Gradient Estimation for Sensitivity Analysis and Adaptive Multiuser Interference Rejection in Code-Division Multiple-Access Systems

Narayan B. Mandayam, *Member, IEEE*, and Behnaam Aazhang, *Senior Member, IEEE*

Abstract—In this paper, we consider a direct-sequence code-division multiple-access (DS-CDMA) system in the framework of a discrete-event dynamic system (DEDS) in order to optimize the system performance. Based on this formulation, we develop infinitesimal perturbation analysis (IPA) for estimating the sensitivity of the average probability of bit error to factors ranging from near-far effects to imperfections in power control. The above estimates are shown to be unbiased, and this technique is then further incorporated into a stochastic gradient algorithm for achieving adaptive multiuser interference rejection for such systems, which is also subject to frequency nonselective slow fading. We use an IPA-based stochastic training algorithm for developing an adaptive linear detector with the average probability of error being the minimization criterion. We also develop a practical implementation of such an adaptive detector where we use a joint estimation-detection algorithm for minimizing the average probability of bit error. A sequential implementation that does not require a stochastic training sequence or a preamble is also developed.

Index Terms—Adaptive detection, direct-sequence CDMA, infinitesimal perturbation analysis, stochastic gradient algorithm.

I. INTRODUCTION

CODE-division multiple access (CDMA) is a multiplexing method that effectively allows many users to commonly use a communication channel. Often, the performance criterion of CDMA systems is specified by means of the average probability of bit error of a given user, and is given by

$$\bar{P}_e = E[L(\theta, \Psi)] \quad (1)$$

where θ is some parameter (possibly vector valued) that could be representing some system specification, and Ψ characterizes the randomness in the system. As an example, θ could represent the characteristics of the desired user codes, and Ψ characterizes the random-valued interferers. For optimum performance of such systems, it is required to address opti-

mization problems that involve minimizing the probability of bit error with respect to certain system parameters, i.e.,

$$\min_{\theta} \bar{P}_e = \min_{\theta} E[L(\theta, \Psi)] \quad (2)$$

which may or may not have additional constraints. Under these circumstances, it is imperative to compute the derivatives of the probability of bit error with respect to these parameters. However, due to the multiple-access nature of these systems, the modulation and detection schemes are quite complex, and computing the bit-error probability in itself is a hard problem. Efforts in performance analysis of CDMA systems have concentrated on obtaining asymptotic approximations and bounds for system error probabilities, and as such, the approximations and bounds cannot capture the sensitivities of the system performance. Owing to this analytical intractability, the optimization of such systems (with respect to any class of parameters) presents itself to no analytical solutions.

An alternative and very effective approach that can be applied for addressing this optimization is to use discrete-event simulations [1], [2]. In this paper, we will formulate a CDMA system in the framework of a discrete-event dynamic system (DEDS), and then based on this formulation, use infinitesimal perturbation analysis (IPA) [3, p. 17] for determining the sensitivity of the average probability of bit error to various system parameters. Specifically, we will derive estimates of the sensitivity of the average probability of bit error to near-far effects. We also outline how this technique can be used to derive sensitivity estimates to fading parameters and imperfections in power control. We will then use this method to derive an adaptive single-user detector for a multiuser channel that is also subject to frequency nonselective slow fading.

Conventionally, demodulation of direct-sequence CDMA (DS-CDMA) signals is achieved with a matched filter receiver. These receivers are optimum only when there is no multiple-access interference, and often suffer from the near-far problem where a nearby interferer can disrupt the reception of a highly attenuated desired signal. To mitigate the near-far problem, several multiuser detection schemes have been proposed [4]–[10]; however, these are more complex than the matched filter detector, and require explicit knowledge or estimates of the interference parameters. Other recent interference rejection schemes include the work in [11] and [12]. Another single-user detector that has been proposed for multiuser channels

Paper approved by J. S. Lehnert, the Editor for Modulation and Signal Design of the IEEE Communications Society. Manuscript received July 5, 1994; revised August 19, 1996 and October 23, 1996. This work was supported by the Advanced Technology Program of the Texas Higher Education Coordinating Board under Grant 003604-018. This paper was presented in part at the International Symposium on Information Theory (ISIT'94), Trondheim, Norway, June 1994.

N. B. Mandayam is with WINLAB, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08855-0909 USA (e-mail: narayan@winlab.rutgers.edu).

B. Aazhang is with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77251-1892 USA.

Publisher Item Identifier S 0090-6778(97)05175-1.

is the optimized single-user detector [13] that minimizes the average probability of bit error which requires knowledge of the interference statistics. To overcome these deficiencies, adaptive single-user detectors have been proposed recently [14], [15] that are based on interference suppression using the minimum mean-squared error (MMSE) criterion. These schemes work in a manner identical to equalization schemes used for combating intersymbol interference (ISI). However, in these schemes, the input to the detector is clocked at the chip rate, while the output is sampled at the bit rate. After every bit interval, the taps of the equalizer are adaptively updated to minimize the MSE. These schemes have been shown to be near-far resistant to varying degrees.

In this paper, we will derive a stochastic gradient algorithm that adaptively minimizes the more relevant performance measure, the average probability of bit error for a DS-CDMA system that is also subject to slow-frequency nonselective fading. We will use IPA estimates to adaptively update the algorithm, which requires the desired user to transmit a training sequence or preamble of bits. We will also show that this method yields a simple recursive update structure, and prove the convergence of this method. Further, it is seen that this adaptive detector is near-far resistant to varying degrees, and can be used as a viable detector for the DS-CDMA channels. However, the adaptive detector proposed here has the drawback that it requires extraction of the transmitted signal samples (the sum of the received signal due to the desired user and the interferers) from the received signal (the sum of the transmitted signal and the additive Gaussian noise). The MMSE [14], [15] detector has the advantage that it works directly with the received signal. To overcome this drawback, we develop a practical implementation of the above adaptive detector that incorporates maximum-likelihood estimates of the transmitted signal in an IPA-based stochastic gradient algorithm. Therefore, we develop an adaptive algorithm that is based on a joint estimation-detection strategy at every iteration. Finally, we present a sequential algorithm that does not require a training sequence, but uses bit decisions from preceding bit intervals in successive bit intervals for updating the algorithm.

II. PRELIMINARIES

A DEDS can be represented by the pair (θ, Ψ) , where θ is a parameter (possibly vector) of the DEDS and Ψ is a random vector representing all of the randomness in the system. Typically, the components of Ψ are random variables uniformly, independently and identically distributed on $[0, 1]$. More precisely, Ψ is defined on the underlying probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and is a measurable mapping $\Psi: \Omega \rightarrow [0, 1]^d$ where d could be finite or infinite. For example, in a DS-CDMA system, θ could be the power of a user's transmission and Ψ could be the random vector whose components are the random powers due to interfering users. Each (θ, Ψ) determines a sample path of the system, and a performance measure obtained from such a sample path is denoted as $L(\theta, \Psi)$. For any realization of Ψ , the performance measure $L(\theta, \Psi)$ is a function of θ and is called the sample performance

function. In most cases, we are interested in the expected value of the performance $\bar{P}_e(\theta) = E[L(\theta, \Psi)]$, and sensitivity analysis is concerned with estimating $(\partial/\partial\theta)\bar{P}_e(\theta)$.

In infinitesimal perturbation analysis (IPA) [3, p. 17], the above quantity is estimated as $(\partial/\partial\theta)\bar{P}_e(\theta) = (\partial L(\theta, \Psi)/\partial\theta)$, where the sample derivative calculated from a single sample path is defined as

$$\frac{\partial L(\theta, \Psi)}{\partial\theta} = \lim_{\Delta\theta \rightarrow 0} \frac{L(\theta + \Delta\theta, \Psi) - L(\theta, \Psi)}{\Delta\theta}. \quad (3)$$

The above estimate is said to be unbiased if and only if

$$E\left[\frac{\partial L(\theta, \Psi)}{\partial\theta}\right] = \frac{\partial}{\partial\theta} \bar{P}_e(\theta) = \frac{\partial}{\partial\theta} E[L(\theta, \Psi)]. \quad (4)$$

Thus, the unbiasedness of the sensitivity estimate is equivalent to the interchangeability of the two operators “ E ” and “ $(\partial/\partial\theta)$.” This can be satisfied by imposing some conditions on $L(\theta, \Psi)$ that can be specified using the Lebesgue-dominated convergence theorem [16, pp. 21–34]. In IPA, a sample path (θ, Ψ) (called the nominal sample path) is observed, and then a perturbed path $(\theta + \Delta\theta, \Psi)$ is constructed via a set of perturbation generation and perturbation propagation rules. Then the sensitivity estimate is given as in (3), where the perturbation in $\Delta\theta$ is such that the nominal and perturbed sample paths are “deterministically similar” [3, pp. 38–74]. In terms of discrete-event simulation, this would amount to two different simulation runs, but both having the same initial seed for the random number generator.

III. SYSTEM DESCRIPTION

We will consider a K -user DS-CDMA system [see Fig. 1(a)] where the received signal in the channel is the sum of the transmissions due to the K users and additive white Gaussian noise. The received signal due to the transmission of the k th user at any receiver is given as

$$\rho_k(t) = \sqrt{2P_k} \sum_{i=-\infty}^{\infty} b_{i,k} a_k(t - iT - \tau_k) \cdot \cos(\omega_c t + \phi_k), \quad 1 \leq k \leq K \quad (5)$$

where $b_{i,k} \in \{-1, +1\}$ is the i th bit of the k th user, T is the bit period, and P_k , ϕ_k , and τ_k are the power, carrier phase, and delay of the k th user, respectively. The carrier frequency is denoted by ω_c , and $a_k(t)$ is the spreading waveform of the k th user given by

$$a_k(t) = \sum_{n=0}^{N-1} a_k[n] \chi(t - nT_c) \quad (6)$$

where $a_k[n]$ is the n th element of the sequence for the k th user, $\chi(t)$ is the chip waveform which is assumed to be a unit pulse of length T_c , and N is the processing gain given by $N = T/T_c$. In the presence of flat fading, the received signal in the channel due to the K users and additive white Gaussian noise is given as

$$r(t) = \sum_{k=1}^K \rho_k(t) + \eta(t) \quad (7)$$

where ϱ_k is the multiplicative attenuation due to flat Rayleigh fading, and $\eta(t)$ is assumed to be white Gaussian with power spectral density $\mathcal{N}_0/2$. We will consider the demodulation of the first bit of the first user, and we will also assume that the receiver is synchronized to this transmission with $\tau_1 = 0$. The n th sample at the output of the chip matched filter is now given as

$$r[n] = \sqrt{2} \int_{nT_c}^{(n+1)T_c} r(t) \chi(t - nT_c) \cos(\omega_c t + \phi_1) dt \quad (8)$$

where the above signal has been converted to baseband via chip filtering [see Fig. 1(b)]. In order to introduce this system into a DEDS framework, we will assume that this sampled value is held and detection is based on the piecewise constant signal $r[n]$, $n \in [0, N-1]$. Hence, the output of the chip matched filter is viewed as the sample path of interest, and this is piecewise constant over the chip times. In a conventional matched filter detector, these samples are multiplied by the spreading sequence for the desired user, and the signal value that is compared to the threshold is given by

$$Y_N = \sum_{n=0}^{N-1} z[n] \quad (9)$$

where $z[n] = a_1[n]r[n]$, and is given by

$$z[n] = a_1[n]\eta[n] + \sum_{k=2}^K \sqrt{P_k} \varrho_k \cos(\phi_k) \int_{nT_c}^{(n+1)T_c} \cdot a_1[n] a_k(t - \tau_k) b_k(t - \tau_k) dt + \sqrt{P_1} \varrho_1 T_c b_{0,1} \quad (10)$$

where the fading coefficients $\{\varrho_k\}$ are assumed to be in the form of a multiplicative attenuation (Rayleigh-distributed random variables) that remain unchanged over a bit interval [17], and the noise samples $\eta[n]$ are i.i.d. with zero mean and variance $\mathcal{N}_0 T_c/2$. In the above, we have assumed without loss of generality that the carrier phase $\phi_1 = 0$. From observing such a sample path and using (6)–(10), we compute the sample performance function, which is the conditional probability of error, as

$$P_e = Q\left(\frac{\sqrt{P_1} \varrho_1 T \pm \mathcal{I}(\mathbf{b}, \boldsymbol{\varrho}, \boldsymbol{\tau}, \boldsymbol{\phi})}{\sqrt{\mathcal{N}_0 T/2}}\right) \quad (11)$$

where $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty \exp(-t^2/2) dt$, the $\pm$ sign in the numerator of the argument of the Q function is determined by the particular sample path, i.e., whether $b_{0,1}$ is +1 or -1, and $\mathcal{I}(\mathbf{b}, \boldsymbol{\varrho}, \boldsymbol{\tau}, \boldsymbol{\phi})$ is evaluated from the sample path as

$$\mathcal{I}(\mathbf{b}, \boldsymbol{\varrho}, \boldsymbol{\tau}, \boldsymbol{\phi}) = \sum_{k=2}^K \cos(\phi_k) \{ \beta_{-1,k} R_{k,1}(\tau_k) + \beta_{0,k} \hat{R}_{k,1}(\tau_k) \} \quad (12)$$

where R and $\hat{R}$ are the continuous-time partial cross-correlation functions, and $\beta_{i,k} = \sqrt{P_k} \varrho_{i,k} b_{i,k}$, $i = -1, 0$. In (12), the vectors $\mathbf{b} = (b_{-1,2}, b_{0,2}, \dots, b_{-1,K}, b_{0,K})$, $\boldsymbol{\tau} = (\tau_2, \dots, \tau_K)$, and $\boldsymbol{\phi} = (\phi_2, \dots, \phi_K)$ denote the bit, delay, and phase of the $K-1$ interfering signals, respectively. The average probability of error is now given

Fig. 1. (a) K -user DS-CDMA system. (b) Receiver structure for the adaptive detector. It is the conventional matched filter receiver when $h_j = a_j^{(1)}$.

by $\bar{P}_e = E_{\mathbf{b}, \boldsymbol{\varrho}, \boldsymbol{\tau}, \boldsymbol{\phi}}[P_e]$, and in this paper, we estimate the sensitivity of this performance criterion to various parameters using IPA. The sensitivity of $\bar{P}_e$ to any parameter β is estimated using IPA as

$$\frac{\partial \bar{P}_e}{\partial \beta} = \frac{P_e(\beta + \Delta\beta) - P_e(\beta)}{\Delta\beta} \quad (13)$$

where $\Delta\beta$ is a small perturbation in the parameter value, and $P_e(\beta + \Delta\beta)$ and $P_e(\beta)$ are the perturbed and nominal sample performance functions evaluated from the same realization. This technique could be used for addressing the sensitivity of system performance to issues such as near-far effects, power control issues for cellular systems, as well as mobility of users in the cellular environment.

IV. SENSITIVITY TO SYSTEM PARAMETERS

In the following, we will show that the IPA estimates of the sensitivity of the average error probability to some system parameter is unbiased. We will first consider the interfering users' powers as the parameter of interest. Let $P_{k'}$ be the power of the specific interfering user, and IPA is used to estimate $(\partial \bar{P}_e / \partial P_{k'})$. Consider the term in $\mathcal{I}(\mathbf{b}, \boldsymbol{\varrho}, \boldsymbol{\tau}, \boldsymbol{\phi})$, where we have grouped the power term with the random variables for the two interfering bits and the fading constants, i.e., $\beta_{i,k'} = \sqrt{P_{k'}} \varrho_{i,k'} b_{i,k'}$, $i = -1, 0$. Since these bits are either -1 or +1 with equal probability, it follows that the nominal value is

$$\beta_{i,k'} = \sqrt{P_{k'}} G_{\varrho_{i,k'}}(u_{k'}) \alpha_{(0,1/2]}(u), \quad i = -1, 0 \quad (14)$$

where u is a uniform random variable on $(0, 1]$, and $\alpha_{(0,1/2]}(u)$ is either +1 or -1 according to whether $u \in (0, 1/2]$ or not. Similarly, $G_{\varrho_{i,k'}}(\cdot)$ is the inverse mapping that generates a Rayleigh random variable from a uniform random seed $u_{k'}$. The perturbation generation rule for the

perturbed realization will be

$$\beta'_{i,k'} = \sqrt{P_{k'} + \Delta P_{k'}} G_{e_{i,k'}}(u_{k'}) \alpha_{(0,1/2]}(u), \quad i = -1, 0 \quad (15)$$

where the same random numbers u and $u_{k'}$ are used as in the nominal realization.

Proposition 1: For the sample performance function defined in (10), the IPA estimate is unbiased, i.e., $(\partial \bar{P}_e / \partial P_{k'}) = E[(\partial \bar{P}_e / \partial P_{k'})]$. $\square$

The proof of the above proposition follows by using continuity and boundedness arguments of the $Q(\cdot)$ function, and a simple application of the Lebesgue-dominated convergence theorem.

In order to reduce the variance of these sensitivity estimates, several sample realizations are used at each value of the parameter, and the gradient estimate is formed by averaging over these sample gradient estimates. The IPA estimation described above is used to estimate the sensitivity of the average error probability to near-far effects. A two-user DS-CDMA system is considered where the users are employing spreading sequences of length $N = 31$ and a conventional matched filter receiver is being employed. The desired user's power is kept fixed at $P_1 = 1.0$, while the interferer's power level is varied. The slope of the average error probability to varying interferer power levels is estimated using IPA by evaluating sample gradients during 100 different runs and averaging over them. The Gaussian noise is not simulated in this model, and only the desired user and multiple-access interference is simulated during each run. This is because the Gaussian noise is analytically taken into account in (11), with the noise level adjusted by means of its variance given by $N_0 T_c / 2$. This was repeated for varying noise levels, and the results are shown in Fig. 2 where the curves are plotted by using a piecewise linear approximation using the slopes estimated at predetermined interferer powers. These variations are consistent with values obtained using simulations, and show that IPA is a very viable method for the sensitivity analysis of such systems. Moreover, it is seen that the IPA estimates obtained from averaging over only 100 sample runs very accurately mimic the Monte Carlo estimates of the error probabilities which require roughly sample runs that are inversely proportional to the actual error probabilities. This is owing to the fact that the IPA estimates are based on common random realizations [18].

A similar estimation procedure can be carried out for studying the sensitivity of the average error probability to flat (slow) fading or alternately imperfections in power control. Once again, the perturbation generation rule is given as in (14) and (15); however, the function $G(\cdot)$ is now perturbed to reflect an infinitesimal change in the parameter of the Rayleigh distribution.

V. ADAPTIVE ONE-SHOT DETECTOR

In the following, we consider a DS-CDMA system that is subject to flat (slow) fading, and we develop an adaptive linear one-shot detector for a multiuser channel. This detection scheme is based on using a stochastic gradient algorithm

Fig. 2. Sensitivity to near-far effects. The linear interpolation with IPA estimates (—) are seen to be consistent with simulation results (x) obtained via Monte Carlo simulations.

for minimizing the average probability of error. However, the gradients required at each step of the algorithm are computed using infinitesimal perturbation analysis from only a single sample realization and without any averaging. The stochastic gradient algorithm requires the extraction of the transmitted signal from the received signal (we will call this detector “clairvoyant”), and is thus similar to multiuser detection schemes [4]–[8] that require information about the interference in the channel. To overcome this drawback, we develop a practical implementation of the above adaptive detector where we use a joint estimation–detection algorithm. A sequential implementation of the above detector which requires no training is also developed.

A. Stochastic Gradient Algorithms Using IPA

Consider a discrete-event dynamic system that is represented by the pair (θ, Ψ) with $\bar{P}_e(\theta) = E[L(\theta, \Psi)]$ as the expected performance measure of interest. It is required to solve an optimization problem of the form in (2), which can be solved via an algorithm of adaptation [19] where the i th step of the iterative algorithm is given as

$$\theta_{i+1} = \theta_i - \gamma_i \nabla_{\theta} L(\theta_i, \Psi_i) \quad (16)$$

where θ_{i+1} is the i th update of the parameter vector, $\nabla_{\theta} L(\theta_i, \Psi_i)$ is the gradient evaluated using infinitesimal perturbation analysis, and γ_i is the step size. We now present the following Lemma [19, p. 56] on the convergence of such an algorithm.

Lemma 1 (Tsytkin): The sufficient conditions for the algorithm of adaptation in (16) to converge almost surely, i.e., $\lim_i \theta_i \xrightarrow{\text{a.s.}} \theta^*$, are

$$\gamma_i > 0, \quad \sum_{i=1}^{\infty} \gamma_i = \infty, \quad \sum_{i=1}^{\infty} \gamma_i^2 < \infty \quad (17)$$

$$\inf_{\epsilon < \|\theta - \theta^*\| < 1/\epsilon} E[(\theta - \theta^*)^T \nabla_{\theta} L(\theta, \Psi)] > 0 \quad (18)$$

$$E[\nabla_{\theta}^T L(\theta, \Psi) \nabla_{\theta} L(\theta, \Psi)] \leq \alpha(1 + \theta^T \theta) \quad (19)$$

where $\boldsymbol{\theta}^* = \arg \{\min_{\boldsymbol{\theta}} \bar{P}_e(\boldsymbol{\theta})\}$, and the operation $\boldsymbol{\theta}^\top$ denotes the transpose of the vector. $\square$

A very simple physical and geometric meaning of the above conditions can be found in [19].

B. A Clairvoyant Minimum Probability of Error Detector

In this section, we will develop an adaptive algorithm for implementing the minimum probability of an error detector. Therefore, we will describe the developments for the i th bit interval of the desired user. This detector requires a training sequence or a preamble of bits to be sent by the desired user. For the adaptive detector proposed here, the samples at the output of the chip matched filter are multiplied by a sequence $h_i[n]$, $n = 0, \dots, N-1$, as opposed to the conventional matched filter receiver where the sequence $h_i[n] = a_1[n]$. As a result, the sampled signal values $z_i[n]$ in (10) are now given as

$$z_i[n] = h_i[n]\eta_i[n] + \sum_{k=2}^K \sqrt{P_k} \varrho_k \cos(\phi_k) \cdot \int_{nT_c}^{(n+1)T_c} h_i[n] a_k(t - \tau_k) b_k(t - \tau_k) dt + \sqrt{P_1} \varrho_1 T_c h_i[n] a_1[n] b_{i,1} \quad (20)$$

where the despreading has been performed with the sequence $h_i[n]$, $n = 0, \dots, N-1$. We will use infinitesimal perturbation analysis to estimate the sensitivity of the average probability of error to the sequence $h_i[n]$, and incorporate these estimates in a stochastic algorithm of the form in (16). In this algorithm, the sequence $h_i[n]$, $n = 0, \dots, N-1$ is adaptively updated from bit period to bit period in order to minimize the average probability of error. Thus, we are proposing a linear detector for minimum probability of error.

In order to formulate this problem in the framework of the stochastic algorithm described in the previous section, we first note that the filtered received signal in (8) can be written in the form of a vector as $\mathbf{r}_i^\top = (r_i[0] \dots r_i[N-1])$. Without loss of generality, we will assume that the receiver is synchronized to the desired user, i.e., $\tau_1 = 0$, and also that $\phi_1 = 0$. For the relative delays of the interferers, we will consider an asynchronous system where we treat each of the delays as $\tau_k = (\nu_k + \delta_k)T_c$, for $2 \leq k \leq K$, where ν_k is an integer between 0 and $N-1$, and δ_k lies in the interval $[0, 1)$. In the detection scheme considered here, the estimate of the desired bit of the desired user depends only on the received signal for $t \in [0, T]$, and the decisions are based on the vector of received samples $\mathbf{r}$ which is given for the i th bit interval as

$$\mathbf{r}_i = \mathbf{s}_i + \boldsymbol{\eta}_i \quad (21)$$

where $\mathbf{s}_i$ is given by

$$\mathbf{s}_i = b_{i,1} \varrho_1 \sqrt{P_1} \mathbf{a}_1 + \sum_{k=2}^K \sqrt{P_k} \cos(\phi_k) \cdot [b_{i,k} \varrho_{i,k} \boldsymbol{\alpha}_{i,k} + b_{i-1,k} \varrho_{i-1,k} \boldsymbol{\alpha}_{i-1,k}] \quad (22)$$

with the k th interfering signature sequence modulated by the symbol $b_{i,k}$ for $\tau_k < t \leq T$, and by the symbol $b_{i-1,k}$ for

$0 < t \leq \tau_k$. The attenuation constants due to slow fading are represented by ϱ_k , and are assumed to be constant over each bit interval of the particular user. In the above, the vectors $\mathbf{a}_k \in \mathbb{R}^N$ are the vectors of spreading sequences for the users, i.e., $\mathbf{a}_k = (a_k[0] \dots a_k[N-1])^\top$, and

$$\alpha_{i,k}[n] = \zeta_{1,k} a_k[n - \nu_k] I_{\{n \geq \nu_k\}} + \zeta_{2,k} a_k[n - \nu_k - 1] I_{\{n \geq \nu_k + 1\}} \quad (23)$$

$$\alpha_{i-1,k}[n] = \zeta_{1,k} a_k[n + N - \nu_k] I_{\{n \leq \nu_k - 1\}} + \zeta_{2,k} a_k[n + N - \nu_k - 1] I_{\{n \leq \nu_k\}} \quad (24)$$

for $0 \leq n \leq N-1$ and $2 \leq k \leq K$. The function I_A is the indicator for the set A , and $\zeta_{1,k}$ and $\zeta_{2,k}$ are defined as

$$\zeta_{1,k} = \int_{nT_c}^{(n+1)T_c} \chi(t) \chi(t + \delta_k T_c) dt \quad (25)$$

$$\zeta_{2,k} = \int_{nT_c}^{(n+1)T_c} \chi(t) \chi(t + (1 - \delta_k) T_c) dt. \quad (26)$$

When $\chi(t)$ is a rectangular pulse of width T_c , the above two expressions are reduced to $\zeta_{1,k} = \delta_k$ and $\zeta_{2,k} = 1 - \delta_k$. In the above formulation, $\boldsymbol{\eta}$ is an N -dimensional Gaussian vector with zero mean and covariance matrix $\sigma^2 \mathbf{I}$, where $\mathbf{I}$ is the identity matrix of dimension N . Based on the above formulation, the sample performance function in the i th bit interval can now be written as

$$P_e = Q\left(\frac{\pm \mathbf{h}_i^\top \mathbf{s}_i}{\sigma \sqrt{\mathbf{h}_i^\top \mathbf{I} \mathbf{h}_i}}\right) \quad (27)$$

where $\mathbf{s}_i$ is the vector of received signals that carry bit information (being the sum of desired user's attenuated signal and the attenuated interfering signals) samples, i.e., the received signal vector (both the desired user and the interfering users' signals) without the additive Gaussian noise samples. The $\pm$ sign in (27) is determined by the particular sample path, that is, whether $b_{i,1}$ in (22) is either $+1$ or -1 . This corresponds to the bits sent by the desired user as part of the training sequence or preamble. The adaptive detection scheme is now based on finding the best set of sequences that minimize¹ $\bar{P}_e$, i.e., the vector $\mathbf{h}^*$ such that

$$\mathbf{h}^* = \arg \{\min_{\mathbf{h}} \bar{P}_e\}. \quad (28)$$

We now adaptively update the vector $\mathbf{h}$, using the stochastic algorithm discussed in the previous section, as

$$\mathbf{h}_{i+1} = \mathbf{h}_i - \gamma_i \nabla_{\mathbf{h}_i} P_e(\mathbf{h}_i, \mathbf{s}_i) \quad (29)$$

where the gradient of the average bit-error rate is estimated as $\nabla_{\mathbf{h}} P_e$ using infinitesimal perturbation analysis, and $\mathbf{s}_i$ is the vector of transmitted signals for the i th iteration, i.e., the i th bit period. The algorithm is initialized by using the matched filter taps for the zeroth bit interval.

¹This is not equivalent to (classical) maximizing SNR since the multiple access interference is non-Gaussian

Proposition 2: For the sample performance function defined in (27), the infinitesimal perturbation analysis estimate of the sensitivity of the average probability of error to the despreading sequence is unbiased, i.e., $E[\nabla_{\mathbf{h}} P_e] = \nabla_{\mathbf{h}} E[P_e]$. $\square$

The proof of the above proposition once again follows by showing that the sample performance function in (27) is continuous and bounded in each of the components of $\mathbf{h}$. We now state the following proposition on the convergence of the proposed adaptive algorithm.

Proposition 3: The algorithm given in (29) converges in the sense that $\lim_{i \rightarrow \infty} \mathbf{h}_i \xrightarrow{\text{a.s.}} \mathbf{h}^*$. $\square$

We sketch a proof of the above proposition in the Appendix. The next issue considered is the implementation of the detector. In the following, we show that the adaptive detector implementation allows a simple recursive update structure. We note that Proposition 2 implies that the IPA estimates of the sensitivity of the bit-error probability are unbiased. Therefore, this implies that the nominal and perturbed sample paths are “deterministically similar.” In other words, these two paths have the same realization of random numbers (vectors) in them. Therefore, the sample gradient can be computed from a single sample path as

$$\nabla_{\mathbf{h}_i} P_e = \left(\frac{\partial P_e}{\partial h_i[0]} \cdots \frac{\partial P_e}{\partial h_i[N-1]} \right)^\top \quad (30)$$

where P_e is as given in (27). Furthermore, it follows by the chain rule of differentiation that $(\partial Q(f(x))/\partial x) = -(1/\sqrt{2\pi}) \exp(-f^2/2) (\partial f(x)/\partial x)$, which implies via (27) that

$$\frac{\partial P_e}{\partial h_i[n]} = -\frac{1}{2\sqrt{2\pi}} \left\{ \exp \left(-\frac{1}{2} \left(\frac{\pm \mathbf{h}_i^\top \mathbf{s}_i}{\sigma \|\mathbf{h}_i\|} \right)^2 \right) \cdot \frac{\partial \left(\frac{\pm \mathbf{h}_i^\top \mathbf{s}_i}{\sigma \|\mathbf{h}_i\|} \right)}{\partial h_i[n]} \right\} \quad (31)$$

where $\mathbf{s}_i$ is as given in (22) and the partial derivatives are given as

$$\frac{\partial \left(\frac{\pm \mathbf{h}_i^\top \mathbf{s}_i}{\sigma \|\mathbf{h}_i\|} \right)}{\partial h_i[n]} = \frac{\pm s_i[n]}{\sigma \|\mathbf{h}_i\|} + \frac{\pm \mathbf{h}_i^\top \mathbf{s}_i}{\sigma^2 \|\mathbf{h}_i\|^3} h_i[n], \quad n = 0, \dots, N-1 \quad (32)$$

where $s_i[n]$ denotes the n th component of the vector $\mathbf{s}_i$. Thus, by observing a single sample path, we see that the sample partial derivatives, and hence the sample gradient required for the adaptive algorithm in (29), can be computed. Therefore, the result in Proposition 2 and the analyticity of the Q function circumvents the requirement of two sample runs for evaluating the gradient at each iteration of the adaptive

algorithm. However, the drawback² of the adaptive detector proposed in this section is that it requires perfect extraction of the transmitted signal from the received signal samples, i.e., the extraction of the vector $\mathbf{s}_i$ from $\mathbf{r}_i = \mathbf{s}_i + \boldsymbol{\eta}_i$. Alternately, if the above information of the users in the channel is available *a priori*, then we could implement the above optimum one-shot detector by using numerical simulations to evaluate (29) and solve the optimization problem in (28).

3) Joint Maximum Likelihood Estimation (MLE) of Signals: In this section, we develop a practical implementation of the adaptive single-user detector proposed in the previous section. Specifically, we propose a joint estimation–detection strategy where we overcome the drawback of the detector proposed in the previous section. We use maximum likelihood (ML) estimates of the transmitted signal which are then incorporated into an IPA-based stochastic gradient algorithm as before. The received signal filtered at each of the chip values forms the observation vector given as $\mathbf{r}_i^\top = (r_i[0] \cdots r_i[N-1])$, which can be rewritten for the i th bit interval via (21) as $\mathbf{r}_i = \mathbf{s}_i + \boldsymbol{\eta}_i$, where the transmitted vector can be parameterized as

$$\mathbf{s}_i = \sum_{k=1}^K \boldsymbol{\xi}_{k,i}(A_k, \tau_k) \quad (33)$$

where A_k and τ_k are the amplitude and delay of the signal due to the k th user, respectively, and $\boldsymbol{\xi}_{k,i}$ is of the form

$$\boldsymbol{\xi}_{k,i} = \begin{cases} \sqrt{P_1} \varrho_1 b_{i,1} \mathbf{a}_1, & k=1 \\ \sqrt{P_k} [b_{i,k} \varrho_{i,k} \mathbf{a}_{i,k} + b_{i-1,k} \varrho_{i-1,k} \mathbf{a}_{i-1,k}], & k=2, \dots, K \end{cases} \quad (34)$$

where the symbols in the above equation are as defined before. Since the symbols $b_{i-1,k}$ and $b_{i,k}$ are independent of each other, we can consider the desired user to be interfered by $2(K-1)$ interferers. If the timing information of the users in the channel is known, then the parameterization in (33) is given as

$$\mathbf{s}_i = \sum_{k=1}^{2K-1} B_{k,i} \mathbf{a}_k^{(\tau_k)} \quad (35)$$

where $B_{1,i}$ accounts for the phase, attenuation, and power of the desired user, and $B_{k,i}$, $k=2 \cdots 2K-1$ accounts for the bit, phase, attenuation, and power of the $2(K-1)$ interferers, respectively. Since the timing information of the interferers is known, the code vectors $\mathbf{a}_k^{(\tau_k)}$, $k=2, \dots, 2K-1$ can be constructed by using the left and right portions of the code sequence for each of the $K-1$ interfering sequences. The maximum likelihood estimates of the parameters $B_{1,i}, \dots, B_{2K-1,i}$ during the i th bit interval are given as the solution of the $2K-1$ linear simultaneous equations given by

$$\frac{1}{\sigma^2} (\mathbf{r}_i - \mathbf{s}_i)^\top \mathbf{I}^{-1} \frac{\partial \mathbf{s}_i}{\partial B_{k,i}} = 0, \quad k=1, \dots, 2K-1 \quad (36)$$

where $\sigma^2 \mathbf{I}$ is the covariance matrix of the additive Gaussian noise vector $\boldsymbol{\eta}$. We then use the maximum likelihood estimates

²The MMSE receiver works directly on $\mathbf{r}_i$, which is the received signal vector, as opposed to $\mathbf{s}_i$, which is the sum of the desired signal and interfering signal vectors.

of the transmitted signal in the stochastic gradient algorithm as in (29), that is,

$$\mathbf{h}_{i+1} = \mathbf{h}_i - \gamma_i \nabla_{\mathbf{h}_i} P_e(\mathbf{h}_i, \hat{\mathbf{s}}_i) \quad (37)$$

where $\hat{\mathbf{s}}_i$ is the maximum likelihood estimate of the transmitted signal from the observation vector $\mathbf{r}_i$ during the i th iteration of the algorithm.

When the timing information of the users in the channel is not known, we use a coarse parameterization of the transmitted signal so that the maximum likelihood estimates can be readily incorporated into an IPA-based stochastic gradient algorithm. Since the proposed adaptive algorithm requires only the raw values of the transmitted signal (i.e., the sum of the desired user and the interfering signals) and not the individual amplitudes and delays of the different users, we use maximum likelihood estimates of $\mathbf{s}_i$ based on observing $\mathbf{r}_i = \mathbf{s}_i + \boldsymbol{\eta}_i$. The maximum likelihood estimate of the transmitted vector during the i th bit interval is given as $\hat{\mathbf{s}}_i = \mathbf{r}_i$. In this case, the adaptive detector is implemented as in (37) with the received signal vector itself incorporated in the sample performance function, and also the sample gradient. This implementation is consistent with that of the MMSE receiver.

C. Sequential Detector

In this section, we will develop a sequential algorithm which overcomes the drawbacks of a training sequence. In this algorithm, the bit decision from the preceding bit interval is used to adaptively update the filter coefficients in every successive bit interval. Therefore, the sample performance function used to evaluate the gradient in the i th bit interval is the one corresponding to the received signal in the $(i-1)$ st bit interval. Based on the formulation in Section V-B, the sample performance function in the i th bit interval can now be written as

$$P_e = Q\left(\frac{\hat{b}_{i-1,1} \mathbf{h}_{i-1}^\top \mathbf{s}_{i-1}}{\sigma \sqrt{\mathbf{h}_{i-1}^\top \mathbf{I} \mathbf{h}_{i-1}}}\right) \quad (38)$$

where $\mathbf{s}_{i-1}$ is the vector of transmitted signals (i.e., the sum of desired user's signal and the interfering signals) samples during the $(i-1)$ st bit interval. The $\pm$ sign in (27) is now replaced in the above equation by $\hat{b}_{i-1,1}$ which is the bit estimate during the $(i-1)$ st bit interval, i.e.,

$$\hat{b}_{i-1,1} = \text{sgn}(\mathbf{h}_{i-1}^\top \mathbf{r}_{i-1}). \quad (39)$$

The numerical results on the performance of the above schemes is described in the next section.

D. Numerical Results

1) *Clairvoyant Detector*: We now compare the performance of the proposed adaptive single-user detector for multiuser channels with that of a conventional matched filter receiver. The signature sequences considered for the different users were m sequences of length $N = 31$, and the interferer powers were set to be equal. The interferers were asynchronous, and the relative delays of the interferers with respect to the desired user are modeled as uniform

random variables on $[0, T)$. The phases are chosen to be identically zero for the simulations. In Fig. 3, the performance of the adaptive detector and the conventional matched filter receiver is shown for a three-user DS-CDMA system with varying interference powers. It is seen that while the matched filter receiver degrades with increasing interferer powers, the adaptive detector degrades very gracefully and is more near-far resistant. In order to compare the performance of the adaptive single-user detector with that of the recently proposed MMSE detector [14], [15], a two-user DS-CDMA employing m sequences of length $N = 31$ was considered. The relative power of the interferer is increased, while the desired user power is kept fixed. The result of such a comparison is shown in Fig. 6 where, for the MMSE detector, we assume perfect extraction of the transmitted signal, and use the closed-form solution for the filter coefficients (see [14]) rather than use an adaptive algorithm. Thus, all of our numerical simulations are biased in favor of the MMSE detector (the implementation of the MMSE receiver is simpler as it works directly with the received signal). While the MMSE detector is perfectly near-far resistant, the adaptive multiuser detector outperforms it in that it is fairly near-far resistant and has the lowest average probability of bit error. However, at very high power levels of the multiple-access interference, the performance of the adaptive detector proposed here approaches that of the MMSE detector (see Fig. 6), while at very low power levels of the multiple-access interference, the adaptive detector performance is the same as that of the conventional receiver (see Fig. 3), which is the optimum detector for channels with no multiple-access interference. Note that when the interference in the channel is Gaussian, the MMSE detector performs similarly to the adaptive detector in that it yields the same average BER. In the above figures, the BER of the adaptive detector is evaluated using Monte Carlo simulations after the taps have converged to the optimum values. As is the practice in MMSE receivers, the step sizes chosen were constant as opposed to the varying step sizes required for theoretical convergence of the adaptive detector. We have not optimized the step sizes for the numerical results since, typically, the optimum step size would have to be chosen according to how fast the fading is.

In order to study the dynamics of the proposed algorithm to step size and near-far effects, we study the variation of the angle between the interferer codes and the filter coefficients $\mathbf{h}$ of the adaptive algorithm. This is more intuitive than studying the convergence of filter coefficients in answering the issue of whether or not the filter taps of the algorithm change so as to reduce their cross correlation with the interferer codes. A pictorial representation of the dynamics is shown in Fig. 4. To illustrate this point, a two-user DS-CDMA system is considered in Fig. 5(a) and (b) where the cosine of the angle between the desired user code $\mathbf{a}_1$ and $\mathbf{a}_2$, i.e., $\cos(\delta) = (\mathbf{a}_1^\top \mathbf{a}_2 / |\mathbf{a}_1| |\mathbf{a}_2|)$, is plotted. The variation of this angle with the number of iterations is shown, where it is seen that the adaptive algorithm at each step tries to make itself orthogonal to the interferer code $\mathbf{a}_2$. At convergence, it is seen that $\cos(\delta') = (\mathbf{h}^{*\top} \mathbf{a}_2 / |\mathbf{h}^*| |\mathbf{a}_2|)$, is approaching zero, i.e., the codes $\mathbf{h}^*$ and $\mathbf{a}_2$ are almost orthogonal. In Fig. 5(a),

Fig. 3. Performance of adaptive detector (clairvoyant). The adaptive detector (clairvoyant) is compared with the conventional matched filter receiver. The SNR is 16 dB. The adaptive detector is fairly near-far resistant with overall better performance.

Fig. 4. Pictorial representation of algorithm dynamics. The algorithm dynamics move the filter taps from $\mathbf{h}_0 = \mathbf{a}_1$ to $\mathbf{h}^*$ which is almost orthogonal to the interference vector $\mathbf{a}_2$, while still retaining maximum "autocorrelation" with $\mathbf{a}_1$.

the convergence of the algorithm is shown for various step sizes. It is seen that the algorithm converges in identical fashion for a range of values of the step size. For the above example, $(P_2/P_1) = 5$ and the SNR is set to 11 dB. In Fig. 5(b), the convergence of the algorithm is shown for severe near-far environments. Once again, the convergence of the algorithm is fairly robust to increasing near-far effects. The rate of convergence of the algorithm can be hastened by choosing varying step sizes, but this would again result in added complexity over the MMSE receiver.

2) *Joint Maximum Likelihood Estimation (MLE) of Signals*: The performance of the practical implementation of the adaptive detection scheme is shown in Fig. 6. It is seen that the adaptive detector using ML estimates performs better than both the MMSE detector and the matched filter receiver. However, there is a degradation in its performance relative to the adaptive detector (clairvoyant) owing to fact that the transmitted signals were parameterized in a coarse manner without any information about the interferers. Even though the received signal vector is used in this implementation, it is seen in Fig. 8 that the detector closely follows the adaptive detector with exact information about the interferers. In the example shown, the interferers are five times stronger than the desired user. When the timing information is available (see Fig. 6), the performance of the detector is enhanced, and it closely mimics the performance of the adaptive detector (clairvoyant).

Fig. 5. (a) Dynamics of adaptive algorithm. The cosine of the angle between the adaptive filter taps and the interference vector is shown for a two-user system with $N = 31$. It is seen that the filter taps try to "orthogonalize" themselves to the interference vector during each step of the algorithm. The results are shown for various values of γ the constant step size chosen. (b) Convergence under near-far effects. The cosine of the angle between the adaptive filter taps and the interference vector is shown for a two-user system with $N = 31$. It is seen that the filter taps try to "orthogonalize" themselves to the interference vector during each step of the algorithm. The results are shown for various values of relative strength of interferer. The constant step size chosen is $\gamma = 0.05$. The SNR is 10 dB.

As in the case of the clairvoyant adaptive detector, this detector approaches either the matched filter detector or the MMSE detector under extreme environments. For the case of a flat fading channel, the numerical results are shown in Fig. 7 where the adaptive detector implemented using the received signal (and without any timing information) performs slightly better than that of the MMSE receiver. The performance under increasing variance (or alternately, the second moment with fixed mean) of the fading constants shows trends similar to that observed in Fig. 8, where the detector performance is illustrated for varying SNR's.

3) *Sequential Detector*: The performance results of the sequential detector proposed in this paper is shown in Fig. 9

Fig. 6. Comparison of detection schemes. The various implementations of the adaptive detectors are shown versus the MMSE detector (with exact information) and the matched filter detector. The SNR is 11 dB. The adaptive detectors outperform both the MMSE and matched filter schemes. The single-user lower bound ($K = 1$) is shown for reference. The step size is fixed at $\gamma = 0.01$.

Fig. 7. Performance of adaptive detector under flat fading. It is seen that the adaptive detector (JMLE and without timing information) performs better than the MMSE detector, but the gains are not as significant as in the case of a nonfading channel. The SNR is 11 dB, and the Rayleigh fading is generated from two i.i.d. normal random variables. The step size is fixed at $\gamma = 0.01$.

where a six-user DS-SS-CDMA system employing m sequences of length 31 is considered. The sequential detector outperforms the MMSE detector, but there is a degradation in performance compared to that of the adaptive detector with training. The sequential algorithm is initialized by using a matched filter detector for the zeroth bit interval, and as a result, the rate of convergence of this algorithm may be affected in severe near-far environments (see Fig. 10).

VI. DISCUSSION AND CONCLUSION

In this work, a DS-SS-CDMA system was considered in the framework of a discrete-event dynamic system, and the technique of infinitesimal perturbation analysis was used to es-

Fig. 8. Performance of adaptive detector with varying SNR. It is seen that the adaptive detector (JMLE and without timing information) closely mimics the adaptive detector (clairvoyant) over a wide range of SNR's. The interferers are five times stronger in this case. The step size is fixed at $\gamma = 0.01$.

Fig. 9. Performance of sequential detection scheme. There is a degradation in performance of the sequential detector with respect to the adaptive detector (clairvoyant) which uses a preamble. However, it still outperforms the MMSE detector (with exact information). The SNR is set to 10 dB. The step size is fixed at $\gamma = 0.01$.

timate the sensitivity of system performance. This method clearly circumvents the constraints imposed on analysis due to modulation and detection schemes. Further, the above IPA analysis can be readily incorporated into stochastic approximation schemes where the average probability of bit error is the performance measure of interest. Thus, unlike existing analysis methods, the above sample realization formulation uniquely provides a convenient handle on the optimization of CDMA systems with regard to minimizing the average probability of bit error. Specifically, an adaptive detector was developed for multiuser channels subject to slow fading, where the performance criterion for this detector was the minimization of the average probability of bit error. The practical implementation of this detector was shown to be superior to the conventional matched filter, and also a viable

Fig. 10. Convergence of sequential detection scheme. A typical convergence plot of a filter tap (value) h for the sequential detector (—) is shown versus the adaptive detector with preamble (---) for the case of constant step size $\gamma = 0.01$. The example is a six-user system with interference five times stronger, and SNR of 10 dB.

alternative to multiuser detection schemes [4]–[8] in that it does not require *a priori* information about the multiple-access interference. The performance of the proposed detector was superior to the MMSE detector [14], [15] as well. However, the MMSE receiver has the advantage that it is easier to implement since it directly uses the received signal. To overcome this drawback, several variations of the adaptive detector were also proposed. It was also shown that the dynamics of the algorithm proposed here are such that the filter coefficients are updated at each stage so as to make them as orthogonal as possible to the interferer codes while still retaining maximum “correlation” with the desired user codes. As a result of this behavior, the proposed adaptive detector performs uniformly better than both the conventional receiver as well as the MMSE detector. Furthermore, the sequential implementation of this detector shows that it is viable even in vastly different multiuser environments, and it is a very good candidate for implementation in systems subject to temporal variations.

APPENDIX

In this appendix, we will sketch a proof of Proposition 3. By Lemma 1, the sufficient conditions for the algorithm in (29) to converge almost surely are given as

$$\inf_{\epsilon < \|\mathbf{h} - \mathbf{h}^*\| < 1/\epsilon} E[(\mathbf{h} - \mathbf{h}^*)^\top \sum_{j=1}^2 \nabla_{\mathbf{h}} P_{e,j}(\mathbf{h}, \Psi)] > 0 \quad (40)$$

$$E \left[\nabla_{\mathbf{h}}^\top \left(\sum_{j=1}^2 P_{e,j}(\mathbf{h}, \Psi) \right) \nabla_{\mathbf{h}} \left(\sum_{j=1}^2 P_{e,j}(\mathbf{h}, \Psi) \right) \right] \leq \alpha(1 + \mathbf{h}^\top \mathbf{h}) \quad (41)$$

where $P_{e,j}(\mathbf{h}, \Psi)$ is as given in (27) with $j = 1$ and $j = 2$ corresponding to the case of the desired user’s bit being -1 or $+1$. The condition in (40) is a requirement on the convexity of the average probability of bit error, and is satisfied whenever

the interference, fading, and noise statistics are such that the likelihood ratio of the underlying sufficient statistics is monotonic. Alternately, this condition requires the near–far resistance of the detector to be nonzero during the iterations. In the absence of this condition, the algorithm would converge to a local minimum. In order to prove (41), which is essentially a condition on the parabolicity of the expectation of the quadratic form, we begin with the following identity:

$$E[\nabla_{\mathbf{h}}^\top Q(f(\mathbf{h})) \nabla_{\mathbf{h}} Q(f(\mathbf{h}))] \leq \frac{1}{2\pi} \sum_{n=1}^N E \left[\left(\frac{\partial f}{\partial h[n]} \right)^2 \right] \quad (42)$$

which follows easily. From (27), it can be seen that $P_{e,j}(\mathbf{h}, \Psi)$ is given as $P_{e,j}(\mathbf{h}, \Psi) = Q(f_j(\mathbf{h}))$, $j = 1, 2$, where $f_j(\mathbf{h})$ is given as $f_j(\mathbf{h}) = (\mathbf{h}^\top \boldsymbol{\beta}_j / \sigma \sqrt{\mathbf{h}^\top \mathbf{h}})$, and

$$\boldsymbol{\beta}_j = \mathbf{a}_1 + (-1)^j \sum_{k=2}^K \sqrt{P_k} (b_{0,k} \boldsymbol{\alpha}_{0,k} + b_{-1,k} \boldsymbol{\alpha}_{-1,k}), \quad j = 1, 2 \quad (43)$$

where the cases of $j = 1$ and $j = 2$ correspond to the desired user sending bit -1 and bit $+1$, respectively. Therefore, we can write

$$\begin{aligned} & \frac{1}{2\pi} \sum_{n=1}^N E \left[\left(\sum_{j=1}^2 \frac{\partial f_j}{\partial h[n]} \right)^2 \right] \\ &= \frac{2}{\pi \sigma^2} \sum_{n=1}^N E \left[\left(\frac{\partial}{\partial h[n]} \frac{\mathbf{h}^\top \mathbf{a}_1}{\sqrt{\mathbf{h}^\top \mathbf{h}}} \right)^2 \right]. \end{aligned} \quad (44)$$

After evaluating the partial derivatives in the above expression, some algebraic manipulation results in

$$\begin{aligned} & \frac{1}{2\pi} \sum_{n=1}^N E \left[\left(\sum_{j=1}^2 \frac{\partial f_j}{\partial h[n]} \right)^2 \right] \\ & \leq \frac{2}{\pi \sigma^2} \frac{1}{(\mathbf{h}^\top \mathbf{h})_{\min}} \sum_{n=1}^N E[(\mathbf{a}_1[n])^2] (1 + \mathbf{h}^\top \mathbf{h}) \end{aligned} \quad (45)$$

where $\mathbf{a}_1[n]$ is the n th entry of the vector $\mathbf{a}_1$, and $(\mathbf{h}^\top \mathbf{h})_{\min}$ is the minimum value of the sum of the squares of the filter taps. Since the user sequences are PN sequences of either $+1$ or -1 , it follows that

$$\begin{aligned} & \frac{1}{2\pi} \sum_{n=1}^N E \left[\left(\sum_{j=1}^2 \frac{\partial f_j}{\partial h[n]} \right)^2 \right] \\ & \leq \frac{2N}{\pi \sigma^2} \frac{1}{(\mathbf{h}^\top \mathbf{h})_{\min}} (1 + \mathbf{h}^\top \mathbf{h}) = \kappa(1 + \mathbf{h}^\top \mathbf{h}) \end{aligned} \quad (46)$$

with $\kappa = 2N/\pi \sigma^2 (\mathbf{h}^\top \mathbf{h})_{\min}$. We now use the identity in (42), and the condition in (41) is satisfied. The same procedure can be used to derive the condition for the case of slow fading under the assumption that the fading is independent of the bits.

ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their useful and constructive comments.

REFERENCES

- [1] R. Suri and Y. T. Leung, "Single run optimization of discrete event simulations—An empirical study using the M/M/1 queue," *IEEE Trans.*, vol. 21, pp. 35–49, Mar. 1989.
- [2] E. K. Chong and P. J. Ramadge, "Optimization of queues using an infinitesimal perturbation analysis-based stochastic algorithm with general update times," *SIAM J. Contr. Opt.*, vol. 31, May 1993.
- [3] Y.-C. Ho and X.-R. Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*. Boston: Kluwer Academic, 1991.
- [4] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 34, pp. 123–136, Jan. 1989.
- [5] ———, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, Apr. 1990.
- [6] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems," *IEEE Trans. Commun.*, vol. 38, pp. 725–736, Apr. 1990.
- [7] ———, "Multistage detection for asynchronous code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, Apr. 1990.
- [8] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.
- [9] Z. Xie, C. Rushforth, and R. Short, "Multiuser signal detection using sequential decoding," *IEEE Trans. Commun.*, vol. 38, pp. 578–583, May 1990.
- [10] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 43, pp. 421–434, Feb./Mar./Apr. 1995.
- [11] M. Rupf, F. Tarkoy, and J. L. Massey, "User-separating demodulation for code-division multiple-access systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 786–795, June 1994.
- [12] T. Felhauer, A. Klein, and P. W. Baier, "A novel method for CDMA and other applications to separate non orthogonal signals," Univ. Kaiserslautern, Tech. Rep., 1994.
- [13] H. V. Poor and S. Verdú, "Single-user detectors for multiuser channels," *IEEE Trans. Commun.*, vol. 36, pp. 50–60, Jan. 1988.
- [14] U. Madhow and M. I. Honig, "MMSE detection of CDMA signals," in *Proc. Int. Symp. Inform. Theory*, 1993, p. 49.
- [15] S. L. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Commun.*, vol. 43, pp. 1746–1755, Feb./Mar./Apr. 1995.
- [16] P. Billingsley, *Probability and Measure*. New York: Wiley, 1979.
- [17] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
- [18] R. Y. Rubinstein and G. Samorodnitsky, "Variance reduction by the use of common and antithetic random variables," *J. Statist. Computation Simulation*, vol. 22, pp. 161–180, 1985.
- [19] Y. Z. Tsypkin, *Adaptation and Learning in Automatic Systems*. New York: Academic, 1971.

Narayan Mandayam received the B.Tech. (Hons.) degree in 1989 from the Indian Institute of Technology, Kharagpur, and the M.S. and Ph.D. degrees in 1991 and 1994 from Rice University, Houston, TX, all in electrical engineering.

From 1994, to 1996, he was a Research Associate at the Wireless Information Network Laboratory (WINLAB), Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ. In September 1996, he joined the ECE Department at Rutgers as an Assistant Professor, and also serves as the Research Director for Radio Systems at WINLAB. His research interests are in communication theory, wireless system modeling and performance, multiaccess protocols, multiuser detection, and multimedia wireless communications for cellular and PCS.

Dr. Mandayam is a recipient of the Government of India National Talent Scholarship, and the Institute Silver Medal from the Indian Institute of Technology, Kharagpur.

Behnaam Aazhang (S'81–M'85–SM'91) was born in Bandar Anzali, Iran, on December 7, 1957. He received the B.S. (with highest honors), M.S., and Ph.D. degrees in electrical computer engineering from the University of Illinois at Urbana-Champaign in 1981, 1983, and 1986, respectively.

From 1981 to 1985, he was a Research Assistant in the Coordinated Science Laboratory, University of Illinois. In August 1985, he joined the faculty of Rice University, Houston, TX, where he is now a Professor in the Department of Electrical and Computer Engineering. He has been a Visiting Professor at IBM Federal Systems Company, Houston, TX, the Institut für Kommunikationstechnik at Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, the Telecommunications Laboratory at University of Oulu, Oulu, Finland, and at the U.S. Air Force Phillips Laboratory, Albuquerque, NM. His research interests are in the areas of communications theory, information theory, and their applications with emphasis on multiple access communications; cellular mobile radio communications, optical communication networks; and efficient simulation of stochastic systems.

Dr. Aazhang is a recipient of the Alcoa Foundation Award (1993), the NSF Engineering Initiation Award (1987–1989), and the IBM Graduate Fellowship (1984–1985), and is a member of Tau Beta Pi and Eta Kappa Nu. He is currently serving as the Editor for Spread Spectrum Networks of IEEE TRANSACTIONS ON COMMUNICATIONS, as the Treasurer of IEEE Information Theory Society, and the Technical Area Chair of 1997 Asilomar Conference, Monterey, CA. He has served as the Secretary of the Information Theory Society (1990–1993), and as the Publications Chairman of the 1993 IEEE International Symposium on Information Theory, San Antonio, TX.