Adaptive Iterative Multiuser Decision Feedback Detection

Michael L. Honig, *Fellow, IEEE*, Graeme K. Woodward, *Member, IEEE*, and Yakun Sun, *Student Member, IEEE*

Abstract—**Adaptive iterative receivers which combine multiuser decision-feedback detection with maximum *a posteriori* (MAP) decoding and soft feedback are presented for synchronous coded direct sequence-code-division multiple access. Both successive and parallel demodulation of users are considered. Optimal filters are derived using both minimum mean squared error and least squares (LS) criteria. The latter assumes short (repeated) spreading codes and that the users to be demodulated simultaneously transmit training sequences. The LS criterion does not require prior knowledge or estimates of spreading codes and channels. Simulation results show that the adaptive receiver can perform significantly better than the standard (soft) interference canceller, since the adaptive algorithm attempts to measure and exploit the second-order statistics between the input and output of the MAP decoder. With limited training, successive feedback and decoding performs significantly better than parallel feedback. The effect of code rate on performance is examined, and reduced-rank versions of the adaptive LS algorithms, which can reduce training overhead, are also presented.**

Index Terms—**Adaptive filters, code division multiaccess (CDMA), interference cancellation, interference suppression, iterative methods, MIMO systems, multiuser detection.**

I. INTRODUCTION

MULTIUSER decision-feedback detectors (DFDs) for direct-sequence (DS)-code-division multiple access (CDMA) based on the minimum mean squared error (MMSE) criterion have been presented in [1]–[3]. The successive-decision feedback detector (S-DFD) was presented in [1], whereas a parallel-DFD (P-DFD) based on the MMSE criterion was presented in [3]. The MMSE criterion has the attractive property that with short or repeated spreading sequences the filter coefficients can be estimated without side information about user signatures, amplitudes, and channels. It has been shown in [4]–[6] that given sufficient E_b/N_0, the MMSE DFD can perform significantly better than the analogous linear detector at relatively high loads (users/bandwidth expansion). However, at low to moderate E_b/N_0, error propagation can severely degrade the performance of the DFD.

M. L. Honig and Y. Sun are with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208 USA (e-mail: mh@ece.northwestern.edu; yksun@ece.northwestern.edu).

G. K. Woodward is with Agere Systems, North Ryde, NSW 2113, Australia (e-mail: graemew@agere.com).

In this paper, we combine *adaptive* multiuser decision feedback detection with repeated spreading sequences [2], [7] with coding and iterative (turbo) decoding as presented in [8]–[12]. Specifically, soft outputs from the DFD are used to estimate likelihoods which are interleaved and input to the MAP decoder for the convolutional code. The MAP decoder computes *a posteriori* probabilities (APPs) for each user's coded bits, which are used to generate soft estimates. These soft estimates are subsequently used to update the DFD filters, deinterleaved, and fed back through the feedback filter. This process is then iterated.

In the adaptive receiver presented, the DFD filters are *jointly optimized* according to a least squares (LS) criterion at each iteration. This is in contrast to prior work in which the feedback filter is fixed for all iterations [10], [11] or where the filters are re-estimated for each symbol [11], [12]. (See also [13], which uses a different adaptation method for filter estimation.) The resulting receiver requires only a training sequence for estimation of the filter coefficients. Prior knowledge of spreading codes, channels, and noise statistics is not required, provided that the channels and relative delays for the users of interest are stationary within each packet.

Our numerical results show that for the system parameters considered, the adaptive iterative receiver performs significantly better than the fixed DFD optimized according to the MMSE criterion, assuming *perfect* feedback [2]. This is because the adaptive algorithm attempts to exploit the joint statistics between the input symbols and the soft decisions, which are fed back through the feedback filter. We derive the MMSE DFD filters in terms of these joint statistics.

The performance of the adaptive iterative receivers is studied by simulation as a function of background signal-to-noise ratio (E_b/N_0), training, load, and code rate. Our results show that at a moderate load ($K/N = 0.75$ where K is the number of users and N is the bandwidth expansion), the iterative receivers achieve near single-user performance at an E_b/N_0 close to the lower bound corresponding to the large system capacity [14]. At larger loads, this gap widens, and the performance becomes much more sensitive to the code properties.

Parallel demodulation is also compared with successive demodulation. We present an iterative successive (IS)-DFD in which the most recently available soft decisions are fed back and used to estimate the filters. (See also [3] and [15], where related S-DFDs are presented.) Given sufficient training and iterations, both the P- and IS-DFDs exhibit the same performance. With limited training, the IS-DFD (or a hybrid which switches from the IS- to the P-DFD) performs substantially better than the P-DFD.

(a) Transmitters and AWGN channel
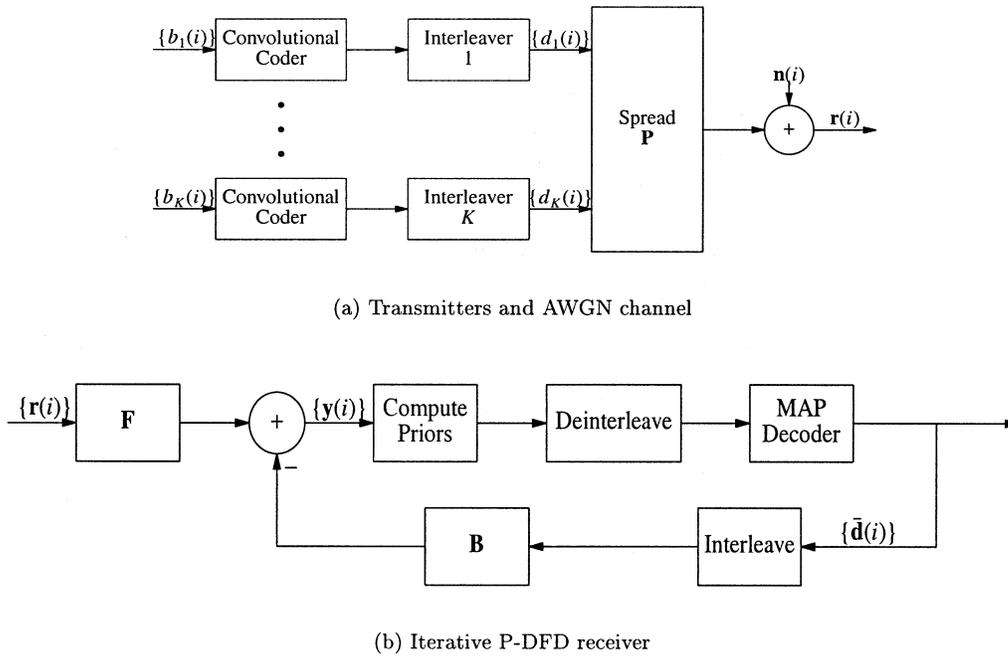

(b) Iterative P-DFD receiver

Fig. 1.　System model.

Finally, we present reduced-rank versions of the adaptive LS algorithms, which can reduce the amount of training overhead required. These algorithms are motivated by related work on reduced-rank linear interference suppression [16]. Numerical results show that with limited training the reduced-rank algorithms offer a signficant performance improvement for large systems.

The adaptive iterative receivers are described in Section II. Section III presents the adaptive algorithms, and simulation results are presented in Section VI. The effect of code rate is examined in Section VII and reduced-rank adaptive algorithms are presented in Section VIII.

II. SYSTEM MODEL AND RECEIVERS

The system model is shown in Fig. 1. For simplicity, we consider an ideal synchronous CDMA system with additive white Gaussian noise (AWGN). The extension to an asynchronous CDMA system with multipath is briefly discussed in Section IX. We will also assume that the users are received with equal powers, although this is not required for the adaptive receivers to be presented.

Each user's sequence of information symbols ($\{b_k(i)\}$ for user k) is the input to a convolutional coder (the same for each user); the coder outputs are sent through a random interleaver (different for each user), and these are transmitted through the ideal synchronous DS-CDMA channel. The received vector of N' samples during symbol interval i is given by

$$\mathbf{r}(i) = \mathbf{P}\mathbf{d}(i) + \mathbf{n}(i) \qquad (1)$$

where $\mathbf{P}$ is the $N' \times K$ matrix of spreading codes, K is the number of users, $\mathbf{d}(i)$ is the ith vector across users of interleaved coded symbols, and $\mathbf{n}(i)$ is the vector of noise samples with covariance matrix $\sigma^2\mathbf{I}$. Let $N = N'/R_C$ be the total bandwidth expansion due to coding at rate R_C and spreading. We assume

binary information and coded symbols so that the kth element of $\mathbf{d}$, $d_k(i) \in \{\pm 1\}$.

The received vector $\mathbf{r}(i)$ is the input to a nonlinear filter, consisting of feedforward and feedback matrices $\mathbf{F}^{(m)}$ and $\mathbf{B}^{(m)}$, where m denotes iteration. Referring to Fig. 1(b), the output of the DFD corresponding to user k at time i is

$$y_k^{(m)}(i) = \left(\mathbf{F}_k^{(m)}\right)^{\dagger} \mathbf{r}(i) - \left(\mathbf{B}_k^{(m)}\right)^{\dagger} \bar{\mathbf{d}}_k^{(m)}(i) \qquad (2)$$

where "$\dagger$" denotes complex conjugate transpose, $\mathbf{F}_k^{(m)}$ and $\mathbf{B}_k^{(m)}$ are the feedforward and feedback filters, respectively, and $\bar{\mathbf{d}}_k^{(m)}(i)$ is the input to the feedback filter at time i, all corresponding to the mth iteration for user k.

For the P-DFD, the vector $\bar{\mathbf{d}}_k^{(m)}$ does not depend on k, i.e., soft decisions from the preceding iteration are used for parallel cancellation. An adaptive linear receiver is used for the first iteration. For the IS-DFD, the users are demodulated successively, and the most recent decisions are used for cancellation. After each iteration, the users are reordered, and the feedback filter $\mathbf{B}_k^{(m)}$ cancels *all* demodulated users, as in the P-DFD. That is, the first $k - 1$ decisions in $\bar{\mathbf{d}}_k^{(m)}$ are from the current iteration, and the remaining decisions are from the preceding iteration.

Ideally, if $\bar{\mathbf{d}}^{(m)}(i) = \mathbf{d}(i)$, then $\mathbf{B}^{(m)}$ can be selected to cancel the multiple-access interference. The feedforward filter $\mathbf{F}^{(m)}$ then suppresses unknown (e.g., other cell) interference. Computation of $\mathbf{F}^{(m)}$, $\mathbf{B}^{(m)}$, and $\bar{\mathbf{d}}^{(m)}$ is described in Section III.

For purposes of MAP decoding, we assume that the interference plus noise at the output of the subtractor shown in Fig. 1 (corresponding to $\mathbf{y}$) is Gaussian. This assumption is reasonable when $\bar{\mathbf{d}}^{(m)}$ is close to $\mathbf{d}$ or when there are many active users. For the kth user we therefore write

$$y_k^{(m)}(i) = A_k^{(m)} d_k(i) + \xi_k^{(m)}(i) \qquad (3)$$

where $A_k^{(m)}$ is a constant and $\xi^{(m)}(i)$ is a Gaussian random variable with variance $\sigma^2_{\xi_k^{(m)}}$. Since

$$A_k^{(m)} = E\left[d_k^*(i)y_k^{(m)}\right] \tag{4}$$

and

$$\sigma^2_{\xi_k^{(m)}} = E\left\{\left[y_k^{(m)}(i) - A_k^{(m)}d_k(i)\right]^2\right\} \tag{5}$$

estimates $\hat{A}_k^{(m)}$ and $\hat{\sigma}^2_{\xi_k^{(m)}}$ can be obtained via the corresponding sample averages over the packet. These estimates are used to compute the detector *a posteriori* probabilities $\Pr\left(d_k(i) = \pm 1 | y_k^{(m)}(i)\right)$ which are deinterleaved and input to the MAP decoder for the convolutional code. In what follows, we assume that the MAP decoder generates APPs $\Pr[d_k(i) = \pm 1]$, which are used to compute the input to the feedback filter $\bar{\mathbf{d}}^{(m)}$.

III. COMPUTATION OF DFD FILTERS

We first consider MMSE optimization of $\mathbf{F}$ and $\mathbf{B}$, followed by the LS adaptive estimation algorithm.

A. MMSE Optimization

For the P-DFD our objective is to select $\mathbf{F}$ and $\mathbf{B}$ to minimize the mean squared error (MSE)

$$\mathcal{E} = E\left\{\|\mathbf{d} - \mathbf{F}^\dagger\mathbf{r} + \mathbf{B}^\dagger\bar{\mathbf{d}}\|^2\right\} \tag{6}$$

where $\bar{\mathbf{d}}$ does not depend on k, and the dependence on iteration m and symbol i is not shown for convenience. With perfect feedback, i.e., $\bar{\mathbf{d}} = \mathbf{d}$, the optimal $\mathbf{F}$ and $\mathbf{B}$ are derived in [2] and [3]. In that case

$$\mathbf{F} = \mathbf{F}_{\text{lin}}(\mathbf{I} + \mathbf{B}) \tag{7}$$

where

$$\mathbf{F}_{\text{lin}} = \mathbf{R}^{-1}\mathbf{P} \tag{8}$$

is the *linear* MMSE filter and

$$\mathbf{R} = E[\mathbf{r}\mathbf{r}^\dagger] = \mathbf{P}\mathbf{P}^\dagger + \sigma^2\mathbf{I}. \tag{9}$$

Let

$$\epsilon = \mathbf{d} - (\mathbf{F}_{\text{lin}})^\dagger\mathbf{r} \tag{10}$$

denote the error corresponding to $\mathbf{F}_{\text{lin}}$, which has covariance matrix

$$\mathbf{Q} = \mathbf{I} - \mathbf{P}^\dagger\mathbf{R}^{-1}\mathbf{P}. \tag{11}$$

Let $\tilde{\mathbf{Q}}_k$ denote the kth column of $\mathbf{Q}$ with the kth element removed ($K-1$ elements), and let $\tilde{\mathbf{Q}}_{k,k}$ denote the $(K-1)\times(K-1)$ matrix, which is $\mathbf{Q}$ with both the kth row and kth column removed. Let the $(K-1)\times K$ matrix $\tilde{\mathbf{B}}$ denote $\mathbf{B}$ with the zeros along the diagonal removed. Then the nonzero entries of the kth column of the feedback filter $\mathbf{B}$ is given by

$$\tilde{\mathbf{B}}_k = \tilde{\mathbf{Q}}_{k,k}^{-1}\tilde{\mathbf{Q}}_k. \tag{12}$$

This filter minimizes the MSE $E\{\|\epsilon_k - \sum_{l\neq k}[\tilde{\mathbf{B}}]_{k,l}\epsilon_l\|^2\}$ over $\tilde{\mathbf{B}}_k$. For the ideal single-cell synchronous CDMA case considered, it can be shown that $\mathbf{F} = \kappa\mathbf{P}$, where κ is a constant, and $\mathbf{B} = \kappa(\mathbf{P}^\dagger\mathbf{P} - \mathbf{I})$. That is, the optimized P-DFD with perfect feedback reduces to the standard interference canceller. In the presence of other-cell interference, $\mathbf{F}$ becomes the linear MMSE filter given that only other-cell users are present, and $\mathbf{B}$ again cancels intracell interference with ideal feedback.

For the S-DFD, we optimize $\mathbf{F}$ and $\mathbf{B}$ column by column, i.e., for user k, the MSE is

$$\mathcal{E}_k = E\{|d_k - \bar{\mathbf{F}}_k^\dagger\bar{\mathbf{r}}|^2\} \tag{13}$$

where

$$\bar{\mathbf{F}}_k = \begin{bmatrix} \mathbf{F}_k \\ -\mathbf{B}_k \end{bmatrix} \quad \bar{\mathbf{r}} = \begin{bmatrix} \mathbf{r} \\ \mathbf{d} \end{bmatrix}. \tag{14}$$

With perfect feedback, the MMSE solution for $\bar{\mathbf{F}}_k$ is

$$\bar{\mathbf{F}}_k = \bar{\mathbf{R}}^{-1}\bar{\mathbf{v}}_k \tag{15}$$

where

$$\bar{\mathbf{R}} = \{E[\bar{\mathbf{r}}\bar{\mathbf{r}}^\dagger]\} = \begin{bmatrix} \mathbf{R} & \mathbf{P} \\ \mathbf{P}^\dagger & \mathbf{I}_K \end{bmatrix} \quad \bar{\mathbf{v}}_k = \begin{bmatrix} \mathbf{p}_k \\ \mathbf{e}_k \end{bmatrix}. \tag{16}$$

$\mathbf{I}_K$ is the $K \times K$ identity matrix, and $\mathbf{e}_k$ is the kth unit vector. It is straightforward to show that this solution is equivalent to (7) and (12) [3].

For the case of interest, where $\bar{\mathbf{d}} \neq \mathbf{d}$, the MMSE solution for $\mathbf{F}$ and $\mathbf{B}$ depends on the *joint (second-order) statistics* of $\bar{\mathbf{d}}$, $\mathbf{d}$, and $\mathbf{r}$. Computation of $\bar{\mathbf{d}}$ will be discussed shortly, but for now we assume that $\bar{\mathbf{d}}$ is a random variable and that correlations with $\mathbf{d}$ and $\mathbf{r}$ are well defined. For the numerical results in Section IV, $\bar{\mathbf{d}}$ are the soft decisions on the coded bits at the output of the MAP decoder. Let

$$\mathbf{R}_{xy} \triangleq E[\mathbf{x}\mathbf{y}^\dagger] \tag{17}$$

where $\mathbf{x}$ and $\mathbf{y}$ are random vectors. Minimizing $\mathcal{E}$ in (6) gives the MMSE feedforward filter

$$\mathbf{F} = \mathbf{R}_{rr}^{-1}(\mathbf{R}_{r\bar{d}}\mathbf{B} + \mathbf{R}_{rd}). \tag{18}$$

The feedback matrix is defined in terms of the matrices

$$\mathbf{M}_1 = \mathbf{R}_{\bar{d}\bar{d}} - \mathbf{R}_{\bar{d}r}\mathbf{R}_{rr}^{-1}\mathbf{R}_{r\bar{d}} \tag{19}$$

$$\mathbf{M}_2 = \mathbf{R}_{\bar{d}d} - \mathbf{R}_{\bar{d}r}\mathbf{R}_{rr}^{-1}\mathbf{R}_{rd}. \tag{20}$$

Let $\tilde{\mathbf{M}}_{1;k,k}$ denote $\mathbf{M}_1$ with the kth row and column removed (a $(K-1)\times(K-1)$ matrix), and let $\tilde{\mathbf{M}}_{2,k}$ denote the kth column of $\mathbf{M}_2$ with the kth component removed ($K-1$ elements). Then

$$\tilde{\mathbf{B}}_k = \tilde{\mathbf{M}}_{1;k,k}^{-1}\tilde{\mathbf{M}}_{2,k} \tag{21}$$

where $\tilde{\mathbf{B}}_k$ is defined as before.

For the S-DFD, we minimize $\mathcal{E}_k$ in (13), which gives (15) where

$$\bar{\mathbf{R}} = \{E[\bar{\mathbf{r}}\bar{\mathbf{r}}^\dagger]\} = \begin{bmatrix} \mathbf{R}_{rr} & \mathbf{R}_{r\bar{d}_k} \\ \mathbf{R}_{\bar{d}_k r} & \mathbf{R}_{\bar{d}_k\bar{d}_k} \end{bmatrix} \quad \bar{\mathbf{v}}_k = \begin{bmatrix} [\mathbf{R}_{\bar{d}_k r}]_k \\ \mathbf{R}_{\bar{d}_k\bar{d}_k;k} \end{bmatrix} \tag{22}$$

where $\mathbf{R}_{\bar{d}_k\bar{d}_k;k}$ denotes the kth column of $\mathbf{R}_{\bar{d}_k\bar{d}_k}$. Unlike the case with perfect feedback, in this case, the solutions corresponding to the P- and S-DFDs are generally different.

B. LS Optimization

Determining $\mathbf{R}_{r\bar{d}}$, $\mathbf{R}_{\bar{d}d}$, and $\mathbf{R}_{\bar{d}\bar{d}}$ in terms of $\mathbf{P}$, the noise statistics, the interleaver, and code properties appears to be difficult, which makes an assessment of true MMSE performance quite difficult. An alternative adaptive approach is to select $\mathbf{F}_k$ and $\mathbf{B}_k$ to minimize the LS cost function computed over all available L input vectors

$$\mathcal{E}_{LS,k} = \sum_{i=0}^{L} \left| d_k(i) - \mathbf{F}_k^\dagger \mathbf{r}(i) + \mathbf{B}_k^\dagger \bar{\mathbf{d}}_k(i) \right|^2. \qquad (23)$$

This gives the same expressions for $\mathbf{F}_k$ and $\mathbf{B}_k$ as given by (15) and (22) (or (18) and (21) for the P-DFD), where the covariance matrices are replaced by the corresponding sample covariance matrices [17]. That is, for sequences $\{\mathbf{x}(i)\}$ and $\{\mathbf{y}(i)\}$, $i = 0, \ldots, J$, we redefine

$$\mathbf{R}_{xy} \triangleq \sum_{i=0}^{J} \mathbf{x}(i)\mathbf{y}^\dagger(i). \qquad (24)$$

The preceding LS algorithm is decision-directed since it depends on the coded bit sequence $\{d(i)\}$. For training, the sequence of training bits would have to be coded and interleaved in order to measure $\mathbf{R}_{r\bar{d}_k}$, $\mathbf{R}_{\bar{d}_k d_k}$, and $\mathbf{R}_{\bar{d}_k \bar{d}_k}$. To avoid this complication, we can instead update $\mathbf{F}$ and $\mathbf{B}$ according to an adaptive LS algorithm based on perfect feedback during training [18] and subsequently switch to decision-directed mode during the remainder of the data packet. Either hard or soft decisions on $\{\mathbf{d}_k(i)\}$ can be used to update $\mathbf{F}_{\text{lin}}$ in (7). For the P-DFD the error covariance matrix at the output of $\mathbf{F}_{\text{lin}}$ is then estimated as

$$\hat{\mathbf{Q}} = \sum_{i=0}^{M} \epsilon(i)\epsilon^\dagger(i) \qquad (25)$$

and $\mathbf{B}$ is computed according to (12) where $\mathbf{Q}$ is replaced by $\hat{\mathbf{Q}}$. Hard or soft decisions can again be used to compute $\hat{\mathbf{Q}}$. This algorithm minimizes the LS cost function $\mathcal{E}_{LS}$ when $\bar{\mathbf{d}} = \mathbf{d}$. Simulation results indicate that using soft decisions gives better performance than using hard decisions. Furthermore, the soft decision-directed version of this simplified LS adaptive algorithm [using (25)] is observed to perform as well as the preceding "exact" soft decision-directed LS algorithm [using (11)]. The former algorithm was therefore simulated to generate the numerical comparisons in Section VI.

IV. Computation of Feedback Symbols

The MAP decoder computes $\Pr[d_k(i) = 1]$ for each $k = 1, \ldots, K$ and i. We wish to use this information to compute the symbols $\bar{\mathbf{d}}(i)$, which are fed back through $\mathbf{B}$. For fixed $\mathbf{F}$ and $\mathbf{B}$, we can choose $\bar{\mathbf{d}}_k$ to minimize the MSE given by (6), which gives

$$\bar{\mathbf{d}}_k = (\mathbf{B}\mathbf{B}^\dagger)^{-1}\mathbf{B}(\mathbf{F}^\dagger\mathbf{P} - \mathbf{I})E[\mathbf{d}_k] \qquad (26)$$

where $E[d_k] = 2\Pr[d_k = 1] - 1$ is the soft estimate of d_k at the output of the MAP decoder. If $\mathbf{P}$ is unknown (e.g., due to multipath), then $\mathbf{P}$ can be replaced by the sample covariance matrix $\mathbf{R}_{br}$. A simplification occurs when $\mathbf{B}$ is selected

to cancel the interference at the output of $\mathbf{F}$ given perfect feedback ($\bar{\mathbf{d}} = \mathbf{d}$). In that case it can be shown that $\mathbf{F}^\dagger\mathbf{P} - \mathbf{I} = \mathbf{B}^\dagger$ so that $\bar{\mathbf{d}} = E[\mathbf{d}]$. The numerical results in Section VI use this simplification, rather than using (26).

It has been pointed out in [19] that using APPs to compute $E[\mathbf{d}]$ in (26) gives a biased estimate of the transmitted symbols, which are fed back for interference cancellation. This bias can be eliminated by replacing the APPs with the extrinsic information computed by the decoder. That is, the extrinsic information is obtained by removing the *a priori* information from the APP. (Extrinsic information is typically used in decoding of turbo codes.) For the adaptive multiuser receivers considered here, numerical experiments have shown that replacing APPs with extrinsic information results in much slower convergence with increasing iterations (≈ 1 dB penalty after five iterations relative to feedback with APPs), although the packet error rate converges to a slightly better result (a fraction of a decibel improvement after 50 iterations).

V. Turbo Iterations

Given a sequence of received vectors $\mathbf{r}(1), \ldots, \mathbf{r}(L)$ and a training sequence $\mathbf{d}(1), \ldots, \mathbf{d}(t)$, the receiver operations for the IS-DFD follow.

For the first iteration $(m = 1)$, do for $k = 1, \ldots, K$:

1) Compute the LS filters $\mathbf{F}_k^{(1)}$ and $\mathbf{B}_k^{(1)}$ with the training data according to (15) and (22), where the sample covariance matrices are defined according to (24).
2) Compute $y_k^{(1)}(i)$, $i = 1, \ldots, L$, from (2).
3) Compute detector *a posteriori* probabilities for user k.
4) Deinterleave.
5) Decode and compute the sequence $\{\bar{d}_k^{(1)}(i)\}$, $i = 1, \ldots, L$, from MAP decoder outputs.

For each $m = 2, \ldots, M$ (number of iterations):

1) Reorder users. Do for $k = 1, \ldots, K$ (according to the new order):
   a) interleave MAP outputs;
   b) compute $\mathbf{F}_k^{(m)}$ and $\mathbf{B}_k^{(m)}$ from the received vectors, training sequence, and $\{\bar{\mathbf{d}}_k^{(m)}\}$;
   c) compute MAP inputs for user k and deinterleave;
   d) decode and compute the sequence $\{\bar{d}_k^{(m)}(i)\}$, or if $m = M$, output hard estimates of information bits $\{b_k(i)\}$.

For the first iteration the elements $k, \ldots, K$ of the vectors $\bar{\mathbf{d}}_k^{(1)}$ and $\mathbf{B}_k^{(1)}$ are zero, since only soft decisions for users $1, \ldots, k-1$ are available. For the mth iteration, $m > 1$, the elements $k, \ldots, K$ of $\bar{\mathbf{d}}_k^{(m)}$ are soft decisions from iteration $m - 1$, and, in general, only the kth entry of $\mathbf{B}_k^{(m)}$ is zero.

Numerical experiments show that the performance of the IS-DFD is sensitive to the criterion used to reorder the users. For the numerical results in Section VII, the users at the beginning of iteration m are ordered according to increasing values of

$$\mathcal{M}_k = \sum_{i=1}^{L} \left| \text{sgn}[d_k(i)] - \bar{d}_k(i) \right|^2 \qquad (27)$$

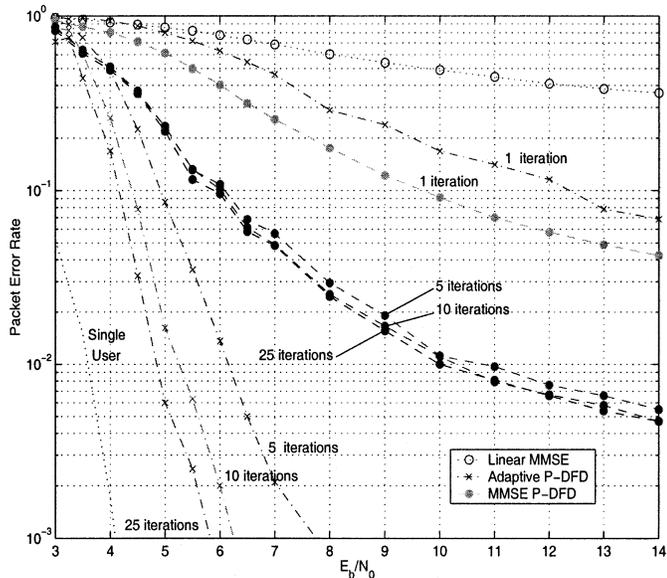where the sum is over all symbols in the packet.

Fig. 2. Receiver comparison for system with 12 users, bandwidth expansion of 16 and 200 training symbols.

TABLE I
CONVOLUTIONAL CODES FROM [25] USED TO GENERATE THE RESULTS IN SECTION VII. d_{free} IS THE FREE DISTANCE AND γ IS THE ASYMPTOTIC CODING GAIN. RATE 3/4 AND 7/8 CODES ARE PUNCTURED RATE 1/4 AND RATE 1/8 CODES, RESPECTIVELY

R_C	# states	d_{free}	γ [dB]
1/8	256	50	8.0
1/4	256	24	7.8
1/2	256	12	7.8
3/4	256	6	6.5
7/8	256	4	5.4

For the P-DFD, the users can be decoded in parallel so that there is no loop over the user index k when computing $\mathbf{F}$ and $\mathbf{B}$. For the first iteration, $\mathbf{F}^{(1)} = \mathbf{F}_{\mathrm{lin}}$ (linear LS filter) and $\mathbf{B}^{(1)} = \mathbf{0}$. The filters are recomputed at each iteration according to the decision-directed algorithm described in Section III.

VI. NUMERICAL COMPARISON

Fig. 2 shows a plot of *packet* error rate versus E_b/N_0 for the following receivers:
1) linear MMSE (no iterations);
2) MMSE P-DFD with coefficients calculated assuming *perfect* feedback ($\bar{\mathbf{d}} = \mathbf{d}$);
3) "approximate" LS P-DFD based on perfect feedback, but using soft decisions [(7), (12), and (25)].

A rate $R_C = 1/2$ code with the properties shown in Table I was used to generate these curves. The LS P-DFD based on both soft and hard decisions [(18) and (21)] was also simulated, but those results are not shown since with soft decisions the results are virtually identical to those of the "approximate" LS algorithm. Hard decisions cause a slight degradation in performance (< 0.2 dB). Plots are shown for 5, 10, and 25 iterations. All results are averaged over the users. Also shown is the single-user bound. The load $K/N = 3/4$, where $K = 12$ is the number of users, and $N = 16$ is the bandwidth expansion factor. Random Gaussian spreading sequences were used with unit energy. This

type of small system might correspond to a wireless local area network.

The data packets contain 500 information symbols, and for the results in Fig. 2, we assume a training sequence of 200 symbols. A packet is in error if it contains one or more bit errors. In addition to being an appropriate performance measure for a packet data network, the packet error rate (PER) is a better performance measure than bit-error rate (BER) for the random short code system considered. This is because a high cross correlation between two spreading sequences may cause a burst of errors in the associated packets, which raises the average BER even though the packet error rate may be low. Changing spreading codes from one packet to the next equalizes the packet error rate over the user population. Additional simulations with larger packet sizes show that these results essentially scale to larger packets so that the fractional overhead due to training can be decreased. (For packets up to four times longer, the performance degradation was approximately 0.5 dB, which is due to the PER performance measure.)

Because of the high load, the linear receiver has a PER near one for the range of E_b/N_0 shown. It is interesting that the optimized (MMSE) P-DFD assuming perfect feedback performs significantly worse than the adaptive P-DFD. This is because the adaptive P-DFD attempts to measure and exploit the joint statistics of the soft estimates with the transmitted symbols. We remark that a similar improvement in performance for MMSE filters relative to the ideal interference canceller is reported in [12]; however, the MMSE filters in [12] are conditioned on the soft decisions and, hence, must be recomputed for each symbol using knowledge of the spreading codes, relative amplitudes and phases, noise variance, and channels. Here, the filters are recomputed adaptively once per turbo iteration. The breakpoint shown in Fig. 2, i.e., where the curves start to fall, is close to the E_b/N_0 corresponding to large system sum capacity. Specifically, evaluating the large system sum capacity expression in [14] shows that the E_b/N_0 must be at least 3.2 dB to achieve a spectral efficiency of 0.75 bit per chip with load 0.75.

Fig. 3 shows PER versus number of training symbols. Curves are included for both the LS algorithms described earlier and stochastic gradient algorithms given by

$$\mathbf{F}_k(i+1) = \mathbf{F}_k(i) + |\bar{d}_k(i)|\mu\epsilon_k(i)\mathbf{r}(i)$$
$$\tilde{\mathbf{B}}_k(i+1) = \tilde{\mathbf{B}}_k(i) + |\bar{d}_k(i)|\mu\epsilon_k(i)\bar{\mathbf{d}}(i) \qquad (28)$$

where

$$\epsilon_k(i) = d_k(i) - \mathbf{F}_k^\dagger(i)\mathbf{r}(i) + \mathbf{B}^\dagger(i)\bar{\mathbf{d}}(i) \qquad (29)$$

and μ is the step-size. The additional factor $|\bar{d}_k(i)|$ is a reliability factor which weights the particular update. The decisions $d_k(i)$ can be either hard or soft. For the simulation results shown, the filters are reinitialized to zero after each turbo iteration, and the step size is constant over all iterations.

Fig. 3 shows that the LS algorithms can achieve a packet error rate of 5% with less than 100 training symbols, whereas the stochastic gradient algorithm requires at least four times as many training symbols. We also observe that there is a tradeoff between the amount of training data and number of iterations needed to achieve a target packet error rate. Specifically, with 200 training symbols four iterations are needed to achieve a PER of 5%, whereas with 100 training symbols, ten iterations are
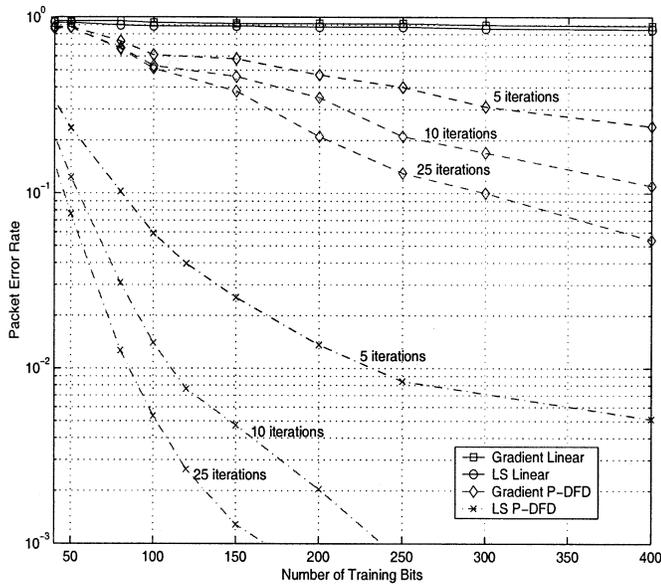
Fig. 3. PER versus training interval for adaptive receivers with 12 users and bandwidth expansion factor of 16 at $E_b/N_0 = 6$ dB.

Fig. 4. Receiver comparison with $K = 12$ users, $N = 16$, code rate $R_c = 1/2$ (eight chips percoded bit), 80 training symbols.

needed. Of course, increasing the number of iterations cannot compensate entirely for an inadequate training period.

Figs. 4 and 5 show the performance of the adaptive P- and IS-DFDs and a hybrid, which switches from the IS- to the P-DFD after the first iteration. Fig. 4 shows PER versus E_b/N_0 with 80 training symbols, and Fig. 5 shows PER versus number of iterations with $E_b/N_0 = 6$ dB and 80 training symbols. The results are averaged over the users. For this case, the hybrid algorithm performs best and offers a substantial gain with respect to the adaptive P-DFD. This performance difference diminishes as the training length increases. Fig. 5 shows that with limited training the IS- and hybrid DFDs require substantially fewer iterations to reach the same target error rate than the P-DFD. Alternatively, for a fixed number of iterations and target PER, the hybrid DFD requires significantly less training than the IS- and P-DFDs.

VII. EFFECT OF CODE RATE

We now study the effect of code rate on performance. It has been observed in [14] and [20]–[24] that for the linear MMSE receiver, the optimal code rate generally increases with load and SNR. For adaptive receivers, the code rate also influences the amount of training required. Namely, a smaller code rate reduces the number of filter coefficients to estimate, which reduces the required training overhead.

Fig. 6 shows plots of PER versus E_b/N_0 for different code rates. In this case, $K = 20$ and $N = 16$, corresponding to a load of 1.25. The large load was chosen to illustrate the substantial gain provided by an increase in code rate. Specifically, the rate 1/2 code requires at least 5 dB more E_b/N_0 than the rate 3/4 code to achieve a PER equal to 0.01. The convolutional codes considered are shown in Table I. These have been selected from [25] on the basis of similar complexity (constraint lengths). A larger code rate allows for more degrees of freedom for interference suppression at the first stage. This effect is more
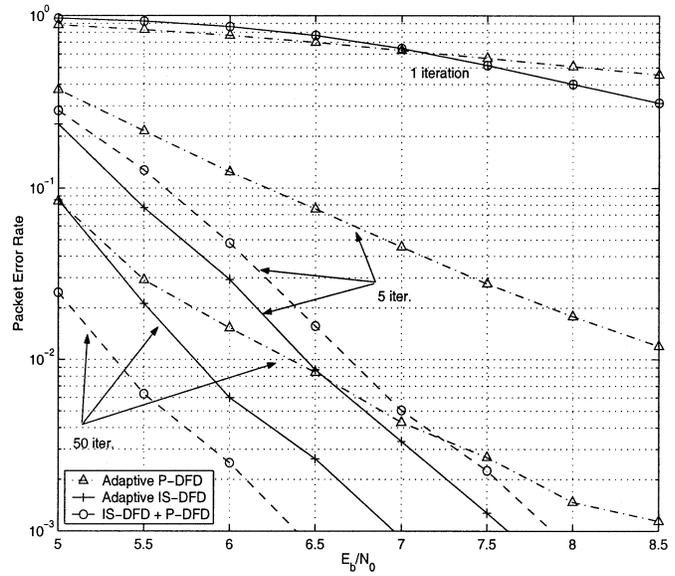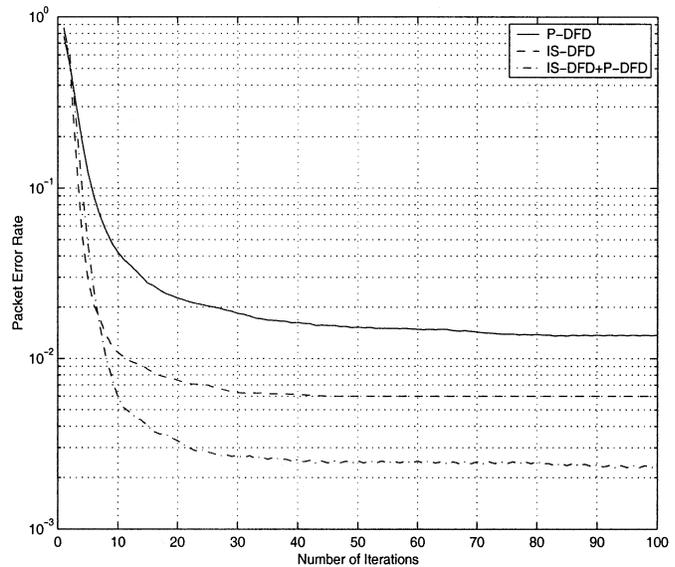

Fig. 5. PER versus turbo iterations for the P-, IS-, and hybrid DFDs with $E_b/N_0 = 6$ dB and 80 training symbols. Remaining parameters are the same as in Fig. 4.

pronounced for the iterative receivers than for linear receivers. Namely, if the code rate is too small, then an iterative receiver has insufficient degrees of freedom to produce soft decisions at the output of the first stage, which are sufficiently reliable to improve the detection at succeeding stages. Increasing the code rate may improve the soft decisions produced by the initial stage enough to achieve a relatively low PER. In contrast to the previous results for smaller load 3/4, here there is a substantial gap between the multi-user PERs and single-user performance (approximately 6 dB with the rate 7/8 code). The breakpoint shown in Fig. 6 is also further away from the E_b/N_0, corresponding to the large system sum capacity. Namely, 4.5 dB is needed to achieve a spectral efficiency of 1.25 chips/bit with load 1.25 [14].
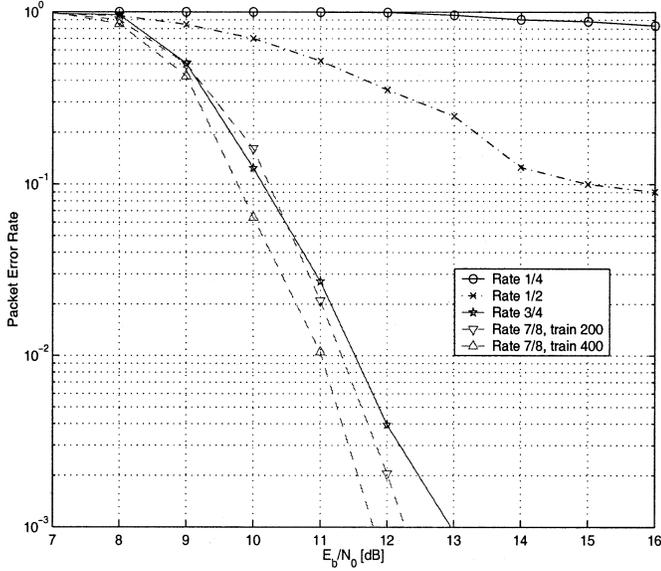
Fig. 6. Performance with different code rates, 20 users, total bandwidth expansion 16 (50 iterations, 200 training symbols).

Fig. 7. Training performance for different code rates, 12 users, bandwidth expansion 16 (25 iterations).

Further numerical experiments show that at the lower load of 0.75, the rate 1/2 code generally performs the best out of the codes in Table I. This is consistent with the large system analysis performance presented in [26].

The effect of code rate on training is illustrated by Fig. 7, which shows PER versus training symbols with the same system parameters used to generate Fig. 3. A similar set of codes to that shown in Table I was used, with all codes having 64 states, and asymptotic coding gains between 5.4 and 6.5 dB. Training symbols are *uncoded* in all cases. These results illustrate that the code rate affects both the length of training required to achieve a target PER and the error floor with unlimited training. Additional numerical examples show that longer training intervals and additional iterations are more beneficial with high loads or low E_b/N_0. For example, a similar experiment with $K/N = 1.25$ and $E_b/N_0 = 11$ dB shows that optimal performance requires nearly twice as much training as for the case shown in Fig. 7.

VIII. REDUCED-RANK ESTIMATION

A property of the LS estimation algorithms previously presented is that the amount of training generally increases proportionally with the filter length. Consequently, for a fixed packet length, a large training overhead may be needed with large filter lengths. One method for reducing the amount of training overhead is reduced-rank estimation of the feedforward and feedback filters [27, Sec. 8.4]. Here, we focus on the class of adaptive reduced-rank algorithms introduced in [16], which are motivated by the multistage Wiener filter [28]. These algorithms can provide a significant reduction in training when used with adaptive linear filters [16].

A reduced-rank filter first projects the received signal onto a lower dimensional subspace. Both the filtering and the filter estimation take place within this subspace. Let $\mathbf{S}_D$ be the $N' \times D$
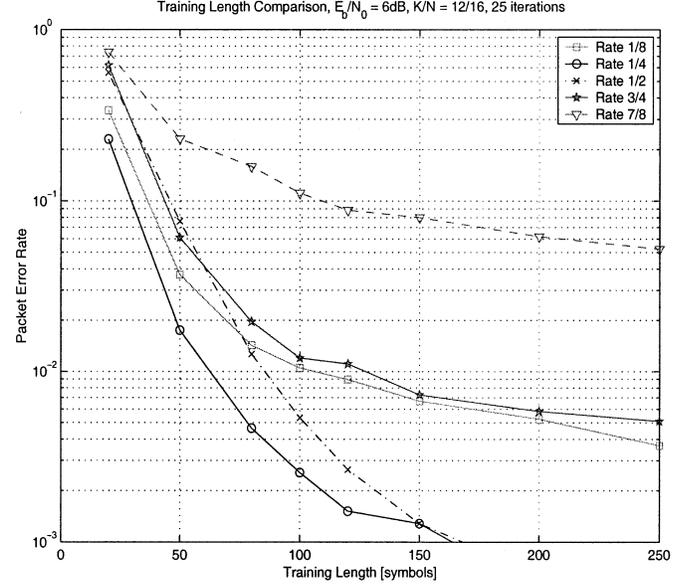
matrix with column vectors which are a basis for a D-dimensional subspace, where $D < N'$. The projected received vector corresponding to symbol i is then given by

$$\tilde{\mathbf{r}}(i) = \mathbf{S}_D^\dagger \mathbf{r}(i). \tag{30}$$

The sequence of projected received vectors $\{\tilde{\mathbf{r}}(i)\}$ is the input to the filter $\tilde{\mathbf{c}}$ ($D \times 1$ vector), which has output

$$z(i) = \tilde{\mathbf{c}}^\dagger \tilde{\mathbf{r}}(i). \tag{31}$$

The vector $\tilde{\mathbf{c}}(i)$, which minimizes the LS cost function $\sum_{i=1}^{M} [|b_1(i) - \tilde{\mathbf{c}}^\dagger(i)\tilde{\mathbf{r}}(i)|^2]$, is

$$\tilde{\mathbf{c}} = (\mathbf{S}_D^\dagger \mathbf{R}_{rr} \mathbf{S}_D)^{-1}(\mathbf{S}_D^\dagger \mathbf{R}_{br}) \tag{32}$$

where $\mathbf{R}_{rr} = \sum_{i=1}^{M} \mathbf{r}(i)\mathbf{r}^\dagger(i)$ ($N' \times N'$ matrix) and $\mathbf{R}_{br} = \sum_{i=1}^{M} b_1^*(i)\mathbf{r}(i)$ ($N' \times 1$ vector). Here, we focus on the Powers of $\mathbf{R}$ (PoR) algorithm, presented in [16], for which

$$\mathbf{S}_D = [\,\mathbf{R}_{br} \quad \mathbf{R}_{rr}\mathbf{R}_{br} \quad \mathbf{R}_{rr}^2\mathbf{R}_{br} \quad \ldots \quad \mathbf{R}_{rr}^{D-1}\mathbf{R}_{br}\,]. \tag{33}$$

For the adaptive iterative P-DFD, we use the decomposition of the feedforward filter, given by (7), and compute reduced-rank estimates for both the linear filter $\mathbf{F}_{\text{lin}}$ and the error estimation filter $\mathbf{I} + \mathbf{B}$. That is, for each user k, we estimate a linear reduced-rank filter $\mathbf{F}_k$ according to (32), and from (12) the reduced-rank error estimation filter is given by

$$\mathbf{B}_k = (\mathbf{S}_{Q,D}^H \tilde{\hat{\mathbf{Q}}}_{k,k}\mathbf{S}_{Q,D})^{-1}(\mathbf{S}_{Q,D}^H \tilde{\hat{\mathbf{Q}}}_k) \tag{34}$$

where $\tilde{\hat{\mathbf{Q}}}_k$ is the kth column of $\hat{\mathbf{Q}}$ excluding the kth component, $\tilde{\hat{\mathbf{Q}}}_{k,k}$ is $\hat{\mathbf{Q}}$ excluding the kth row and column, $\hat{\mathbf{Q}}$ is given by (25), and

$$\mathbf{S}_{Q,D} = \left[\, \tilde{\hat{\mathbf{Q}}}_k \quad \tilde{\hat{\mathbf{Q}}}_{k,k}\tilde{\hat{\mathbf{Q}}}_k \quad \tilde{\hat{\mathbf{Q}}}_{k,k}^2\tilde{\hat{\mathbf{Q}}}_k \quad \ldots \quad \tilde{\hat{\mathbf{Q}}}_{k,k}^{D-1} \quad \tilde{\hat{\mathbf{Q}}}_k \,\right]. \tag{35}$$

Figs. 8 and 9 compare the performance of full- and reduced-rank iterative DFDs for a system with $(K, N) = (48, 64)$ and 120 training symbols. Fig. 8 shows PER versus E_b/N_0 with five and 50 iterations, and Fig. 9 shows PER versus number of iterations with $E_b/N_0 = 6$ dB. For the reduced-rank
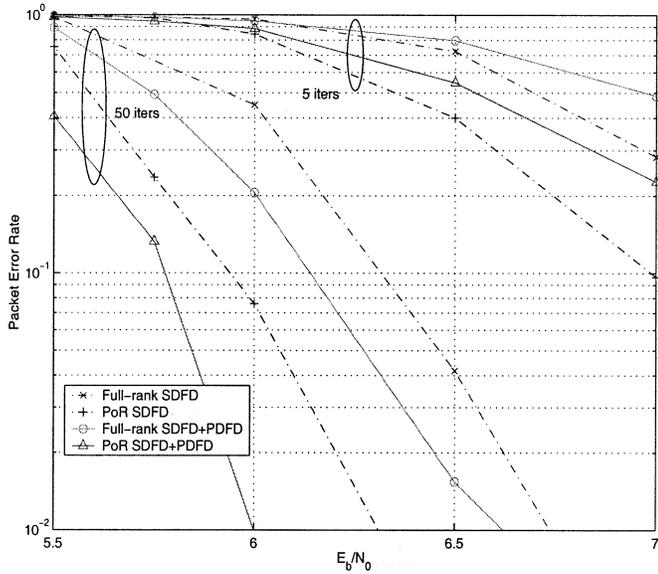
Fig. 8.   Comparison of PER versus E_b/N_0 for full- and reduced-rank DFDs with $K = 48$, $N = 64$, and 120 training symbols.
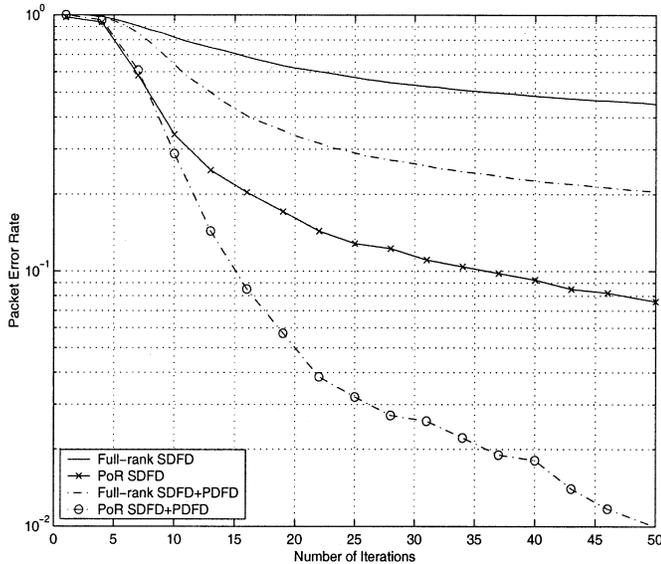

Fig. 9.   Comparison of PER versus iterations for full- and reduced-rank DFDs with $K = 48$, $N = 64$, 120 training symbols, and $E_b/N_0 = 6$ dB.

results, the feedforward and feedback ranks are optimized for each user. Specifically, for the P-DFD, the two ranks for user k are chosen to minimize

$$\mathcal{M}_k = \sum_{i=1}^{L} |y_k(i) - \text{sgn}(y_k)|^2 \qquad (36)$$

where $y_k(i)$ is the P-DFD output for user k at time i [see Fig. 1(b)]. For the IS-DFD, the two ranks optimized at a particular iteration for user k are used at the next iteration.

The results show that the reduced-rank algorithms offer a significant performance gain relative to the full-rank algorithms for the parameters selected. In general, for a fixed training length this gain increases with the size of the filters. For a smaller system size, e.g., $N = 16$, reduced-rank estimation does not offer a significant performance advantage. For fixed filter lengths, the performance gain decreases as the training

length increases, since in both cases the filters converge to the MMSE filters with perfect feedback.

IX. EXTENSIONS

Here, we briefly indicate how the adaptive iterative receivers presented can be extended to asynchronous CDMA with multipath. We first remark that these receivers can be directly applied to a synchronous CDMA system with multipath. In that case, the matrix $\mathbf{P}$ can be interpreted as the *received* signature during the observation window, i.e., the transmitted signature convolved with the channel response. For low delay spreads, the multipath is coherently combined by the MMSE estimation algorithm, as in a linear MMSE receiver, assuming the channels are stationary during the observation window. With large delay spreads, the associated intersymbol interference appears as unknown interference, which may be partially suppressed by the feedforward filter.

For a quasisynchronous chip-asynchronous system in which the users are offset by small delays (less than a few chips), chip oversampling can compensate for the random timing offsets. For example, with n samples per chip, both the received signatures and feedforward filters have dimension nN'. This increase in number of filter coefficients may increase the amount of training overhead needed. With larger relative delays among users, it becomes necessary to expand the observation window, as in a linear MMSE receiver [18]. Of course, the increase in filter length again increases the required training overhead.

Finally, we mention that although this work was motivated by the application to CDMA, the model also applies to a narrow-band system with multiple transmitter and receiver antennas [29]. Hence, the adaptive approach presented here can also applied to space–time iterative detection.

X. CONCLUSION

Adaptive multiuser decision feedback receivers with iterative soft MAP decoding have been presented for synchronous DS-CDMA. Estimation of the filter coefficients is performed once per iteration and relies upon training sequences, which are simultaneously transmitted by the users to be demodulated. The receiver complexity is relatively low. Specifically, the P-DFD filters for each user are computed by solving a set of N' linear equations for the forward linear filter and a set of $K - 1$ linear equations for the backward filter. The receiver is appropriate for the reverse link of a cellular system since the adaptive feedforward filter suppresses interference from all users which are not jointly decoded (e.g., from other cells).

The performance of the adaptive receivers has been studied by computer simulation and shows that for a small system (e.g., $N = 16$), the adaptive receiver can perform substantially better than a fixed DFD optimized with known signatures, assuming perfect feedback. We also find that with limited training or iterations, successive demodulation can significantly improve upon the performance of the iterative P-DFD. For moderate to large filter lengths, reduced-rank estimation can further reduce the required training overhead. Combining these methods with other enhancements, such as optimizing the tradeoff between extrinsic information and APPs, and filtering decision statistics over multiple iterations, as proposed in [30], may yield further improvements.

REFERENCES

[1] 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. 1995.

[2] G. Woodward, R. Ratasuk, M. L. Honig, and P. Rapajic, "Multistage decision-feedback detection for DS-CDMA," in *Conf. Rec. Int. Conf. Communications (ICC)*, Vancouver, BC, Canada, June 1999.

[3] ——, "Minimum mean squared error multiuser decision-feedback detectors for DS-CDMA," *IEEE Trans. Commun.*, 2002.

[4] M. K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the gaussian multiple-access channel," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Mar. 1998, pp. 1405–1409.

[5] R. Müller, "Multiuser receivers for randomly spread signals: Fundamental limits with and without decision-feedback," *IEEE Trans. Inform. Theory*, vol. 47, pp. 268–283, Jan. 2001.

[6] P. Rapajic, M. L. Honig, and G. Woodward, "Multiuser decision feedback detection: Performance bounds and adaptive algorithms," in *Proc. Int. Symp. Information Theory*, Cambridge, MA, Aug. 1998.

[7] R. Ratasuk, G. Woodward, and M. L. Honig, "Adaptive multiuser parallel-decision-feedback demodulation," in *Proc. Allerton Conf. Communication, Control, Computing*, Monticello, IL, Oct. 1999.

[8] P. D. Alexander, A. J. Grant, and M. C. Reed, "Iterative detection in code-division multiple-access with error control coding," *Eur. Trans. Telecommun.*, vol. 9, pp. 419–425, Sept.-Oct. 1998.

[9] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. Commun.*, vol. 46, pp. 1693–1699, Dec. 1998.

[10] P. D. Alexander, M. C. Reed, J. Asenstorfer, and C. B. Schlegel, "Iterative multiuser interference reduction: Turbo CDMA," *IEEE Trans. Commun.*, vol. 47, July 1999.

[11] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1061, July 1999.

[12] H. E. Galmal and E. Geroniotis, "Iterative multiuser detection for coded CDMA signals in AWGN and fading channels," *IEEE J. Select. Areas Commun.*, vol. 47, pp. 30–41, Jan. 2000.

[13] Q. Zhao and D. K. Borah, "Adaptive iterative CDMA multiuser detection in unknown multipath channels," in *Conf. Rec. Int. Conf. Communications (ICC) 2002*, vol. 1, New York, Apr. 2002, pp. 361–365.

[14] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.

[15] M. Kopbayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1450–1460, Aug. 2001.

[16] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.*, vol. 50, pp. 986–994, June 2002.

[17] M. L. Honig and D. G. Messerschmitt, *Adaptive Filters: Structures, Algorithms, and Applications*. Boston, Ma: Kluwer, 1985.

[18] M. L. Honig and R. Ratasuk, "Adaptive multiuser decision-feedback demodulation for GSM," in *Proc. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 1998.

[19] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1772–1793, July 2002.

[20] J. Foerster and L. B. Milstein, "Coding for a coherent DS-CDMA system employing an MMSE receiver in a Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 48, pp. 1909–1918, June 2000.

[21] W. G. Phoel, M. L. Honig, and B. R. Vojcic, "Coded performance of MMSE receivers for DS-CDMA in the presence of fading," in *Proc. CISS*, Mar. 1999.

[22] W. G. Phoel and M. L. Honig, "Performance of coded DS-CDMA with pilot-assisted channel estimation and linear interference suppression," *IEEE Trans. Commun.*, vol. 50, pp. 822–832, May 2002.

[23] W. Xiao and M. L. Honig, "Forward link performance of satellite CDMA with linear interference suppression and one-step power control," IEEE Trans. Wireless Commun., 2002, to be published.

[24] E. Biglieri, G. Caire, and G. Taricco, "CDMA system design through asymptotic analysis," *IEEE Trans. Commun.*, vol. 48, pp. 1882–1896, Nov. 2000.

[25] I. E. Bocharova and B. D. Kudryashov, "Rational rate punctured convolutional codes for soft-decision viterbi decoding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1305–1313, July 1997.

[26] M. L. Honig and R. Ratasuk, "Large system performance of iterative multiuser decision-feedback detectors," *IEEE Trans. Commun.*, Aug. 2001.

[27] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison Wesley, 1993.

[28] J. Goldstein, I. Reed, and L. Scharf, "A multistage representation of the wiener filter based on orthogonal projections," *IEEE Trans. Inform. Theory*, vol. 44, Nov. 1998.

[29] A. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–83, Nov. 1997.

[30] S. Marinkovic, B. Vucetic, and A. Ushirokawa, "Space-time iterative and multistage receiver structures for CDMA mobile communication systems," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1594–1603, Aug. 2001.

Michael L. Honig (S'80–M'81–SM'92–F'97) received the B.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1977 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1978 and 1981, respectively.

He joined Bell Laboratories, Holmdel, NJ, where he worked on local area networks and voiceband data transmission. In 1983, he joined the Systems Principles Research Division, Bellcore, where he worked on digital subscriber lines and wireless communications. Since 1994, he has been with Northwestern University, Evanston, IL, where he is a Professor in the Electrical and Computer Engineering Department. He was a Visiting Lecturer at Princeton University, Princeton, NJ, in 1993, a Visiting Mackay Professor at the University of California, Berkeley, in 2000, and a Visiting Scholar at the University of Sydney, Sydney, Australia, during 2001.

He has served as an Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY and the IEEE TRANSACTIONS ON COMMUNICATIONS, and was a Guest Editor for the *European Transactions on Telecommunications* and *Wireless Personal Communications*. He has also served as a member of the Digital Signal Processing Technical Committee for the IEEE Signal Processing Society and as a member of the Board of Governors for the Information Theory Society. He is the Corecipient of the 2002 IEEE Communications Society and Information Theory Society Joint Paper Award.

Graeme K. Woodward (S'94–M'99) received the B.Sc., B.E. (with honors and University Medal), and Ph.D. degrees from the University of Sydney, Sydney, Australia, in 1990, 1992, and 1999, respectively.

From 1991 to 1994, he worked at Alcatel Australia in digital switching systems. During his Ph.D. studies he was a Visiting Researcher at Yokohama National University, Yokohama, Japan, during 1996 and at Northwestern University, Evanston, IL, during 1998. From 1999 to 2000, he worked with Southern-Poro Communications, Sydney, in channel modeling, 3G cellular systems, and multiuser detection for CDMA, In 2001, he joined Bell Laboratories Research, Lucent Technologies, Sydney, Australia, researching 3G systems, multi-antenna systems, multiuser detection, interference suppression and adaptive digital filter algorithms for equalization.

Yakun Sun (S'02) received the B.S. degree from Tsinghua University, Beijing, China, in 1999 and the M.S. degree in electrical engineering from Northwestern University, Evanston, IL, in 2001. He is currently working toward the Ph.D. degree at Northwestern University.

His research interests include wireless communications and signal processing.