra paths of the desired user as interference. Here again, the ideal Rake and MMSE receivers are equivalent, but only approximately 3 dB better than the eigenrake receivers operating with knowledge only of the desired user's training sequence. In Fig. 8 for $K = 4$, we see that both the standard and selective eigenrake receivers are robust to strong MAI (the ideal Rake receiver is omitted, due to its poor performance). At higher $E_b/N_0$, when the effect of MAI is dominant, adaptive selection of eigenrake correlators is 0.5–1 dB better than fixing the number of correlators to equal the number of paths. We attribute this to the ability of the selective eigenrake to adaptively "deselect" paths suffering too great a level of interference.

Finally, Figs. 9 and 10 show the BER for the eigenrake receiver (with equal gain combining) as the number of correlators and the number of multipath components for the desired user are varied, without recourse to an adaptive selection mechanism. There are three strong interferers $(K = 4)$, each with a single path. In Fig. 9, $M_1$ is varied from 1 to 4, with the number of correlators always set as $N_{Ci} = M_1$. The performance of the ideal MMSE receiver improves monotonically with $M_1$, while the performance of the eigenrake improves until $M_1 = 3$, but degrades for $M_1 = 4$. This is probably due to the fact that as the number of paths increases with $N = 10$ fixed, the likelihood of one of the eigenrake correlators seeing a bad crosscorrelation pattern also increases. On the other hand, the ideal MMSE receiver sees a single effective spreading waveform, regardless of the number of paths. In Fig. 10, we fix $M_1 = 3$, and vary $N_{Ci}$ from 1 to 4. The performance improves as $N_{Ci}$ increases, until $N_{Ci} = M_1 = 3$. However, when $N_{Ci} = 4 > M_1$, the additional correlator cannot be linearly independent by *Theorem 3*, and the performance degrades significantly, compared with that for $N_{Ci} = 2, 3$, while still showing diversity gains relative to a single correlator. We conclude, therefore, that adaptive selection

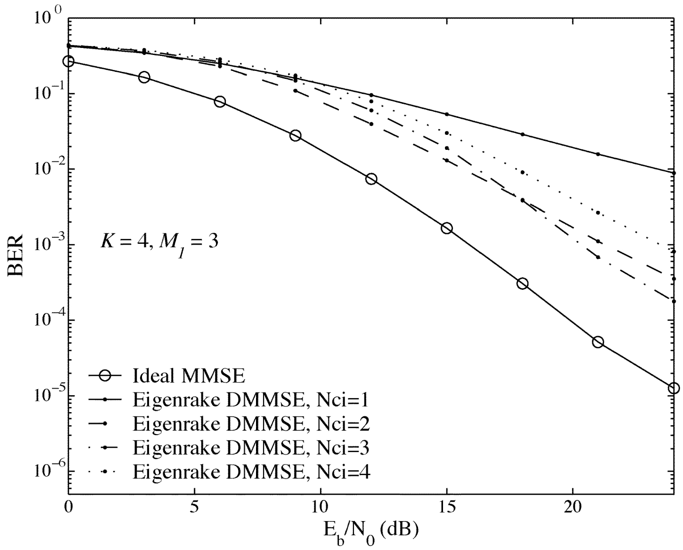


Fig. 10. BER for $K = 4$ users with $M_1 = 3$ multipath components for the desired user using the eigenrake DMMSE receiver with $N_{Ci} = \{1, 2, 3, 4\}$ correlators for detection, with the ideal MMSE solution included for comparison.

of the number of correlators, as done by the selective eigenrake, is crucial to balancing the effects of diversity gains and residual interference, and that mismatch between the number of correlators and paths severely degrades performance.

## VI. Conclusions

We have shown that the DMMSE criterion leads to adaptive linear receivers which are robust to rapid channel time variations, unlike adaptive algorithms based on the conventional MMSE criterion. The near–far resistance and interference-suppression properties of the resulting DMMSE solution are shown by establishing equivalence with a channel-compensated MMSE solution. For frequency-selective fading, the DMMSE criterion is extended to obtain the *eigenrake* receiver, which provides implicit timing acquisition, diversity, and interference suppression.

In addition to its robustness to fading, DMMSE reception is also robust to lack of carrier synchronization, and is, therefore, attractive for packetized transmission in rapidly varying network topologies. For example, it has recently been employed in cross-layer design of medium-access-control mechanisms in wireless networks, where it has proven effective in permitting a number of rapidly moving terminals to randomly access an access point employing a DMMSE receiver [40], [41], [47].

Beyond the initial exposition of the DMMSE criterion in this paper, much further work remains on detailed receiver design, comparison with other approaches, and development of efficient numerical techniques. For example, it is necessary to resolve practical issues, such as whether to use noncoherent techniques in conjunction with DMMSE, possibly with multiple symbol detection [26], [27], or whether to use DMMSE for robust interference suppression, followed by separate channel-gain recovery for coherent detection. Another possible direction for future work is integration of DMMSE with sophisticated coding techniques, such as those used for turbo multiuser detection [42], [43], and single-user noncoherent communication based on joint channel and data estimation [44], [45].

While we consider short spreading sequences in this paper, a novel interpretation of the DMMSE criterion for systems with long spreading sequences yields rapidly converging adaptive beamformers that do not require overhead in terms of training symbols [9], [46]. A detailed exploration of the practical implications of this (e.g., to commercial cellular CDMA systems) is an important topic for future work.

## References

[1] M. Abdulrahman, A. U. H. Sheikh, and D. D. Falconer, "Decision feedback equalization for CDMA in indoor wireless communication," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 4, pp. 698–704, May 1994.

[2] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.

[3] S. L. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Commun.*, vol. 43, no. 2–4, pp. 1746–1755, Feb.-Apr. 1995.

[4] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 4, pp. 685–697, May 1994.

[5] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 944–960, Jul. 1995.

[6] L. J. Zhu and U. Madhow, "Adaptive interference suppression for direct sequence CDMA over severely time-varying channels," in *Proc. IEEE Global Telecommun. Conf.*, vol. 2, Phoenix, AZ, Nov. 1997, pp. 917–922.

[7] U. Madhow, "Blind adaptive interference suppression for the near–far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 124–136, Jan. 1997.

[8] L. J. Zhu, "Differential MMSE: A new approach for adaptive interference suppression over time-varying channels," Ph.D. dissertation, Univ. Illinois, Urbana-Champaign, Urbana, IL, 1999.

[9] U. Madhow, L. J. Zhu, and L. Galup, "Differential MMSE: New adaptive algorithms for equalization, interference suppression, and beamforming," in *Proc. 32nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1998, pp. 640–644.

[10] M. L. Honig, M. J. Shensa, S. L. Miller, and L. B. Milstein, "Performance of adaptive linear interference suppression for DS-CDMA in the presence of flat Rayleigh fading," in *Proc. IEEE Veh. Technol. Conf*, vol. 3, Phoenix, AZ, May 1997, pp. 2191–2195.

[11] R. Schober, W. H. Gerstacker, and A. Lampe, "Noncoherent MMSE interference suppression for DS-CDMA," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 577–587, Apr. 2002.

[12] H. V. Poor and X. Wang, "Adaptive multiuser detection in fading channels," in *Proc. 34th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Oct. 1996, pp. 603–612.

[13] A. N. Barbosa and S. L. Miller, "Adaptive detection of DS/CDMA signals in fading channels," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 115–124, Jan. 1998.

[14] M. Latva-aho and M. J. Juntti, "LMMSE detection for DS-CDMA systems in fading channels," *IEEE. Trans. Commun.*, vol. 48, no. 2, pp. 194–199, Feb. 2000.

[15] R. Fantacci, S. Morosi, and M. Bonechi, "Adaptive MMSE receivers for communications in non-stationary multipath fading channel," in *Proc. IEEE 6th Int. Symp. Spread Spectrum Tech., Applic.*, vol. 2, Parsippany, NJ, Sep. 2000, pp. 545–549.

[16] M. K. Varanasi, "Noncoherent detection in asynchronous multiuser channels," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 157–176, Jan. 1993.

[17] E. Visotsky and U. Madhow, "Noncoherent multiuser detection for CDMA systems with nonlinear modulation: A non-Bayesian approach," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1352–1367, May 2001.

[18] M. Brehler and M. K. Varanasi, "Optimum multiuser noncoherent DPSK detection in generalized diversity Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 6, pp. 1565–1574, Jun. 2003.

[19] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.

[20] U. Madhow, "Blind adaptive interference suppression for direct-sequence CDMA," *Proc. IEEE*, vol. 86, no. 10, pp. 2049–2069, Oct. 1998.

[21] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[22] S. L. Miller, M. L. Honig, and L. B. Milstein, “Performance analysis of MMSE receivers for DS-CDMA in frequency-selective fading channels,” *IEEE Trans. Commun.*, vol. 48, no. 11, pp. 1919–1929, Nov. 2000.

[23] M. L. Honig, S. L. Miller, M. J. Shensa, and L. B. Milstein, “Performance of adaptive linear interference suppression in the presence of dynamic fading,” *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 635–645, Apr. 2001.

[24] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[25] L. J. Zhu and U. Madhow, “MMSE interference suppression for rapidly faded CDMA systems: A new formulation for differentially modulated signals,” in *Proc. IEEE Int. Symp. Inf. Theory*. Cambridge, MA, Aug. 1998, p. 182.

[26] D. Divsalar and M. K. Simon, “Multiple-symbol differential detection of MPSK,” *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

[27] D. Warrier and U. Madhow, “Noncoherent communication in space and time,” in *Proc. 33rd Conf. Inf. Sci. Syst*. Baltimore, MD, Mar. 1999, [CD-ROM].

[28] L. J. Zhu and U. Madhow, “Adaptive interference suppression for DS-CDMA over a Rayleigh fading channel,” in *Proc. 31st Conf. Inf. Sci. Syst*. Baltimore, MD, Mar. 1997, [CD-ROM].

[29] B. Yang, “Projection approximation subspace tracking,” *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, Jan. 1995.

[30] X. Wang and H. V. Poor, “Blind multiuser detection: A subspace approach,” *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 677–690, Mar. 1998.

[31] ——, “Blind adaptive multiuser detection in multipath CDMA channels based on subspace tracking,” *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3030–3044, Nov. 1998.

[32] S. Roy, “Subspace blind adaptive detection for multiuser CDMA,” *IEEE Trans. Commun.*, vol. 48, no. 1, pp. 169–175, Jan. 2000.

[33] D. G. Brennan, “Linear diversity combining techniques,” *Proc. IRE*, vol. 47, no. 6, pp. 1075–1102, Jun. 1959.

[34] ——, “Linear diversity combining techniques,” *Proc. IEEE*, vol. 91, no. 2, pp. 331–356, Feb. 2003.

[35] V. Y. Pan and Z. Q. Chen, “The complexity of the matrix eigenproblem,” in *Proc. 31st ACM Symp. Theory Comput.*, Atlanta, GA, May 1999, pp. 507–516.

[36] Y. N. Rao, J. C. Principe, and T. F. Wong, “Fast RLS-like algorithm for generalized eigendecomposition and its applications,” *J. VLSI Signal Process.*, vol. 37, no. 2, pp. 333–344, Jun. 2004.

[37] W. C. Jakes, *Microwave Mobile Communications*. New York: IEEE Press, 1994.

[38] M. F. Pop and N. C. Beaulieu, “Limitations of sum-of-sinusoids fading channel simulators,” *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 699–708, Apr. 2001.

[39] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, “Statistical simulation models for Rayleigh and Rician fading,” in *Proc. IEEE Int. Conf. Commun.*, vol. 5, Anchorage, AK, May 2003, pp. 3524–3529.

[40] K. Bruvold and U. Madhow, “Adaptive multiuser detection for mobile-centric fast handoffs in pseudocellular wireless networks,” in *Proc. 58th IEEE Veh. Technol. Conf.*, vol. 2, Orlando, FL, Oct. 2003, pp. 796–800.

[41] ——, “Cross-layer optimization of the reservation channel in a pseudo-cellular network: Mobile-centric fast handoffs via multiuser detection,” in *Proc. 37th Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Pacific Grove, CA, Nov. 2003, pp. 1733–1737.

[42] X. Wang and H. V. Poor, “Iterative (turbo) soft interference cancellation and decoding for coded CDMA,” *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.

[43] S. Parkvall, E. G. Ström, L. B. Milstein, and B. E. Ottersten, “Asynchronous near–far resistant DS-CDMA receivers without *a priori* synchronization,” *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 78–88, Jan. 1999.

[44] H.-J. Su and E. Geraniotis, “Low-complexity joint channel estimation and decoding for pilot symbol-assisted modulation and multiple differential detection systems with correlated Rayleigh fading,” *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 249–261, Feb. 2002.

[45] R. R. Chen, R. Koetter, U. Madhow, and D. Agrawal, “Joint noncoherent demodulation and decoding for the block fading channel: A practical framework for approaching Shannon capacity,” *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1676–1689, Oct. 2003.

[46] L. Galup and U. Madhow, “Blind spatial interference suppression of DS-CDMA with long spreading sequences,” in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, Aug. 1998, p. 312.

[47] K. Bruvold, R. Mudumbai, and U. Madhow, “A QoS framework for stabilized collision channels with multiuser detection,” in *Proc. IEEE Int. Conf. Commun.*, Seoul, Korea, May 2005.


**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.

**Kristoffer Bruvold** (S’98) received the B.S. degree in electrical and computer engineering from The Ohio State University (OSU), Columbus, in 1998, and the M.S. degree in electrical and computer engineering in 2000 from the University of California, Santa Barbara (UCSB), where he is currently working toward the Ph.D. degree.

Since 1999, he has also been part-time with the Jet Propulsion Laboratory, Pasadena, CA. His current research interests are in cross-layer design for wireless communications and networking.

Mr. Bruvold was a National Merit Scholar while at OSU, and was awarded Teaching Assistant of the Year in both 2000 and 2001 in the Department of Electrical and Computer Engineering at UCSB.

**Liping Julia Zhu**, photograph and biography unavailable at the time of publication.