MMSE Interference Suppression for Timing Acquisition and Demodulation in Direct-Sequence CDMA Systems

Upamanyu Madhow, *Senior Member, IEEE*

***Abstract*— It has been recently shown that minimum-mean-squared-error (MMSE) demodulators are effective means of interference suppression in code-division multiple-access (CDMA) systems. The MMSE demodulator can be implemented adaptively using an initial training sequence, followed by decision-directed adaptation. This requires that the symbol-level timing of the desired user be known prior to training. In this paper we remove this requirement by providing a method for timing acquisition in which the output of the acquisition process is a near–far-resistant demodulator which automatically accounts for the delays and amplitudes of both the desired signal and the interference without explicitly estimating these parameters. The only requirements are a training sequence for the desired user and a finite uncertainty regarding the symbol timing. The latter condition can be realized by using a periodic training sequence even if the absolute timing uncertainty is arbitrarily large.**

***Index Terms*—CDMA, direct sequence, equalization, interference suppression, MMSE, near–far problem, timing acquisition, training.**

I. Introduction

THE minimum-mean-squared-error (MMSE) criterion leads to demodulators for direct-sequence code-division multiple-access (CDMA) which do not suffer from the near–far problem or from the interference floor in performance exhibited by conventional matched-filter demodulation. Application of the linear MMSE criterion to centralized multiuser detection was suggested in [21]. More recently, it was recognized that, since the received signal in linearly modulated CDMA systems with *short* spreading waveforms (i.e., with spreading waveforms whose period equals the symbol interval) is cyclostationary, the linear MMSE criterion is amenable to decentralized adaptive implementation without explicit knowledge of the interference parameters. See [1], [10], [12], and [17] for low-complexity near–far-resistant demodulators based on this idea.

Paper approved by S. L. Miller, the Editor for Spread Spectrum of the IEEE Communications Society. Manuscript received June 25, 1996; revised June 24, 1997 and February 10, 1998. This work was supported by the Office of Naval Research under Grant N00014-95-1-0647. This paper was presented in part at the 1995 IEEE International Symposium for Information Theory, Whistler, Canada, September 1995, and at the 1995 IEEE Military Communications Conference, San Diego, CA, November 1995.

The author is with the Electrical and Computer Engineering Department and the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801 USA (e-mail: madhow@tesla.csl.uiuc.edu).

Publisher Item Identifier S 0090-6778(98)05726-2.

Adaptive implementation of the MMSE demodulator typically involves initial adaptation using a training sequence of symbols transmitted by the desired user, followed by continuing adaptation in decision-directed mode. Such adaptation, however, requires that the receiver know the (symbol-level) *timing* of the desired user, since it must know which symbol of the training sequence falls within the current observation interval. Conventional methods of timing acquisition are as interference-limited and as vulnerable to the near–far problem as conventional demodulation, and, until recently [2], [3], [8], [9], [14], [19], there were no schemes with reasonable complexity for near–far-resistant timing acquisition. Moreover, near–far-resistant timing acquisition for the desired user alone is not enough for arriving at near–far-resistant demodulators, which, in general, must take into account the timings of *all* users. In this paper we present a timing acquisition method whose output is a demodulator that *implicitly* accounts for the timings, spreading sequences, and powers of all users.

Our method consists of quantizing the timing uncertainty into a finite number P of hypotheses regarding the *phase* of the training sequence, running an adaptive demodulator under each hypothesis, and combining the hypotheses that yield the best estimated mean-squared errors (MSE's) to obtain a near–far-resistant demodulator. The only requirements are a known training sequence for the desired user and a timing uncertainty which is a small multiple P of the symbol interval. Since the timing uncertainty is only relevant in that the receiver does not know which of P training symbols falls in a given observation interval, large initial timing uncertainties can be handled by the use of a periodic training sequence with a small period P.

The use of periodic training sequences by multiple users would imply that these users' bits could have nonzero periodic cross correlations. The near–far resistance of the MMSE detector has been previously demonstrated [10] under the assumption that all users' bits are uncorrelated. One of the contributions of this paper is to prove that, under mild conditions on the bit sequences, the MMSE detector is near–far resistant even when the bits of the desired and interfering users are correlated. This in turn implies the near–far resistance of our acquisition scheme. The main practical advantage of our method is that it provides the capability for rapid timing acquisition without coordination between the transmitter and receiver, thus eliminating the need for a pilot signal, a side channel, or a reverse link. As discussed in the conclusions, it therefore opens up possibilities for robust packet radio networks with *ad hoc* architectures. A drawback of our method is that it is not designed to recover chip timing, and the resulting chip asynchronism causes some performance (and hence capacity) loss. As discussed in Section III, it is possible to alleviate the problem by sampling faster than the chip rate.

The blind adaptive version of the MMSE demodulator proposed in [6] can also be extended to a method for timing acquisition that yields a near–far-resistant demodulator. See [8] and [9] for a parallel development of the blind scheme, which is similar in philosophy but differs in detail from the training-based scheme considered in this paper. See [3] and [19] for a subspace-based approach to blind timing acquisition. Another, simpler, approach to acquisition based on MMSE adaptation using an all-one training sequence has been proposed in [14]. However, this method does not yield information sufficient to compute a near–far-resistant demodulator and cannot be used to acquire more than one user at a time. A similar comment applies to a maximum-likelihood timing estimation algorithm proposed in [2], which also relies on an all-one training sequence. Finally, while the proposed acquisition scheme bears some resemblance to cyclic equalization techniques for narrow-band singleuser applications [13], [16] in its use of periodic training sequences, as explained in Section III, the ideas of cyclic equalization do not apply to a wide-band multiuser system.

Section II contains background material, including the system model and a review of the MMSE demodulator. Section III provides a description of our acquisition scheme. Section IV contains performance analysis of the scheme, including the proof of its near–far resistance with periodic training sequences. Numerical results are given in Section V, and Section VI contains our conclusions.

II. Background

After describing an asynchronous CDMA system, we obtain an *equivalent synchronous model* that enables the design of finite-complexity adaptive interference suppression schemes. We conclude this section by reviewing the linear MMSE demodulator, which forms the basis for our joint acquisition and demodulation scheme.

A. Asynchronous CDMA

Consider an asynchronous direct-sequence CDMA system with K simultaneous antipodal users over an additive white Gaussian noise (AWGN) baseband channel.[1] The received signal due to the kth user ($1 \le k \le K$) is given by

$$r^{(k)}(t) = \sum_{n=-\infty}^{\infty} b_{k,n} A_k s_k(t - nT - \tau_k) \tag{1}$$

where T is the *symbol interval*, $b_{k,n} \in \{-1, 1\}$ is the nth symbol of the kth user, A_k is its amplitude, τ_k is its relative delay with respect to the receiver, and $s_k(t)$ is its *spreading waveform*, given by

$$s_k(t) = \sum_{j=0}^{N-1} a_k[j]\psi(t - jT_c). \tag{2}$$

Here $a_k[j] \in \{-1, 1\}$ is the jth element of the *spreading sequence* for the kth user, $\psi(t)$ is the *chip waveform*, T_c is the *chip interval*, and $N = T/T_c$ is the *processing gain*.

The net received signal is given by

$$r(t) = \sum_{k=1}^{K} r^{(k)}(t) + n(t) \tag{3}$$

where $n(t)$ is AWGN. Taking the first user to be the *desired* user, our objective is to demodulate its bit sequence $\{b_{1,n}\}$. We assume that, in the acquisition phase, a training sequence for the bits of the desired user is available, and that there is an uncertainty of at most P bit intervals regarding the phase of the training sequence. Thus, the bit of interest in the nth observation interval is known to be $b_{1,n+i^*}$, where i^* is an unknown "phase" which takes values in $\{0, 1, \cdots, P-1\}$. Other than the assumption of a training sequence and a finite timing uncertainty for the desired user, all other parameters (i.e., the delays, amplitudes, and spreading sequences for all users) are unknown.

B. The Equivalent Synchronous Discrete-Time Model

The receiver does not know either the chip or symbol timing of the desired user. The received signal is passed through a chip matched filter and sampled at the chip rate (or faster). Limiting attention to a finite observation interval for each bit decision[2] then yields an *equivalent synchronous discrete-time* model. All results in this paper are for a rectangular chip waveform and chip rate sampling, although generalizations are straightforward. Since the delay of the desired user is unknown, the receiver's observation interval need not be aligned with the received signal for the desired user in any way. In order to fully utilize the desired signal's energy, the observation interval must be chosen such that at least one complete bit of the desired user falls in the interval regardless of its delay relative to the receiver. The minimum length of the observation interval such that this property holds is $2T$, and for concreteness, this is the length assumed throughout this paper.[3] Longer observation intervals yield better steady-state performance but lead to higher complexity and slower convergence. Choosing an observation interval of length at least $2T$ also has the desirable consequence that a common observation interval can be used for all users being acquired or demodulated by the receiver. For least-squares adaptation, this implies that the computation of the inverse of the correlation matrix can be used for all users, so that an efficient centralized implementation of the decentralized adaptive method proposed here is possible.

[1] The proposed methods also apply to complex baseband models that could incorporate, for instance, two-dimensional signal constellations or different carrier phases for different users (e.g., see [22] and [23] for application of MMSE detectors to a complex baseband model that includes Rayleigh fading).

[2] A finite observation interval is suboptimal for asynchronous CDMA, but has the advantage of limiting the detector complexity.

[3] For dispersive channels, the minimum length is $2T$ plus the channel delay spread.

The lth discrete time sample can be written as

$$r[l] = \int_{lT_c}^{(l+1)T_c} r(t)\,dt. \tag{4}$$

For making a decision on the nth symbol of the desired user, we may consider an *equivalent synchronous system* with received vector $\mathbf{r}_n \in \mathcal{R}^{2N}$ corresponding to the nth observation interval $\mathbf{r}_n = (r[nN], r[nN+1], \cdots, r[nN+2N-1])^T$.

We now express $\mathbf{r}_n$ in terms of the parameters of the asynchronous CDMA model (1)–(3). Without loss of generality, let $b_{k,n}$ denote the bit of the kth user that falls completely in the nth observation interval, and let τ_k denote the delay of this bit relative to the left edge of the nth observation interval. Since $\tau_k \in [0, T)$, we may write it as $\tau_k = (n_k + \delta_k)T_c$, where n_k is an integer between zero and $N-1$, and $\delta_k \in [0, 1)$. Let $\mathbf{a}_k$ denote a vector of length $2N$, consisting of the N elements of the spreading sequence of the kth user followed by N zeros, i.e., $\mathbf{a}_k = (a_k[0], \cdots, a_k[N-1], 0, \cdots, 0)^T$. Let $\mathbf{T}_L$ denote the acyclic left shift operator and let $\mathbf{T}_R$ denote the acyclic right shift operator, both operating on vectors of length $2N$. Thus, for a vector $\mathbf{x} = (x_0, \cdots, x_{2N-1})^T$, we have $\mathbf{T}_L\mathbf{x} = (x_1, \cdots, x_{2N-1}, 0)^T$ and $\mathbf{T}_R\mathbf{x} = (0, x_0, \cdots, x_{2N-2})^T$. Let T_L^n, T_R^n denote n applications of these operators, resulting in left and right shifts by n, respectively.

For each asynchronous user, three consecutive bit intervals overlap with a given observation interval of length $2T$. Furthermore, since the system is chip asynchronous, two adjacent chips contribute to each chip sample. The contribution of the kth user to the received vector $\mathbf{r}_n \in \mathcal{R}^{2N}$ of $2N$ samples for the nth observation is therefore given by

$$\mathbf{r}_n^{(k)} = b_{k,n-1}\mathbf{v}_k^{-1} + b_{k,n}\mathbf{v}_k^0 + b_{k,n+1}\mathbf{v}_k^1 \tag{5}$$

where

$$\begin{aligned}\mathbf{v}_k^{-1} &= A_k[(1-\delta_k)\mathbf{T}_L^{N-n_k}\mathbf{a}_k + \delta_k\mathbf{T}_L^{N-n_k-1}\mathbf{a}_k]\\ \mathbf{v}_k^0 &= A_k[(1-\delta_k)\mathbf{T}_R^{n_k}\mathbf{a}_k + \delta_k\mathbf{T}_R^{n_k+1}\mathbf{a}_k]\\ \mathbf{v}_k^1 &= A_k[(1-\delta_k)\mathbf{T}_R^{n_k+N}\mathbf{a}_k + \delta_k\mathbf{T}_R^{n_k+N+1}\mathbf{a}_k].\end{aligned} \tag{6}$$

The net received vector is given by

$$\mathbf{r}_n = \sum_{k=1}^{K} \mathbf{r}_n^{(k)} + \mathbf{w}_n \tag{7}$$

where $\mathbf{w}_n$ is white Gaussian noise with covariance $\sigma^2\mathbf{I}$. Our task is to demodulate bit $b_{1,n}$ based on the observation vector $\mathbf{r}_n$. The observation vectors $\{\mathbf{r}_n\}$ are stationary, which means that adaptive mechanisms that exploit the structure of $\mathbf{r}_n$ for demodulation can be devised. Note that, since adjacent observation intervals overlap, $\{\mathbf{r}_n\}$ are not independent.

Using (5)–(7), we may write

$$\mathbf{r}_n = b_{1,n}\mathbf{v}_1^0 + \sum_{k=1}^{K}(b_{k,n-1}\mathbf{v}_k^{-1} + b_{k,n+1}\mathbf{v}_k^1) + \sum_{k=2}^{K} b_{k,n}\mathbf{v}_k^0 + \mathbf{w}_n. \tag{8}$$

The preceding equation shows that the received vector may be written in terms of a *generic* equivalent synchronous model

$$\mathbf{r}_n = b_0[n]\mathbf{u}_0 + \sum_{j=1}^{J} b_j[n]\mathbf{u}_j + \mathbf{w}_n \tag{9}$$

where $b_0[n]$ is the *desired bit* that we wish to demodulate, $\mathbf{u}_0$ is the vector modulating it, and, for $1 \le j \le J$, $b_j[n]$ are interfering bits due to intersymbol interference and multiple-access interference, and $\mathbf{u}_j$ are interference vectors modulating these bits. The correspondence between (8) and (9) is immediate. Note that the desired bit $b_0[n] = b_{1,n}$ and that the number of interference vectors is $J = 3K - 1$.

For convenience of notation, we mostly work with the generic model (9), hiding the fine structure of (8) unless it is called for.

C. MMSE Demodulation

Letting $\langle , \rangle$ denote the standard inner product in Euclidean space, the linear MMSE demodulator for the equivalent synchronous system (9) corresponds to an estimate

$$\hat{b}_0[n] = \mathrm{sgn}(\langle \mathbf{c}, \mathbf{r}_n \rangle) \tag{10}$$

where the correlator $\mathbf{c}$ minimizes the MSE $E\{(\langle \mathbf{c}, \mathbf{r}_n\rangle - b_0[n])^2\}$ between the decision statistic and the desired bit $b_0[n]$. The desired bit would ideally be taken to be the bit of the desired user that falls completely within the current observation interval.

The MMSE solution is given by

$$\mathbf{c}_{\mathrm{MMSE}} = \mathbf{R}^{-1}\tilde{\mathbf{u}} \tag{11}$$

where $\mathbf{R} = E[\mathbf{r}_n\mathbf{r}_n^T]$ is the correlation matrix for the observation vectors and $\tilde{\mathbf{u}} = E[b_0[n]\mathbf{r}_n]$ is the result of correlating the desired bit sequence with the observation vectors. For uncorrelated $\{b_j[n]\}$ (which is the case if all users transmit random bits), $\tilde{\mathbf{u}} = \mathbf{u}_0$.[4] The MSE achieved by the MMSE solution is given by [10]

$$\eta = 1 - \tilde{\mathbf{u}}^T\mathbf{R}^{-1}\tilde{\mathbf{u}} = 1 - \mathbf{c}_{\mathrm{MMSE}}^T\tilde{\mathbf{u}}. \tag{12}$$

In practice, the MMSE receiver is obtained by training with respect to a known sequence $b_0[n]$, followed by decision-directed adaptation using the estimate $\hat{b}_0[n]$ for further adaptation. The adaptive implementation can be realized using a number of standard algorithms, e.g., stochastic gradient, least squares, and recursive least squares (see [4] for descriptions of these algorithms).

III. Acquisition Algorithm

In the acquisition phase the desired user sends a training sequence $t[n]$ known to the receiver except for a timing uncertainty bounded by P bit intervals. In terms of the equivalent synchronous model (8), the bit of the desired user falling completely within the nth observation interval is $b_{1,n} = t[n+i^*]$, where $i^* \in \{0, 1, \cdots, P-1\}$ is unknown and the signal vector modulated by this bit is $\mathbf{v}_1^0$. The other two bits

[4] The case of periodic bit sequences is addressed in Sections III and IV.

of the desired user falling in this observation interval are then given by $b_{1,\,n-1} = t[n+i^*-1]$ and $b_{1,\,n+1} = t[n+i^*+1]$, and the signal vectors modulating these bits are given by $\mathbf{v}_1^{-1}$ and $\mathbf{v}_1^1$, respectively. Since the phase i^* of the training sequence is *not* known while in acquisition mode, we run P adaptive algorithms, each corresponding to one of the following P hypotheses about the phase of the training sequence:

$$H_i\colon b_0[n] = t[n+i], \qquad i = 0, 1, \cdots, P-1. \tag{13}$$

Assuming that the bits of the training sequence are uncorrelated with those of other users, the MMSE solution corresponding to the ith hypothesis is given by

$$\mathbf{c}_i = \mathbf{R}^{-1}\tilde{\mathbf{u}}^{(i)} \tag{14}$$

where

$$\begin{aligned}\tilde{\mathbf{u}}^{(i)} &= E\{t[n+i]\mathbf{r}_n\}\\ &= \rho(i^*, i)\mathbf{v}_1^0 + \rho(i^*-1, i)\mathbf{v}_1^{-1} + \rho(i^*+1, i)\mathbf{v}_1^1\end{aligned}$$

where $\rho(k, i) = E\{t[n+k]t[n+i]\}$ is the autocorrelation between the kth and ith phases of the training sequence. If the training sequence consists of random uncorrelated bits, then $\rho(k, i) = \delta_{k,\,i}$, where the latter denotes the Kronecker delta. In this case we obtain $\tilde{\mathbf{u}}^{(i)} = \mathbf{0}$ for $i \neq i^*-1, i^*, i^*+1$, $\tilde{\mathbf{u}}^{(i^*)} = \mathbf{v}_1^0$, $\tilde{\mathbf{u}}^{(i^*-1)} = \mathbf{v}_1^{-1}$, and $\tilde{\mathbf{u}}^{(i^*+1)} = \mathbf{v}_1^1$. Thus, the solutions for the hypotheses H_{i^*}, H_{i^*-1}, and H_{i^*+1} are simply the standard MMSE demodulators for the current, past, and future bits of the desired user, respectively. The solution under all other hypotheses is the zero correlator, which yields $\mathrm{MSE} = 1$.

If the training sequence is periodic with period P, the statistical correlations $\rho(k, i)$ are interpreted as empirical correlations over a multiple of the period of the training sequence, so that $\rho(k, i) = \rho_p(k-i)$, where $\rho_p(k)$ is the periodic autocorrelation of the training sequence, given by

$$\rho_p(k) = (1/P)\sum_{i=0}^{P-1} t[i]t[i+k]. \tag{15}$$

In this case we obtain that

$$\tilde{\mathbf{u}}^{(i)} = \rho_p(i^*-i)\mathbf{v}_1^0 + \rho_p(i^*-1-i)\mathbf{v}_1^{-1} + \rho_p(i^*+1-i)\mathbf{v}_1^1.$$

Note that addition and subtraction of indexes for a periodic training sequence are always modulo P.

For binary training sequences, $\rho_p(k)$ will not, in general, be zero for all nonzero k. However, we can design the training sequence such that the magnitude of these values are much less than $\rho_p(0) = 1$. Such good periodic autocorrelation properties could be achieved, for instance, by choosing the training sequence to be (a segment of) a pseudonoise sequence. While optimization of training sequences is not our focus here, note that it is possible to obtain nonbinary training sequences whose periodic autocorrelation function vanishes at all nonzero time lags (e.g., see [18] for bounds on achievable autocorrelation and cross-correlation functions, as well as for sets of sequences that achieve the bounds).

For any well-designed training sequence, therefore, observations similar to those for random training sequences would apply, so that the MSE under the hypotheses H_i, $i \neq i^*-1, i^*, i^*+1$ is expected to be significantly larger than that under the hypotheses H_{i^*}, H_{i^*+1}, and H_{i^*-1}. In most cases the MSE under hypothesis H_{i^*} will be the smallest because of the larger signal energy in the bit of the desired user that falls completely in the observation interval. In some cases, however, this larger signal energy may be offset by bad cross-correlation peaks due to the interference, so that H_{i^*+1} or H_{i^*-1} may yield a smaller MSE than H_{i^*}. In order to avoid having to make a possibly bad choice between comparable hypotheses, we combine the decision statistic for the best hypothesis with that of the best *adjacent* hypothesis. For instance, suppose that hypotheses i and $i-1$ are being combined. If $\tilde{\mathbf{c}}_i^T\mathbf{r}_n$ is a decision statistic for bit $b_0[n]$, then $\tilde{\mathbf{c}}_{i-1}^T\mathbf{r}_n$ is the decision statistic for $b_0[n-1]$. Thus, in order to generate a decision statistic for the bit $b_0[n]$ using $\tilde{\mathbf{c}}_{i-1}$, we must use $\tilde{\mathbf{c}}_{i-1}^T\mathbf{r}_{n+1}$. This leads to the following bit estimate

$$\hat{b}_0[n] = \mathrm{sgn}(\tilde{\mathbf{c}}_i^T\mathbf{r}_n + \tilde{\mathbf{c}}_{i-1}^T\mathbf{r}_{n+1}). \tag{16}$$

The acquisition algorithm can now be described using the following steps.

Step 1 (Computation of MMSE Solutions): Compute the MMSE solution and corresponding MSE under each hypothesis H_i

$$\begin{aligned}\mathbf{c}_i &= \mathbf{R}^{-1}\tilde{\mathbf{u}}^{(i)}\\ \eta_i &= 1 - \tilde{\mathbf{c}}_i^T\tilde{\mathbf{u}}^{(i)}.\end{aligned} \tag{17}$$

Step 2 (Finding the Best Hypothesis): Let $H_{i_{\min}}$ denote the hypothesis with the smallest MSE, i.e.,

$$i_{\min} = \arg\min_i \eta_i.$$

Step 3 (Combining Rule): If $\eta_{i_{\min}-1} < \eta_{i_{\min}+1}$, combine $H_{i_{\min}}$ and $H_{i_{\min}-1}$ to obtain the bit estimate

$$\hat{b}_0[n] = \mathrm{sgn}(\tilde{\mathbf{c}}_{i_{\min}}^T\mathbf{r}_n + \tilde{\mathbf{c}}_{i_{\min}-1}^T\mathbf{r}_{n+1}). \tag{18}$$

If $\eta_{i_{\min}-1} \geq \eta_{i_{\min}+1}$, then combine $H_{i_{\min}}$ and $H_{i_{\min}+1}$ to obtain the bit estimate

$$\hat{b}_0[n] = \mathrm{sgn}(\tilde{\mathbf{c}}_{i_{\min}}^T\mathbf{r}_n + \tilde{\mathbf{c}}_{i_{\min}+1}^T\mathbf{r}_{n-1}). \tag{19}$$

The preceding combining rule means that the demodulator is effectively using a larger observation interval of length $3T$. The combined solution is *not* equivalent to the MMSE receiver corresponding to this larger observation interval because it ignores the correlation between the observation vectors in the two intervals being combined. However, assuming correct acquisition, it is expected to provide better performance than the MMSE solution for a length $2T$ observation interval.

We have found that the stochastic gradient algorithm converges too slowly to provide good performance for a large number of interfering users, especially in a near–far situation. We therefore consider faster, but more complex, least-squares adaptation. Note that the complexity of P (block or recursive) least-squares algorithms that *share* the computation of the inverse of the correlation matrix is not much larger than that of P stochastic gradient algorithms if P is comparable to N.

For least-squares adaptation, in step 1 of the algorithm, $\mathbf{R}$ is replaced by the empirical average

$$\hat{\mathbf{R}} = (1/M) \sum_{n=1}^{M} \mathbf{r}_n \mathbf{r}_n^T \tag{20}$$

and $\tilde{\mathbf{u}}^i$ by the empirical average

$$\hat{\mathbf{u}}^{(i)} = (1/M) \sum_{n=1}^{M} t[n+i]\mathbf{r}_n. \tag{21}$$

The corresponding correlators and MSE estimates, denoted by $\hat{\mathbf{c}}^{(i)}$ and $\hat{\eta}_i$, respectively, are used in steps 2 and 3 of the algorithm. Note that the factors of $1/M$ in (20) and (21) are redundant, but are inserted to make the averaging interpretation transparent.

1) Effect of Chip Asynchronism: For the chip-asynchronous system with chip rate sampling, the desired signal vector is given by [see (6) and (8)]

$$\mathbf{u}_0 = A_1[(1-\delta_1)\mathbf{T}_R^{n_1}\mathbf{a}_1 + \delta_1 \mathbf{T}_R^{n_1+1}\mathbf{a}_1].$$

Assuming that the adjacent shifts of the spreading sequence appearing above are approximately orthogonal, the ratio r of the signal energy compared to a chip-synchronous system is

$$r = 10 \log_{10}[(1-\delta_1)^2 + \delta_1^2] \text{ dB}. \tag{22}$$

Evaluating (22) for $\delta_1 = 0.5$ (the worst-case value for chip rate sampling), $r \approx -3$, i.e., we incur a loss of about 3 dB relative to a system with perfect chip timing for the desired user. One method of alleviating this problem is to sample at rate m/T_c, i.e., at a multiple of the chip rate. A rough idea of the performance loss associated with a given value of m can be obtained as follows. Decomposing the samples into m chip rate substreams, the substream that is closest in chip timing to the desired user will have a worst-case chip offset $\delta_1 = 1/(2m)$. Assuming that the detector manages to extract the signal energy from this substream, the expected performance loss due to asynchronism can be computed using (22). For $m = 2$, the performance loss is 2 dB, while for $m = 4$, the loss goes down to 1 dB.

Direct application of our acquisition algorithm to the received vector obtained over a given observation interval using m/T_c sampling would cause an m-fold increase in the correlator lengths, leading to higher complexity and slower convergence. A simpler solution is to run the acquisition algorithm in parallel on m chip rate substreams and to somehow combine the hypotheses obtained from each of the m algorithms. One possibility, for example, is to sum up the combined decision statistics (18) or (19) from all m algorithms (using majority logic to eliminate inconsistent outcomes if necessary). Optimization and detailed performance evaluation of algorithms for faster-than-chip-rate sampling (and determining their impact on system capacity) is an interesting topic for future work.

Note that standard feedback-based delay tracking (e.g., [20, Sec. 3.5]) does not work for chip timing recovery in this context, since even a small adjustment of the chip sample times leads to a completely different correlation matrix. Thus, an interference-suppressing correlator that works well for a given set of relative delays would not, in general, work well if the sampling times are adjusted by a feedback loop. This problem does not occur for the conventional receiver because it does not attempt to suppress interference.

2) Relation to Cyclic Equalization: Cyclic equalization [13], [16] is similar to our scheme in that a training sequence of period P is used for fast startup of adaptive equalizers in single-user narrow-band systems. Ignoring noise, this results in a periodic received signal with period PT. The observation interval for cyclic equalization is taken to be of length PT (corresponding to P symbol-spaced taps), so that one complete period of the received signal is observed. This implies that equalizers obtained under different hypothesized phases for the training sequence would be cyclically shifted versions of each other, so that it is only necessary to consider one such phase. After the initial adaptation, the "correct" symbol timing and equalizer setting for random bits is obtained by cyclically rotating the taps so that the largest tap is at the center. These ideas do not apply to the multiuser wide-band system considered here for two key reasons: 1) transmission of a periodic training sequence by the desired user need not lead to a periodic received signal since there could be interferers transmitting random bits and 2) even if the received signal were periodic with period PT, an observation interval large enough to accommodate a complete period would correspond to an excessively large number of taps (of the order of PN) because of the wide-band nature of the signal. These differences necessitate the more general hypothesis testing approach adopted in this paper.

IV. Performance Analysis

Acquisition is successful if the estimates $\hat{\eta}_i$ of the MSE are good enough that the combined decision rule extracts the signal energy from at least one complete symbol while providing interference suppression. Thus, the hypothesis H_{i^*} corresponding to one complete symbol falling in the observation interval should be among the two being combined in step 3 of the algorithm. We provide an approximate analysis for estimation of the acquisition error probability in Section IV-A. The demodulation performance for the combined decision rule is analyzed in Section IV-B. In Section IV-C we show that the MMSE detector is near–far resistant even when there are multiple users transmitting periodic training sequences.

A. Acquisition Performance

Recall that i^* denotes the hypothesis corresponding to the symbol that falls completely in the observation interval. Let $i_{\min}$ and i_{next} denote the best hypothesis and best adjacent hypothesis, respectively, chosen by an adaptive implementation. *Correct acquisition* is defined to occur if *either* $i_{\min} = i^*$ *or* $i_{\text{next}} = i^*$, i.e., if i^* is included among the two hypotheses being combined. This guarantees that the signal energy from one complete bit is extracted by the combining rule, thus ensuring good demodulation performance.

We now estimate the probability of *acquisition error*, which occurs if $i_{\min} \neq i^*$ *and* $i_{\text{next}} \neq i^*$. Let ϵ_i denote the error in estimating the MSE under hypothesis H_i using an adaptive implementation, i.e., let

$$\epsilon_i = \hat{\eta}_i - \eta_i. \tag{23}$$

For least-squares adaptation, examination of the detailed structure of the estimates $\hat{\mathbf{R}}$, $\hat{\mathbf{u}}^{(i)}$, and $\hat{\mathbf{c}}^{(i)}$ yields that the order $1/M$ term in ϵ_i is roughly zero-mean Gaussian with variance of order $1/M^2$, using the central limit theorem. The mean of the order $1/M^2$ term yields a nonzero bias of order $1/M^2$ in the estimated MSE. Neglecting terms smaller than order $1/M^2$, we arrive at the approximation that the errors $\{\epsilon_i\}$ in estimating the MSE under the different hypotheses (and therefore the MSE estimates $\{\hat{\eta}_i\}$) are well modeled as jointly Gaussian. This approximation is used to provide an estimate of the probability of acquisition error, as described in the following.

For each $i \neq i^*$, denote the mean and variance of the random variable $\hat{\eta}_i - \hat{\eta}_{i^*}$ by μ_i and s_i^2, respectively. The preceding random variable is Gaussian if the MSE estimates $\{\hat{\eta}_i\}$ are jointly Gaussian. Under this Gaussian approximation, therefore, we obtain that the probability of choosing H_i over H_{i^*}

$$q(i) = P[\hat{\eta}_{i^*} > \hat{\eta}_i] \approx Q\left(\frac{\mu_i}{s_i}\right). \tag{24}$$

Assuming that (24) provides an accurate estimate of the probabilities $q(i)$, we can obtain a union bound on the probability of acquisition error as follows:

$$q_{e,\,acq} = P[i_{\min} \neq i^*,\, i_{\text{next}} \neq i^*] \leq \sum_{i \neq i^*,\, i^* \pm 1} q(i). \tag{25}$$

The reasoning is as follows. If $i_{\min} \neq i^*$, then either 1) $i_{\min} = i^* \pm 1$ or 2) $i_{\min} = i$, for some $i \neq i^*, i^* \pm 1$. The probability of event 2) occurring is clearly included in the preceding union bound. On the other hand, if 1) occurs, then $i_{\text{next}} \neq i^*$ only if $\hat{\eta}_i < \hat{\eta}_{i^*}$ for some $i \neq i^*$, $i^* \pm 1$, which is again an event included in the union bound.

Since analytical computation of $\{\mu_i\}$ and $\{s_i^2\}$ is difficult, we estimate them empirically using multiple simulation runs of the least-squares implementation with M training symbols. Note that reliable estimates of these quantities require much fewer runs than direct simulation for estimation of small acquisition error probabilities (see Section V for a numerical comparison of estimates from approximate analysis and direct simulation).

B. Demodulation Performance

We are interested in analyzing the combined decision rule [(18) or (19)]. For example, consider the decision statistic

$$Z_n = \tilde{\mathbf{c}}_{i_{\min}}^T \mathbf{r}_n + \tilde{\mathbf{c}}_{i_{\min}-1}^T \mathbf{r}_{n+1}.$$

We may rewrite $Z_n = \mathbf{c}^T \tilde{\mathbf{r}}_n$, where $\tilde{\mathbf{r}}_n$ is the received vector corresponding to the union of the observation intervals corresponding to $\mathbf{r}_n$ and $\mathbf{r}_{n+1}$ (the union is an observation interval of length $3T$), and $\mathbf{c}$ is the sum of suitably translated versions of $\tilde{\mathbf{c}}_{i_{\min}}$ and $\tilde{\mathbf{c}}_{i_{\min}-1}$. It is straighforward to derive an equivalent synchronous model of the generic form (9) for this longer observation interval. We may now apply the standard analysis for a linear correlator for the generic equivalent synchronous model as follows.

For the model (9), the average error probability for a linear correlator can be computed by conditioning and then by averaging over the bits $\{b_j[n]\}$ [10]. This is computationally intensive for a large number of interference vectors. An approximation that has been found to give excellent results [15] is $Q(\sqrt{\text{SIR}})$, where SIR denotes the signal-to-interference ratio at the output of the correlator, given for a linear demodulator $\mathbf{c}$ by

$$\text{SIR} = \frac{|\langle \mathbf{c},\, \mathbf{u}_0\rangle|^2}{\sum_{j=1}^{J} |\langle \mathbf{c},\, \mathbf{u}_j\rangle|^2 + \sigma^2 \|\mathbf{c}\|^2}.$$

We restrict attention to SIR in the numerical results reported in Section V, having checked via a few sample computations that it is an excellent predictor of error probability performance.

C. Periodic Training for Multiple Users

Periodic training sequences may be used during acquisition either for convenience or because of large initial timing uncertainties. Suppose that several users (including the desired user) are in acquisition mode and are therefore transmitting periodic bit sequences. In this subsection we prove that, under a mild condition on the transmitted bit sequences, the MMSE solution obtained in this situation for a given user in acquisition mode is near–far resistant against all users, whether they transmit periodic or random bit sequences.

Consider first the standard setting of uncorrelated $\{b_j[n]\}$ in the equivalent synchronous model (9). The MSE corresponding to a linear correlator $\mathbf{c}$ can be shown to be

$$\text{MSE(random)} = E[(\langle \mathbf{c},\, \mathbf{r}_n\rangle - b_0[n])^2] = \mathbf{a}^T\mathbf{a} + \sigma^2\|\mathbf{c}\|^2 \tag{26}$$

where $\mathbf{a} = (\langle \mathbf{c}, \mathbf{u}_0\rangle - 1, \langle \mathbf{c}, \mathbf{u}_1\rangle, \cdots, \langle \mathbf{c}, \mathbf{u}_J\rangle)^T$. The first component of $\mathbf{a}$ represents the error in the scaling for the desired symbol at the output, while the others represent the contribution of the interference vectors to the output. The cost function in (26) therefore trades off interference suppression represented by the energy of $\mathbf{a}$ versus the noise enhancement represented by the energy of $\mathbf{c}$. The key to our proof of near–far resistance of the MMSE solution for correlated bit sequences is to show that the MSE in the latter situation exhibits a form similar to (26).

For periodic training sequences, we may assume (with appropriate numbering of the user vectors) that the bit sequences $\{b_j[n]\}$ in (9) are periodic with period P for $0 \leq j \leq L$, and are random for $L+1 \leq j \leq J$. If there are K_a users (including the desired user) transmitting periodic bit sequences, we have $L + 1 = 3K_a$ for an observation interval of length $2T$, since three bits of each user fall into a given interval. In this setting the statistical average for computing the MSE must consider P consecutive observation intervals in order to account for the periodicity, so that

$$\text{MSE(periodic)} = (1/P) \sum_{n=0}^{P-1} E[(\langle \mathbf{c},\, \mathbf{r}_n\rangle - b_0[n])^2]. \tag{27}$$

For $0 \leq j \leq J$, let $\mathbf{b}_j = (b_j[0], \cdots, b_j[P-1])^T$ denote the sequence of bits modulating the interference vector $\mathbf{u}_j$ in P successive observation intervals. The vectors $\mathbf{b}_j$ are deterministic for $0 \leq j \leq L$ (because of periodicity) and random for $L+1 \leq j \leq J$. Let $\mathbf{B}$ denote the (partly random) $P \times (J+1)$ matrix with columns $\mathbf{b}_0, \cdots, \mathbf{b}_J$ and let $\mathbf{B}_p$ denote the (deterministic) $P \times (L+1)$ matrix with columns $\mathbf{b}_0, \cdots, \mathbf{b}_L$ corresponding to the periodic bit sequences. The MSE (27) can be now rewritten as follows:

$$\text{MSE(periodic)} = \mathbf{a}^T \mathbf{R}_b \mathbf{a} + \sigma^2 \|\mathbf{c}\|^2 \tag{28}$$

where $\mathbf{R}_b = (1/P)E[\mathbf{B}^T\mathbf{B}]$ is the matrix of normalized cross correlations among the bit sequences modulating the $J+1$ signal vectors. Using the uncorrelatedness of the random bit sequences, the latter matrix can be written as

$$\mathbf{R}_b = \begin{bmatrix} \mathbf{R}_{b,\,p} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{29}$$

where $\mathbf{R}_{b,\,p} = (1/P)\mathbf{B}_p^T\mathbf{B}_p$. [If all users transmit random bits, then $\mathbf{R}_b = \mathbf{I}$, which yields (26).]

The MSE in (28) exhibits the same tradeoff as (26) between interference suppression and noise enhancement, except that the energy of $\mathbf{a}$ is now computed using a "twisted" norm that reflects the periodic cross correlations among the bit sequences. We therefore expect the properties of the MMSE solution to carry over to this situation as well, as long as the twisted norm is nondegenerate (i.e., as long as only the zero vector can have norm zero). The latter is equivalent to $\mathbf{R}_b$ being positive definite, which, using (29), is true if and only if $\mathbf{R}_{b,\,p}$ is positive definite (i.e., if and only if the periodic bit sequences $\mathbf{b}_0, \cdots, \mathbf{b}_L$ are linearly independent). This leads to the following result, which, like the formula (28), applies in general to correlated bit sequences, although our focus here is on correlations arising due to periodicity.

1) Proposition: The MMSE detector has nonzero near–far resistance [7] if the following conditions hold.

Condition C1 (Linear Independence of Signal Vectors): The desired signal vector $\mathbf{u}_0$ is linearly independent of the interference vectors $\mathbf{u}_1, \cdots, \mathbf{u}_J$.

Condition C2 (Nondegenerate Bit Correlations): The matrix $\mathbf{R}_b$ of bit correlations is positive definite. For periodic bit sequences, this is equivalent to the desired bit vector $\mathbf{b}_0$ being linearly independent of the bit vectors $\mathbf{b}_1, \cdots, \mathbf{b}_L$.

The proof is sketched in the Appendix.

2) Caveat: The following stronger notion of near–far resistance is proved in [10] under the assumption of uncorrelated bits: even for nonvanishing noise levels, the MMSE detector forces the contribution of a strong interferer to the output to zero (i.e., $\langle \mathbf{c}, \mathbf{u}_j \rangle \rightarrow 0$ as $\|\mathbf{u}_j\|/\|\mathbf{u}_0\| \rightarrow \infty$). This result *need not hold* for correlated bits—while the output due to a strong interferer is small by virtue of the cost function (28), it need not tend to zero if the interference contributions to the output are coupled to the contribution of the desired vector by a nondiagonal $\mathbf{R}_b$.

3) Choice of Periodic Training Sequences: Since the number of P-dimensional bit vectors $\mathbf{b}_j$ involved in condition C2 is $3K_a$, a rule of thumb for choice of P is that $P \geq 3K_a$. For example, if the total number of users K is of the order of the processing gain N, and if the number of users in acquisition mode is about $K/3$, then P can be chosen to be of the same order as N (e.g., an arbitrary, but convenient, choice of training *bit* sequence equal to spreading *chip* sequence would be expected to work). If different users use periodic sequences of different periods, P must be replaced by the lowest common multiple of these periods, which can be quite large. Thus, condition C2 becomes easier to satisfy if different periods are used (rather than the arbitrary choice of $P = N$). In general, design of training sequences for a multiuser application is an interesting problem for further investigation.

V. Numerical Results

We consider a symbol- and chip-asynchronous system with processing gain $N = 10$ and number of users $K = 8$. The spreading sequences and delays are chosen randomly and then fixed. Other choices of spreading sequences and delays have been found to yield qualitatively similar results. The desired user has E_b/N_0 of 17 dB (corresponding to an SIR of 20 dB in a chip-synchronous system with no interference). Unless specified otherwise, we assume a severe near–far problem, with four interferers with power 20 dB higher than that of the desired user and three interferers with power equal to that of the desired user. We do not optimize over training sequences—the training sequence for any user in acquisition mode is a periodic sequence arbitrarily chosen to be exactly the spreading sequence for that user, so that the training sequence period $P = 10$.

Performance (averaged over a number of simulation runs) is plotted as a function of the number of training symbols M, where M ranges from 20 to 100 (the correlators under each hypothesis have $2N = 20$ taps). The performance of the algorithm using the steady-state solutions (14) is used to provide a benchmark for the best possible SIR achievable using least-squares adaptation.

We assume first that only the desired user is in acquisition mode, with the remaining users transmitting random bits. The selection probability for each hypothesis, as well as the acquisition error probability, are estimated via direct simulation using 100 000 runs of the least-squares adaptation. These quantities are also estimated via the union bound estimate (24) and (25), where the means μ_i and variances s_i^2 are obtained using 1000 simulation runs. Fig. 1 shows the estimates of acquisition error probability. Since the union bound estimate is consistently larger than the estimate from direct simulations while exhibiting a similar steep drop in acquisition error probability with M, the former represents a computationally attractive tool for conservative system design (i.e., for choosing the duration of training).

In Fig. 2 we plot estimates of the probability that the least-squares algorithm selects each of the adjacent hypotheses ($H_{i^*\pm 1}$) over the hypothesis H_{i^*} corresponding to a complete bit falling in the observation interval (by virtue of the combining rule, such a selection would not necessarily lead to

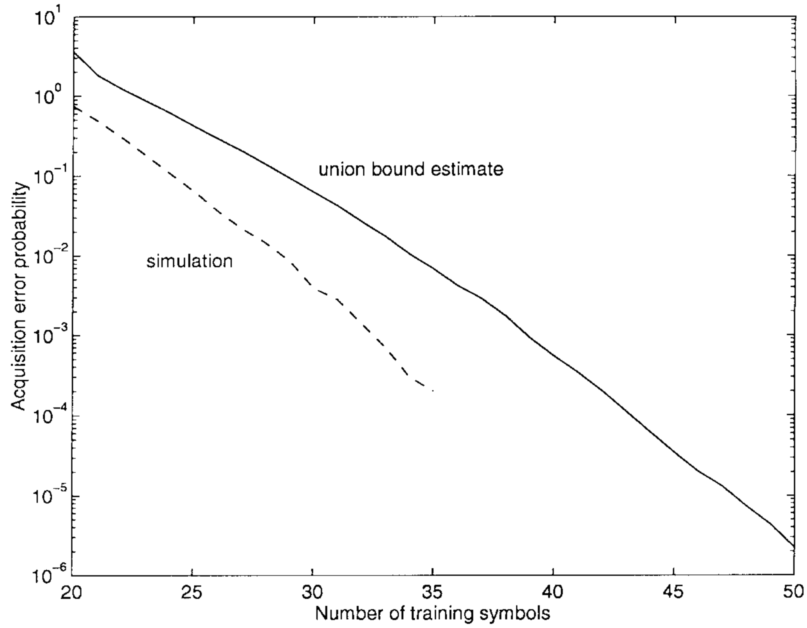

Fig. 1. Acquisition error probability as a function of the number of training symbols.

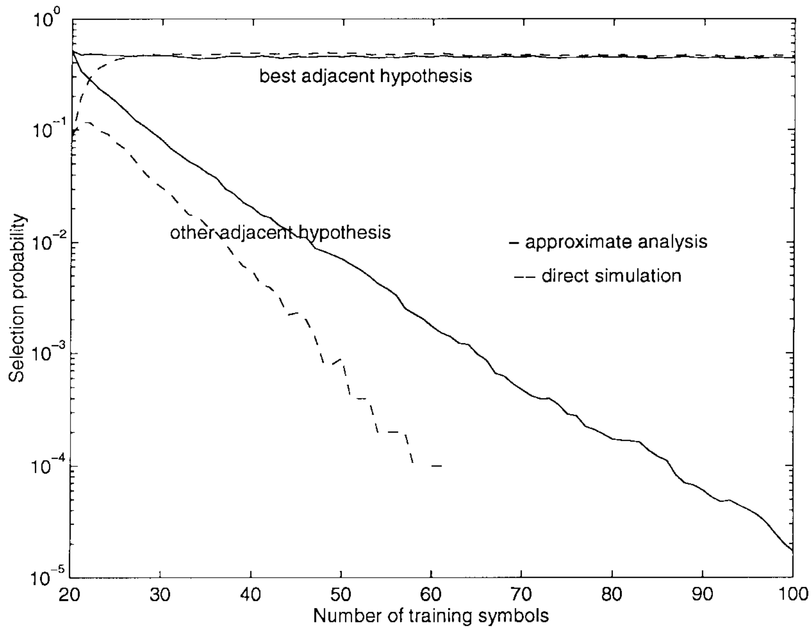

Fig. 2. Probability of selecting adjacent hypotheses $H_{i^*\pm 1}$ over the hypothesis H_{i^*} as a function of the number of training symbols.

acquisition error). There are two important points brought out by the figure.

1) The adaptive implementation prefers the best adjacent hypothesis frequently (with probability around 0.46) over the hypothesis H_{i^*}. Since these two hypotheses are comparable in terms of MSE (the steady-state values are 0.1498 and 0.1465, respectively), the combining rule in step 3 of the algorithm plays a valuable role in ensuring that both are selected.
2) The approximate analysis is again conservative, overestimating the probability of choosing the poorer adjacent hypothesis, which corresponds to a steady-state MSE of 0.3817.

Demodulation performance is quantified by calculating the SIR (absolute values averaged over 100 runs and the resulting average expressed in decibels) as a function of the number of symbols M used in the least-squares algorithm. For our combined decision rule, the averaging accounts for acquisition errors by setting the SIR to zero (the lowest possible value) in the event of an acquisition error. Fig. 3 displays the following SIR's:[5]

1) SIR versus M for our acquisition-based combined decision rule (which effectively spans a $3T$ observation

[5]The conventional matched filter receiver is useless in the setting of Fig. 3 and is therefore not shown in the plot.

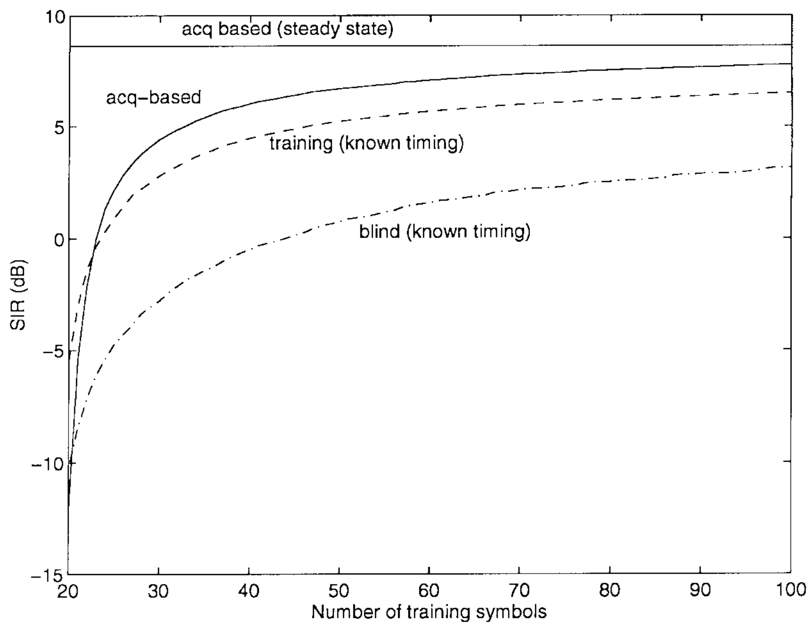

Fig. 3. SIR (in decibels) as a function of the number of training symbols with a severe near–far problem, with all interferers transmitting random bits.

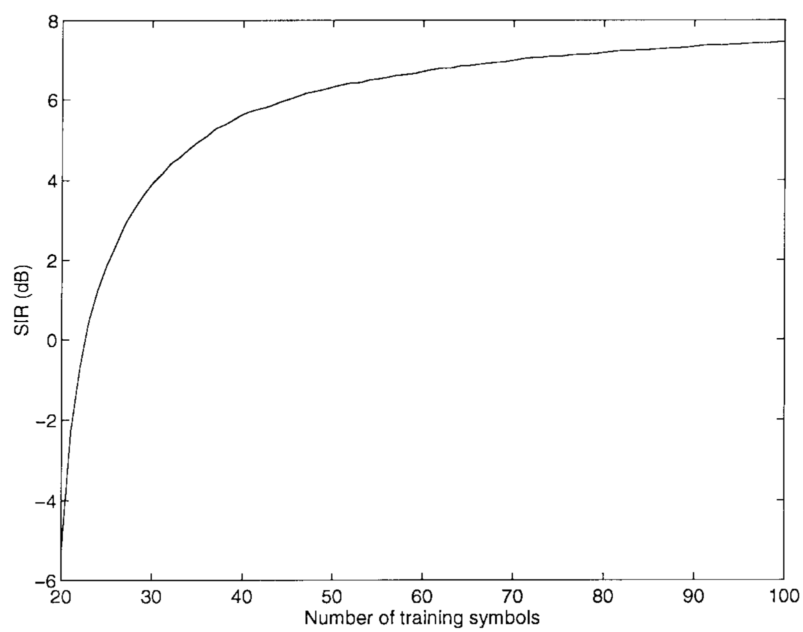

Fig. 4. SIR (in decibels) as a function of the number of training symbols with a severe near–far problem, with two of the strong interferers transmitting periodic bit sequences.

interval), together with the corresponding steady-state SIR;

2) SIR for training-based least-squares adaptation with $2T$ interval and known symbol timing;
3) SIR for blind least-squares adaptation (as in [6]) with length $2T$ observation interval and *known spreading waveform and chip timing.*

Fig. 3 illustrates that, after a short initial transient of poor SIR due to acquisition errors, the performance of the acquisition-based combined decision rule is better than that of a training-based correlator with known symbol timing and a length $2T$ observation interval. This is because of the larger (length $3T$) observation interval used by the combined rule. Indeed, the SIR for the ideal MMSE solution for a $2T$ interval (not shown in the plot) is 1 dB worse than the steady-state SIR for the acquisition-based scheme shown in the figure. Note also that the blind detector converges much slower than either training-based scheme.

The preceding results can be used to determine the training duration required for reliable acquisition and demodulation. From the approximate analysis of acquisition error probability (Fig. 1), we obtain, for instance, that an acquisition error probability of 10^{-6} is obtained when $M = 55$. From the simulation results displayed in Fig. 3, the average SIR is 6.8

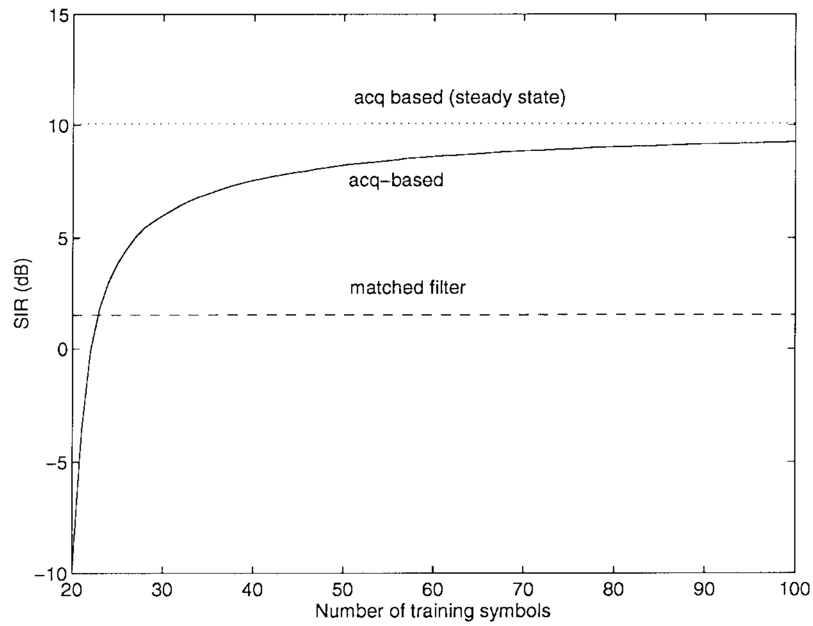

Fig. 5. SIR (in decibels) as a function of the number of training symbols with all users at equal power.

dB for $M = 55$. Using the $Q(\sqrt{\mathrm{SIR}})$ approximation, this corresponds to an average error probability of about 10^{-2}, which is the value recommended typically for switching to decision-directed adaptation. Thus, by the time the SIR is large enough to switch to decision-directed adaptation, the probability of acquisition error is negligible. This implies that there is essentially no performance penalty due to lack of knowledge of symbol timing.

1) Multiple Users in Acquisition Mode: Assume now that two of the strong interferers are also in acquisition mode and are transmitting periodic training sequences with period $P = 10$. The averaged SIR for our combined decision rule, together with the corresponding steady-state value, is plotted in Fig. 4. Comparing with Fig. 3, we notice little degradation in performance due to the correlation between the bit sequences of the desired and interfering users.

2) Performance with Equal Powers: In Fig. 5 we compare the averaged SIR of the combined decision rule with the SIR of the conventional matched filter receiver when all eight users are at equal powers. Clearly, there is a large performance gain due to interference suppression even with perfect power control.

3) Combating Chip Asynchronism: Sampling at twice the chip rate and combining the decision statistics from two parallel algorithms working with chip rate substreams, as described in Section III, leads to an SIR gain of about 1.4 dB relative to the chip rate algorithm, while being 1 dB worse than the SIR of a hypothetical chip rate algorithm chip synchronized to the desired user. These values are for the specific setting considered in Fig. 3 but, as argued in Section III, we would typically expect about 1-dB gain from sampling at twice the chip rate.

VI. Conclusions

The timing acquisition scheme presented here enables asynchronous near–far-resistant multiple-access communication without the need for feedback between receiver and transmitter. While it applies to any CDMA system with short spreading sequences, we anticipate that a major application could be as a means of simplifying network protocols and architectures for unstructured (or *ad hoc*) peer-to-peer wireless networks. While exploring this and other applications is a topic for future research, some advantages of using our scheme are sketched below. Our algorithm makes it possible for a receiver to look for packets sent by a given transmitter by using its periodic training sequence as a signature, which eliminates the need for a side channel (secure, robust, and efficient implementation of side channels in *ad hoc* networks is still an open issue).[6] Further, even if approximate symbol timing is maintained in some fashion, relative mobility between transmitter and receiver would lead to symbol timing uncertainty, especially for high symbol rates.

Optimization of the set of periodic training sequences used for acquisition with multiple users and large timing uncertainties is an interesting topic for further investigation. Another important topic is extending the acquisition algorithm to incorporate modifications to the MMSE adaptive algorithm required to deal with fast fading [5], [22], [23] (standard MMSE decision-directed adaptation does not work with fast fading [22]). Finally, the issue of enhancing performance at reasonable complexity using antenna diversity and faster sampling deserves further thought.

Appendix

We sketch the proof of the proposition in Section IV-C here. Near–far resistance is the worst-case (over interference

[6] While the training and spreading sequences for a given transmitter must be the same over a packet for our adaptive method to work, security issues could be addressed by varying them over different packets, e.g., as a function of the identity of the transmitter or receiver, and the packet number or time of day.

amplitudes) asymptotic rate of exponential decay of the error probability of a multiuser detector as the noise level vanishes [7], [10]. Condition C1 guarantees the existence (and nonzero near–far resistance) of a zero-forcing (ZF) solution [7], [10], which corresponds to $\mathbf{a} = 0$. For uncorrelated bit sequences, the MMSE detector degenerates to the ZF solution [i.e., $\mathbf{a} = 0$ minimizes (26)] as the noise level $\sigma^2 \to 0$ [10], and therefore inherits the near–far resistance of the ZF solution. More generally, we obtain from (28) that $\mathrm{MSE} \to \mathbf{a}^T \mathbf{R}_b \mathbf{a}$ as $\sigma^2 \to 0$. For $\mathbf{R}_b$ positive definite (true under C2), the MSE can be made equal to zero if and only if $\mathbf{a} = 0$, which again corresponds to the ZF solution. Note that if C2 were not satisfied, then it is possible to find nonzero $\mathbf{a}$ (which correspond to undesirable solutions $\mathbf{c}$ that are not zero-forcing) that yield $\mathrm{MSE} = 0$.

REFERENCES

[1] A. Abdulrahman, D. D. Falconer, and A. U. Sheikh, "Decision feedback equalization for CDMA in indoor wireless communications," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 698–706, May 1994.

[2] S. E. Bensley and B. Aazhang, "Maximum likelihood synchronization of a single user for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 46, pp. 392–399, March 1998.

[3] ——, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, no. 8, pp. 1009–1020, Aug. 1996.

[4] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1991.

[5] M. Honig, M. Shensa, S. Miller, and L. Milstein, "Performance of adaptive linear interference suppression for DS-CDMA in the presence of flat Rayleigh fading," in *Proc. IEEE VTC'97*, Phoenix, AZ, May 1997, pp. 2191–2195.

[6] M. L. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[7] R. Lupas and S. Verdu, "Near–far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, Apr. 1990.

[8] U. Madhow, "Blind adaptive interference suppression for the near–far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 102–112, Jan. 1997.

[9] ——, "Blind adaptive interference suppression for direct-sequence CDMA," *Proc. IEEE*, to be published.

[10] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.

[11] U. Madhow and M. B. Pursley, "Acquisition in direct-sequence spread-spectrum communication networks: An asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 39, pp. 903–912, May 1993.

[12] 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.

[13] K. H. Mueller and D. A. Spaulding, "Cyclic equalization—A new rapidly converging equalization technique for synchronous data transmission," *Bell Syst. Tech. J.*, vol. 54, pp. 369–406, Feb. 1975.

[14] R. F. Smith and S. L. Miller, "Code timing estimation in a near–far environment for direct-sequence code-division multiple-access," in *Proc. IEEE MILCOM'94*, Fort Monmouth, NJ, Oct. 1994, pp. 47–51.

[15] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, pp. 858–871, May 1997.

[16] S. U. H. Qureshi, "Fast start-up equalization with periodic training sequences," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 553–563, Sept. 1977.

[17] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685–697, May 1994.

[18] D. V. Sarwate, "Bounds on crosscorrelation and autocorrelation of sequences," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 720–724, Nov. 1979.

[19] E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 84–93, Jan. 1996.

[20] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication.* Reading, MA: Addison-Wesley, 1995.

[21] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 683–690, May 1990.

[22] L. J. Zhu and U. Madhow, "Adaptive interference suppression for DS-CDMA over a Rayleigh fading channel," in *Proc. 1997 Conf. Information Science Systems (CISS'97)*, Baltimore, MD, Mar. 1997, pp. 98–103.

[23] ——, "MMSE interference suppression for rapidly faded CDMA systems: A new formulation for differentially modulated signals," in *Proc. Int. Symp. Information Theory (ISIT'98)*, Cambridge, MA, Aug. 1998, to be published.

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

From 1990 to 1992 he was a with the University of Illinois as a Visiting Assistant Professor. From 1991 to 1994 he was with Bell Communications Research, Morristown, NJ, as a Research Scientist. Since 1994 he has been with the Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, where he is currently an Associate Professor. His research interests are in communication systems and networking, with current emphasis on wireless communcations and high-speed networks. He is an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and an Associate Editor for Detection for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Dr. Madhow is a recipient of the NSF CAREER Award.