Decision Feedback Sequence Estimation for Unwhitened ISI Channels with Applications to Multiuser Detection

Abdulrauf Hafeez and Wayne E. Stark, *Fellow, IEEE*

Abstract—Decision feedback sequence estimation (DFSE), which is a reduced-complexity alternative to maximum likelihood sequence estimation (MLSE), can be used effectively for equalization of intersymbol interference (ISI) as well as for multiuser detection. The algorithm performs very well for whitened (minimum-phase) channels. For nonminimum-phase channels, however, the algorithm is not very effective. Moreover, DFSE requires a noise-whitening filter, which may not be feasible to compute for time-varying channels such as a multiuser direct-sequence code division multiple access (DS-CDMA) channel. Noise-whitening is also cumbersome for applications that involve bidirectional equalization such as the global system for mobile communication (GSM) system. In such conditions, it is desirable to use the Ungerboeck formulation for sequence estimation, which operates directly on the discrete-time unwhitened statistic obtained from conventional matched filtering. Unfortunately, DFSE based on matched filter statistics is severely limited by untreated interference components. In this paper, we identify the anticausal interference components, using an error probability analysis. This leads us to a modified unwhitened decision feedback sequence estimator (MUDFSE) in which the components are canceled, using tentative decisions. We obtain approximate error probability bounds for the proposed algorithm. Performance results indicate that the modified algorithm, used on unwhitened channels with relatively small channel correlations, provides similar performance/complexity tradeoffs as the DFSE used on the corresponding whitened *minimum-phase* channels. The algorithm is especially attractive for multiuser detection for asynchronous DS-CDMA channels with long spreading codes, where it can achieve near-MLSE performance with exponentially lower complexity.

Index Terms—Code division multiple access (CDMA), decision feedback equalizers, equalizers, intersymbol interference (ISI), multiuser channels, multiuser detection, sequence estimation, Viterbi detection.

I. INTRODUCTION

THERE are two main approaches to maximum likelihood sequence estimation (MLSE) for intersymbol interference (ISI) channels with additive noise. Forney's approach [1] consists of using the Viterbi algorithm on the sampled output of whitened matched filters, while Ungerboeck's formulation [2] operates directly on conventional matched filter output samples

without the need for noise whitening. Due to the exponential complexity of these methods, several low-complexity suboptimal schemes have been developed. A promising reduced-complexity alternative to MLSE using Forney's approach is decision feedback sequence estimation (DFSE) [3], [4]. This scheme provides an excellent tradeoff between performance and complexity by reducing the memory order of the Viterbi algorithm and employing conditional decision feedback to cancel the tail of the ISI.

A drawback to DFSE is that its performance is sensitive to channel phase. While the best performance is obtained for minimum-phase channels, the performance may not be adequate for nonminimum phase channels. In the global system for mobile communication (GSM) system, training sequences are inserted in the middle of the burst to aid channel estimation. Equalization over the data field is then performed bidirectionally, starting from the training sequence [5]. If the channel response is minimum-phase in one direction then it is maximum-phase in the other.[1] DFSE performs poorly in maximum-phase conditions and is thus unsuitable for equalization in this case.

Besides channels with ISI, DFSE can be employed for multiuser detection. In direct-sequence code division multiple access (DS-CDMA) systems the use of long spreading codes and the arrival and departure of users gives rise to a time-varying multiuser interference channel. The sequence of statistics obtained at the output of a bank of matched filters is affected by noncausal multiple access interference and correlated noise [6]. In order to use DFSE, the statistic has to be filtered to make the noise white and the interference causal. Noise whitening filters have been proposed for a time-varying environment [7]. The operation, however, is computationally expensive.

Ungerboeck's formulation for sequence estimation is better suited in conditions such as those mentioned above. A reduced-state Ungerboeck-type variant to DFSE, which we call unwhitened decision feedback sequence estimation (UDFSE), was derived in [8][2] and [5]. It operates on discrete-time unwhitened statistics obtained from conventional matched filtering. UDFSE is thus insensitive to channel phase and does not require a noise-whitening filter. The scheme, however,

[1] Minimum-phase response can be obtained for both directions by means of appropriate all-pass filters.

[2] The scheme is referred to as delayed DFSE with *a standard matched filter* in [8].

does not perform as well as DFSE on most channels. This is because, unlike whitened statistics which depend on the past input symbols only, unwhitened statistics depend on both the past and future input symbols. The decision rule of DFSE selects the survivor path with the best accumulated (Euclidean distance) metric leading to a reduced state. The same decision rule when used with an Ungerboeck metric (as in UDFSE) fails to account for interference from the future inputs not represented in the reduced state (but which would have been represented in the full-blown state). As a result, the decisions obtained in UDFSE are affected by untreated interference, even in the absence of feedback errors. The M-algorithm [9] used on unwhitened statistics suffers from a similar problem: survivor paths are chosen on the basis of an accumulated metric which does not reflect the effects of some anticausal interference components. Thus, the performance of the unwhitened M-algorithm is rather poor, as was noted by Wei *et al.* for CDMA systems [10].

In the case of UDFSE, the decision rule can be modified to take into account the effect of the interfering anticausal components. The decision rule in the modified UDFSE (MUDFSE) algorithm selects survivor paths based on the accumulated Ungerboeck metric plus a *bias* term that precancels the effect of the interfering anticausal inputs on the choice of survivor paths, using tentative decisions. The *bias* term is determined by examining the pairwise error probability of the UDFSE algorithm. It can be computed by using conventional decisions based on matched filter outputs or by using decisions obtained at the output of the preceding stage in a multistage scheme. A reduced computation form of the modified detector was proposed in [11].

The rest of the paper is organized as follows. We present the system model in Section II. The two classic approaches to MLSE are reviewed in Section III. In Section IV we describe various methods for DFSE in detail. We show that UDFSE performance is limited by untreated ISI, which leads us to the MUDFSE algorithm. In Section V, we obtain approximate bounds on the symbol error probability of UDFSE and the modified algorithm. In Section VI, we illustrate how the bounds can be evaluated using error-state diagrams. Performance results are presented in Section VII, where the BER performance of the various schemes is compared for single user ISI and multiuser DS-CDMA channels using analysis and simulation.

II. SYSTEM MODEL

Consider the transmission of linearly modulated digital data over a (time-varying) time-dispersive additive white Gaussian noise (AWGN) channel. Assume that the receiver has perfect knowledge of the carrier phase, the symbol timing and the impulse response of the channel. After coherent down conversion, the receiver employs a filter matched to the cascade of the transmit pulse-shaping filter and the channel impulse response (assumed to be fixed for the duration of the transmit filter) and samples the output at symbol spaced intervals. The sequence of sampled matched filter outputs is known to be a sufficient statistic for estimating the transmitted sequence.

Thus, an equivalent discrete-time channel model is obtained. The matched filter output at time n is given by

$$y_n = \sum_{i=-L}^{L} s_{n,i} b_{n-i} + z_n \tag{1}$$

where $\{b_n\}_{n=1}^{N}$ is the transmitted data sequence assumed to be independent and identically distributed (i.i.d.) ($b_n \in A_M$, an M-ary alphabet), $\{z_n\}$ is a complex Gaussian noise process with mean 0 and covariance $E[z_n z_{n-i}^*] = N_0 s_{n,i}$, and $s_{n,i}$ is the sampled channel autocorrelation function, given by

$$s_{n,i} = \begin{cases} \int h^*(t - nT; t) h(t - nT + iT; t)\, dt, & |i| \leq L \\ 0, & |i| > L \end{cases} \tag{2}$$

where $h(\tau; t)$ is the overall channel impulse response with (finite) span L in symbol intervals, where τ represents delay and t represents time variation. We will refer to the above model as the unwhitened model.

Let $s_n(D) = \sum_i s_{n,i} D^i$, then an equivalent discrete-time white Gaussian noise model is obtained by noting that

$$s_n(D) = f_n^*\left(D^{-1}\right) f_n(D)$$

where $f_n(D) = \sum_{i=0}^{L} f_{n,i} D^i$ has all its roots outside the unit circle (minimum-phase channel). The statistic obtained by filtering $\{y_n\}$ with the anticausal noise-whitening filter $1/f_n^*(D^{-1})$ (assuming that it exists) is given by

$$y_n' = \sum_{i=0}^{L} f_{n,i} b_{n-i} + z_n' \tag{3}$$

where $\{z_n'\}$ is a proper[3] complex white Gaussian noise process with mean zero and variance N_0.

III. MLSE

An MLSE based on the unwhitened model determines, as the most likely sequence transmitted, the sequence $\{b\}_{n=1}^{N}$ that maximizes the metric [2], [12]

$$U(\underline{y}_N, \underline{b}_N) = \sum_{n=1}^{N} 2\,\mathrm{Re}\{b_n^* y_n\} - \sum_{i=1}^{N} \sum_{k=1}^{N} b_i^* s_{i,k} b_k \tag{4}$$

where $\underline{x}_n = x_1, x_2, \ldots, x_n$. Due to the symmetry in the sampled channel correlations: $s_{n,i} = s_{n-i,-i}^*$, the above metric can be computed recursively as

$$U(\underline{y}_n, \underline{b}_n) = U(\underline{y}_{n-1}, \underline{b}_{n-1}) + V(y_n, b_n, \sigma_n) \tag{5}$$

where $V(\cdot)$ is known as the branch metric and is given by

$$V(y_n, b_n, \sigma_n)$$
$$= \mathrm{Re}\left\{ b_n^* \left(2y_n - s_{n,0} b_n - 2 \sum_{i=1}^{L} s_{n,i} b_{n-i} \right) \right\} \tag{6}$$

and $\sigma_n = b_{n-1}, b_{n-2}, \ldots, b_{n-L}$ represents the state at time n.

[3] $E[\mathrm{Re}\{z_n'\}\,\mathrm{Im}\{z_n'\}] = 0$.

The above is a generalized form of the Ungerboeck formulation (where $s_{n,i} = s_i$, i.e., the channel is fixed) independently obtained in [12]. An optimum multiuser detector proposed by Verdu [6] for asynchronous DS-CDMA systems can be obtained from this formulation.

The maximum likelihood sequence can be determined equivalently by applying the Viterbi algorithm to the whitened statistic $\{y'_n\}$, as shown in [1], using the metric

$$U'(\underline{y}'_N, \underline{b}_N) = \sum_{n=1}^{N} \left[2 \operatorname{Re} \left(y'_n \sum_{i=0}^{L} f^*_{n,i} b^*_{n-i} \right) - \left| \sum_{i=0}^{L} f_{n,i} b_{n-i} \right|^2 \right] \quad (7)$$

which can be computed recursively as

$$U'(\underline{y}'_n, \underline{b}_n) = U'(\underline{y}'_{n-1}, \underline{b}_{n-1}) + V'(y'_n, b_n, \sigma_n) \quad (8)$$

where

$$V'(y'_n, b_n, \sigma_n) = 2 \operatorname{Re} \left(y'_n \sum_{i=0}^{L} f^*_{n,i} b^*_{n-i} \right) - \left| \sum_{i=0}^{L} f_{n,i} b_{n-i} \right|^2. \quad (9)$$

The two algorithms require the same order of computational complexity and storage proportional to M^L.

IV. DFSE

A reduced-complexity suboptimal alternative to MLSE using the Viterbi algorithm is DFSE [3], [4]. A parameter called memory order J is chosen arbitrarily smaller than channel memory L, and the trellis is collapsed into M^J states corresponding to the J most recent hypothesized symbols. Since the reduced state falls short in providing all the information needed to compute branch metrics, the algorithm uses the path with the best accumulated metric leading to each state to extract the rest of the information. Note that this algorithm operates on whitened statistics $\{y'_n\}$ which depend on past and present inputs only [cf. (3)].

A. UDFSE

The UDFSE algorithm [8] follows the Ungerboeck formulation for sequence estimation. The algorithm is given by the recursion

$$m(\beta_{n+1}) = \max_{b_{n-J}} [m(\beta_n) + W(y_n, b_n, \beta_n)] \quad (10)$$

where $\beta_n = b_{n-1}, b_{n-2}, \ldots, b_{n-J}$ represents the reduced state at time n, $m(\beta_n)$ is the accumulated metric of the path associated with the state β_n, and $W(y_n, b_n, \beta_n)$ is the branch metric given by

$$W(y_n, b_n, \beta_n) = \operatorname{Re} \left\{ b^*_n \left(2y_n - s_{n,0} b_n - 2 \sum_{i=1}^{J} s_{n,i} b_{n-i} \right. \right.$$
$$\left. \left. - 2 \sum_{i=J+1}^{L} s_{n,i} \hat{b}_{n-i}(\beta_n) \right) \right\}. \quad (11)$$

In (11), $\{\hat{b}_{n-i}(\beta_n)\}$ are tentative conditional decisions on inputs more than J samples in the past, obtained from the path with the best accumulated metric, leading to the state β_n, as

$$\hat{b}_{n-J}(\beta_{n+1}) = \arg \max_{b_{n-J}} [m(\beta_n) + W(y_n, b_n, \beta_n)]. \quad (12)$$

Although both full-state formulations in Section III yield the MLSE, it has been noted that the reduced-state derivation UDFSE does not perform nearly as well as its whitened channel counterpart (DFSE) on most channels [10], [11]. The problem arises from the nature of unwhitened statistics $\{y_n\}$ which depend on both the past and future input symbols. Note that the UDFSE algorithm, when deciding b_{n-J} (conditionally) in the nth step (12), ignores interference from inputs $b_{n+1}, b_{n+2}, \ldots, b_{n-J+L}$ which directly affect output sample y_{n-J} (corresponding to b_{n-J}). Thus, the decision rule of (12), which is based on the knowledge of only J inputs in the future, is inherently near sighted. Unlike DFSE, the decisions obtained in UDFSE are affected by untreated interference components, even in the absence of feedback errors. This observation is quantified by the following analysis.

Let $\{\dot{b}_n\}$ be the transmitted sequence of symbols and $\{\dot{\beta}_n\}$ be the sequence of states in the path of $\{\dot{b}_n\}$ in the reduced trellis. Let $\{\bar{b}_n\}$ be a hypothetical sequence of symbols and $\{\bar{\beta}_n\}$ be the corresponding sequence of states in the reduced trellis that diverges from the correct sequence of states at time n_1 and remerges with it at a later time n_2, i.e.,

$$\bar{\beta}_n = \dot{\beta}_n \quad \text{for } n = n_1, n_2 \quad \text{and}$$
$$\bar{\beta}_n \neq \dot{\beta}_n \quad \text{for } n_1 < n < n_2. \quad (13)$$

An error event occurs at time n_2 if the algorithm picks $\{\bar{b}_n\}$ as the correct sequence over $\{\dot{b}_n\}$.

Proposition 4.1: The occurrence of the error event depends on the value of inputs $\dot{b}_{n_2}, \dot{b}_{n_2+1}, \ldots, \dot{b}_{n_2-1+L-J}$ in the absence of error propagation (from any previous error events).

Proof: The error event occurs if the accumulated metric on the incorrect path is greater than that on the correct path at the point where the two paths merge, i.e.,

$$\sum_{n=n_1}^{n_2-1} W\left(y_n, \bar{b}_n, \bar{\beta}_n\right) > \sum_{n=n_1}^{n_2-1} W\left(y_n, \dot{b}_n, \dot{\beta}_n\right). \quad (14)$$

Assume

$$\bar{b}_n = \dot{b}_n \quad \text{for } n_1 - L \leq n \leq n_1 - J - 1 \quad (15)$$

i.e., there are no errors $L - J$ steps prior to the error event. Substituting branch metrics from (11) and noting that $\{\hat{b}_{n-i}(\bar{\beta}_n)\}$ and $\{\hat{b}_{n-i}(\dot{\beta}_n)\}$ are decisions taken from the paths corresponding to the sequences of states $\{\bar{\beta}_n\}$ and $\{\dot{\beta}_n\}$, respectively, (14) can be written as

$$2 \operatorname{Re} \left\{ \underline{\bar{b}}^H \underline{y} \right\} - \underline{\bar{b}}^H S \underline{\bar{b}} > 2 \operatorname{Re} \left\{ \underline{\dot{b}}^H \underline{y} \right\} - \underline{\dot{b}}^H S \underline{\dot{b}} \quad (16)$$

where $\underline{x} = [x_{n_1}, x_{n_1+1}, \ldots, x_{n_2-1}]^T$ and S is a banded Hermitian matrix given by (17) at the bottom of the following page.

Let $\underline{e} = \frac{1}{2}(\underline{\bar{b}} - \underline{\dot{b}})$ be the error sequence. Using (1) and (15) and noting that $e_n = 0$ for $n_2 - J \leq n \leq n_2 - 1$ as $\bar{\beta}_{n_2} = \dot{\beta}_{n_2}$ it can be shown that the above inequality holds if

$$2\,\text{Re}\,\{\underline{e}^H \underline{y}\} > \underline{e}^H S \underline{e} - \underline{e}'^H S' \underline{\dot{b}}' \tag{18}$$

where $\underline{\dot{b}}' = [\dot{b}_{n_2}, \dot{b}_{n_2+1}, \ldots, \dot{b}_{n_2-1+L-J}]^T$, $\underline{e}' = [e_{n_2-L}, e_{n_2-L+1}, \ldots, e_{n_2-1-J}]^T$ and S' is a lower triangular matrix given by (19) at the bottom of this page.

Conditioned on the error sequence $\underline{e}$ the left-hand side of (18) is a Gaussian random variable with mean zero and variance that depends on the error sequence. While the first term on the right-hand side of (18) depends on the error sequence only, the second term in addition depends on the inputs $\dot{b}_{n_2}, \dot{b}_{n_2+1}, \cdots, \dot{b}_{n_2-1+L-J}$. $\qquad\square$

The above analysis suggests that if the term which depends on the $L - J$ future inputs is absorbed in the decision rule of UDFSE, the occurrence of the error event in the modified detector will depend on the error sequence only. This, of course, requires knowledge of some future inputs. In practice, tentative decisions can be used instead as described in the following section.

B. The MUDFSE Algorithm

The MUDFSE algorithm can be outlined as follows. Path metrics are computed as in UDFSE using (10). Conditional decisions are made (and the corresponding survivor paths are chosen) using the modified rule

$$\hat{b}_{n-J}(\beta_{n+1}) = \arg\max_{b_{n-J}}[m(\beta_n) + W(y_n, b_n, \beta_n) - \text{bias}(\beta_n)] \tag{20}$$

where

$$\text{bias}(\beta_n) = \sum_{j=n-L+1}^{n-J} \text{Re}\left\{2b_j^* \sum_{i=n-j+1}^{L} s_{j,-i}\tilde{b}_{j+i}\right\} \tag{21}$$

$$\tilde{b}_n = \text{sign}(y_n). \tag{22}$$

The algorithm is delayed $L - J$ samples as y_{n+L-J} is needed in the nth step. Note that the bias in (22) depends on b_{n-J} (the bit falling out of the state) and $b_{n-J-1}, \ldots, b_{n-L+1}$ for which conditional decisions taken from the path history of state β_n are used. Note that the bias is used only for survivor path selection and does not constitute the accumulated path metric. The bias can be simplified to include only the leading term which depends on b_{n-J}, as follows:

$$\text{bias}(\beta_n) \approx \text{bias}(b_{n-J})$$
$$= \text{Re}\left\{2b_{n-J}^* \sum_{i=J+1}^{L} s_{n-J,-i}\tilde{b}_{n-J+i}\right\}. \tag{23}$$

The approximate bias is independent of the state. It does not contribute significantly to the computational load and storage requirement of the algorithm which is on the order of $(G - J)M^J$ (the same as DFSE) where G is the decision lag. This reduced computation form of MUDFSE was first proposed in [11].

C. Multistage MUDFSE

The MUDFSE algorithm can be run in a multistage configuration where decisions obtained at the output of the first stage are fed back to compute the bias in the second stage and so on, i.e.,

$$(\tilde{b}_n)_1 = \text{sign}(y_n) \quad \text{and} \quad (\tilde{b}_n)_i = (\bar{b}_n)_{i-1}, \qquad i > 1$$

where $\{(\bar{b}_n)_i\}$ are decisions obtained from the ith stage at lag G_i ($G_i \gg L$). Note that the decisions $\{(\bar{b}_n)_1\}$ are likely to be much more reliable than $\{(\tilde{b}_n)_1\}$. The complexity and delay of a Q-stage scheme is given by $\sum_{i=1}^{Q}(G_i - J_i)M^{J_i}$ and $\sum_{i=1}^{Q} G_i + L - J_1$, respectively.

V. ERROR PERFORMANCE

Assuming stationarity (i.e., $s_{n,i} = s_i$) and the absence of decision errors prior to the start of an error event, the beginning

$$S = \begin{bmatrix} s_{n_1,0} & \cdots & s_{n_1,-L} & 0 & 0 & \cdots & 0 \\ s_{n_1+1,1} & s_{n_1+1,0} & \cdots & s_{n_1+1,-L} & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & & & \vdots \\ s_{n_1+L,L} & \cdots & s_{n_1+L,0} & \cdots & s_{n_1+L,-L} & \ddots & 0 \\ 0 & \ddots & & & & \ddots & \vdots \\ \vdots & & & \ddots & & \ddots & 0 \\ 0 & \ddots & s_{n_2-1-L,L} & \cdots & s_{n_2-1-L,0} & \cdots & s_{n_2-1-L,-L} \\ \vdots & & \ddots & & \ddots & & \vdots \\ 0 & \cdots & 0 & s_{n_2-2,L} & \cdots & s_{n_2-2,0} & s_{n_2-2,-1} \\ 0 & \cdots & 0 & 0 & s_{n_2-1,L} & \cdots & s_{n_2-1,0} \end{bmatrix} \tag{17}$$

$$S' = \begin{bmatrix} s_{n_2-L,-L} & 0 & \cdots & 0 \\ s_{n_2-L+1,-L+1} & s_{n_2-L+1,-L} & 0 & 0 \\ \vdots & & & \vdots \\ \vdots & & \ddots & 0 \\ s_{n_2-1-J,-J-1} & \cdots & s_{n_2-1-J,-L+1} & s_{n_2-1-J,-L} \end{bmatrix}. \tag{19}$$

of the error event can be aligned with time $n = 1$ without loss of generality. In the following sections, we find the pairwise error probability and an upper bound on the symbol error probability for UDFSE and MUDFSE.

A. Genie-Aided MUDFSE (GA-MUDFSE)

An error event ε of length n_ε symbols occurs in MUDFSE if

$$\sum_{n=1}^{n_\varepsilon+J} W(y_n, \bar{b}_n, \bar{\beta}_n) - \text{bias}(\bar{\beta}_{n_\varepsilon+J})$$
$$> \sum_{n=1}^{n_\varepsilon+J} W(y_n, \dot{b}_n, \dot{\beta}_n) - \text{bias}(\dot{\beta}_{n_\varepsilon+J}). \quad (24)$$

Assume that perfect information is provided by a genie on the future inputs needed to compute the bias in MUDFSE. Then, using (21) with $\tilde{b}_n$ replaced by $\dot{b}_n$ in the above inequality, it follows that the error event occurs in the GA-MUDFSE if

$$\sum_{n=1}^{n_\varepsilon} \text{Re}\{e_n^* y_n\} > \delta^2(\boldsymbol{e}) \quad (25)$$

where $\delta(\boldsymbol{e})$ is called the distance[4] of the error sequence $\boldsymbol{e} = \{e_1, e_2, \ldots, e_{n_\varepsilon}\}$ and is given by

$$\delta^2(\boldsymbol{e}) = \sum_{n=1}^{n_\varepsilon} \sum_{k=1}^{n_\varepsilon} e_n^* s_{n-k} e_k. \quad (26)$$

The left-hand side of (25) is a Gaussian random variable with mean zero and variance $\delta^2(\boldsymbol{e})N_0$. Thus, the probability that the error event ε occurs is given by

$$\Pr(\varepsilon) = Q\left(\frac{\delta(\boldsymbol{e})}{\sqrt{N_0}}\right). \quad (27)$$

Note that this expression for the pairwise event error probability of GA-MUDFSE is identical to the expression derived in [2] for MLSE. There is, however, an important distinction in the definition of error events for MUDFSE and MLSE. A valid error sequence for a Jth order MUDFSE ($1 \le J \le L$) can have no more than $J - 1$ consecutive zeros according to the definition in (13), whereas $J = L$ for MLSE. When $J = 0$, all error sequences have length $n_\varepsilon = 1$ symbol. In this case $\delta^2(\boldsymbol{e}) = |e|^2 s_0$. In other words, a zeroth order GA-MUDFSE approaches the performance of the ISI-free channel. Notice that the analysis does not consider decision error propagation.

B. MUDFSE

In the absence of ideal bias, the occurrence of an error event in MUDFSE depends, in addition to the distance of the error sequence, on the residual interference arising from tentative decision errors on the $L - J$ inputs $(\dot{b}_{n_\varepsilon+J+1}, \ldots, \dot{b}_{n_\varepsilon+L})$ following the error event. Note that the tentative decision errors following an error event in the main detector are correlated with the main decision errors due to memory in the channel. In order to simplify analysis, we assume that tentative and main decision errors occur independently and the tentative

[4]Note that this definition of distance is different from the one given in [2].

decision error process is stationary. Thus, we get the pairwise error probability for MUDFSE using (24), (18), and (21) as

$$\Pr(\varepsilon) \approx \sum_{\boldsymbol{t}} Q\left(\frac{\delta(\boldsymbol{e}) + \gamma(\boldsymbol{e}, \boldsymbol{t})}{\sqrt{N_0}}\right) P_{\boldsymbol{t}} \quad (28)$$

where $P_{\boldsymbol{t}}$ is the probability that the sequence of tentative decision errors $\boldsymbol{t} = \{t_{n_\varepsilon+J+1}, \ldots, t_{n_\varepsilon+L}\}$ (where $t_n = \frac{1}{2}(\bar{b}_n - \dot{b}_n)$) follows the error event ε and $\gamma(\boldsymbol{e}, \boldsymbol{t})$ is the residual interference, given by

$$\gamma(\boldsymbol{e}, \boldsymbol{t}) = \sum_{n=\max(n_\varepsilon+J+1-L, 1)}^{n_\varepsilon}$$
$$\cdot \text{Re}\left\{e_n^*\left(2 \sum_{k=n_\varepsilon+J+1}^{n+L} s_{n-k} t_k\right)\right\} \Big/ \delta(\boldsymbol{e}). \quad (29)$$

The residual interference $\gamma(\boldsymbol{e}, \boldsymbol{t})$ can be viewed as the projection of the tentative decision error vector $\boldsymbol{t}$ onto the main decision error vector $\boldsymbol{e}$ as determined by the channel correlation spectrum $\boldsymbol{s}$.

Note that, unlike DFSE where error distances diminish, in general, when memory order J is decreased as compared to channel memory L, the distances for MUDFSE remain the same as in the case $J = L$ (MLSE). The variance of the residual interference $\gamma(\boldsymbol{e}, \boldsymbol{t})$, however, increases from zero as the memory order J is made smaller than L in MUDFSE. This sharp contrast between the behavior of MUDFSE and DFSE is consistent with the fact that the operation of matched filtering collects all the energy w_n of the pulse transmitted at time n into the corresponding output sample y_n $(s_{n,0} = w_n)$ [2], while the operation of noise whitening results in the scattering of some of this energy into L subsequent output samples $(\Sigma_{i=0}^{L} |f_{n,i}|^2 = w_n)$.

When $J = 0, n_\varepsilon = 1, \delta^2(\boldsymbol{e}) = |e_1|^2 s_0$ and $\gamma(\boldsymbol{e}, \boldsymbol{t})$ is given by

$$\gamma(\boldsymbol{e}, \boldsymbol{t}) = \text{Re}\left\{e_1^*\left(2 \sum_{k=2}^{L+1} s_{1-k} t_k\right)\right\} \Big/ |e_1|\sqrt{s_0} \quad (30)$$

i.e., MUDFSE reduces to a zero-forcing decision-feedback detector for unwhitened channels (where feed-forward filtering is replaced by tentative decision feedback).

C. UDFSE

Using (18), the pairwise error probability in the case of UDFSE can be written as

$$\Pr(\varepsilon) = \sum_{\boldsymbol{b}} Q\left(\frac{\delta(\boldsymbol{e}) + \mu(\boldsymbol{e}, \boldsymbol{b})}{\sqrt{N_0}}\right) P_{\boldsymbol{b}} \quad (31)$$

where $P_{\boldsymbol{b}}$ is the probability that the sequence of inputs $\boldsymbol{b} = \{b_{n_\varepsilon+J+1}, \ldots, b_{n_\varepsilon+L}\}$ follows the error event ε ($P_{\boldsymbol{b}} = M^{J-L}$ for i.i.d. equiprobable inputs) and $\mu(\boldsymbol{e}, \boldsymbol{b})$ is the untreated interference given by

$$\mu(\boldsymbol{e}, \boldsymbol{b}) = \sum_{n=\max(n_\varepsilon+J+1-L, 1)}^{n_\varepsilon}$$
$$\cdot \text{Re}\left\{e_n^*\left(-\sum_{k=n_\varepsilon+J+1}^{n+L} s_{n-k} b_k\right)\right\} \Big/ \delta(\boldsymbol{e}). \quad (32)$$

$$a_0 = \tfrac{1}{2}D^{s_0}$$
$$a_1 = \tfrac{1}{2}D^{(s_0 + 2s_1)}$$
$$a_2 = \tfrac{1}{2}D^{(s_0 - 2s_1)}$$
$$a_3 = \tfrac{1}{2}D^{(s_0 + 2s_1 + 2s_2)}$$
$$a_4 = \tfrac{1}{2}D^{(s_0 - 2s_1 - 2s_2)}$$
$$a_5 = \tfrac{1}{2}D^{(s_0 - 2s_1 + 2s_2)}$$
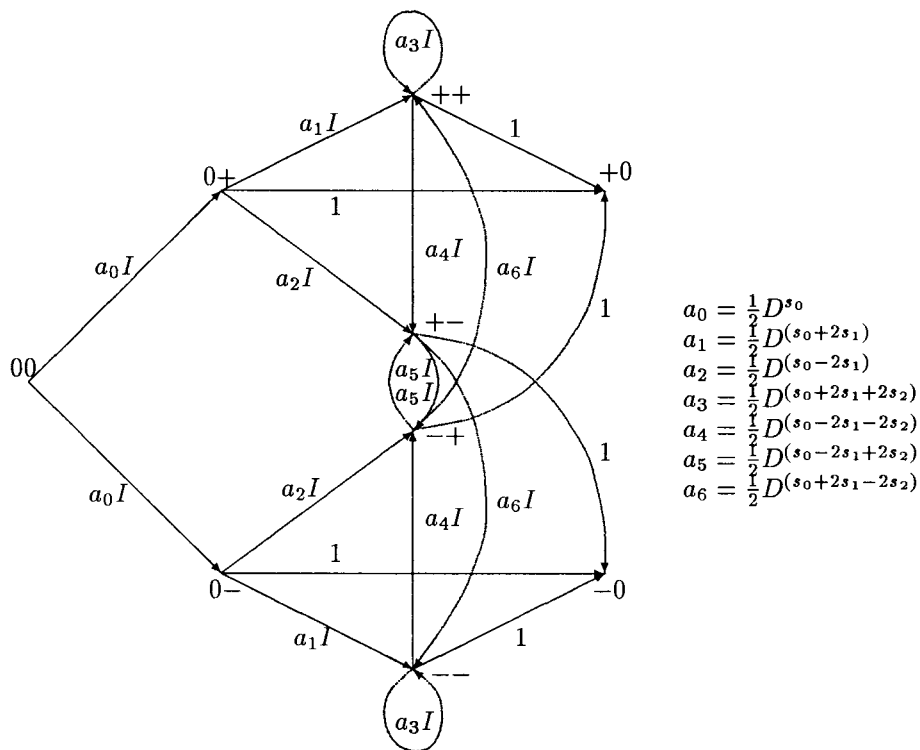$$a_6 = \tfrac{1}{2}D^{(s_0 + 2s_1 - 2s_2)}$$

Fig. 1.   Error-state diagram for $L = 2, J = 1$.

Note that (31) provides an exact expression for the pairwise error probability of UDFSE unlike (28) for MUDFSE.

D. Symbol Error Probability

Assuming that all input sequences are equally likely, the symbol error probability for UDFSE, GA-MUDFSE, and MUDFSE can be union-bounded as [3]

$$P_s \leq \sum_{\varepsilon} w(e) \left(\prod_{n=1}^{n_\varepsilon} \frac{M - |e_n|}{M} \right) \Pr(\varepsilon) \qquad (33)$$

where $w(e)$ is the number of symbol errors entailed by the error sequence e and $\prod_{n=1}^{n_\varepsilon} (M - |e_n|)/M$ is the number of input sequences that can have e as the error sequence where the symbol alphabet is given as $A_M = \{\pm 1, \pm 3, \ldots, \pm(M-1)\}$ for M even. In the case of binary phase-shift keying (BPSK) modulation, (33) simplifies to

$$P_s \leq \sum_{\varepsilon} \frac{w(e)}{2^{w(e)}} \Pr(\varepsilon). \qquad (34)$$

Note that the above bounds on the probability of symbol error are obtained assuming no decision error propagation (a separation of more than $L - J$ symbols between error events). The effect of error propagation, however, is usually small at medium to high signal-to-noise ratio (SNR).

VI. BOUND EVALUATION

To evaluate the symbol error probability for MUDFSE, we make use of the error-state diagram as in [13]. The error-state diagram used for determining the generating function of channel codes for MLSE [13, pp. 283] has to be modified

in the case of MUDFSE so that the error sequences satisfy (13). The modified diagram is shown in Fig. 1 for the case of binary signaling over an AWGN channel with memory $L = 2$ and memory order $J = 1$ chosen for MUDFSE. It has dimensionality 3^L. The nodes are labeled with the error states which are ternary L-tuples with components that take values in $\{0, +1, -1\}$. The transitions are labeled with the branch generating function $\mathrm{Re}\,\{e_n^*(s_0 e_n + 2 \Sigma_{i=1}^{L} s_i e_{n-i})\}$ as the exponent of dummy variable D, and the number of symbol errors entailed by the transition as the exponent of variable I. A factor of $1/2$ is used to account for the weighing factor in (34) if the transition involves an error. The transitions which lead back to the $0+$ and $0-$ states (except from the 00 state) have been eliminated because error sequences which contain a zero in the middle are not allowed in the case $J = 1$, according to the definition in (13).

Instead of enumerating all allowable error paths through the channel code trellis, we are interested in enumerating the paths that terminate with a $+1$ and -1 separately because the residual interference given by (29) depends on the tail of the error path $e' = \{e_{\max(n_\varepsilon + J + 1 - L,1)}, \ldots, e_{n_\varepsilon}\}$. Thus, in general, we seek $2 \times 3^{L-J-1}$ generating functions ($J < L$) corresponding to the paths which terminate with a given tail (such that $e_{n_\varepsilon} \neq 0$).

Let $T_j(D, I)$ be the generating function for the error paths which terminate in the tail e'_j ($j = 1, \ldots, 2 \times 3^{L-J-1}$) found by solving the state equations simultaneously. Each generating function can be series expanded as follows:

$$\frac{\partial}{\partial I} T_j(D, I) \bigg|_{I=1} = \sum_k B_{j,k} D^{\delta_{j,k}^2} \qquad (35)$$

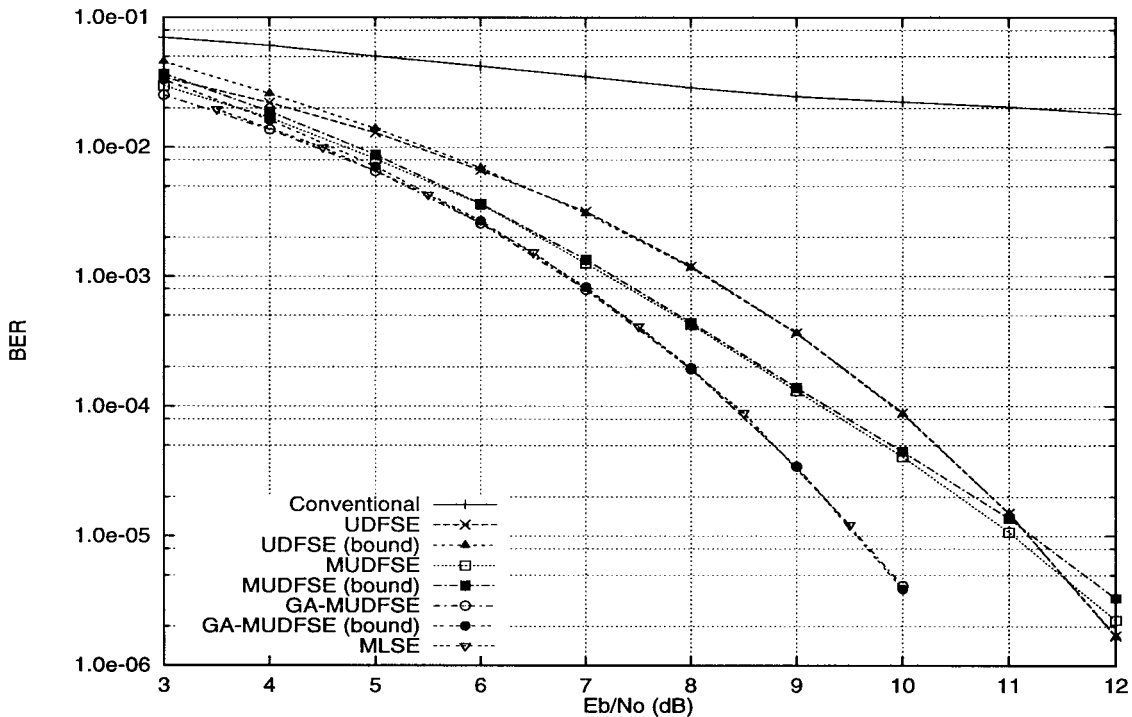where $B_{j,k}$ is the number of error paths (weighted by the

Fig. 2. BER performance of various detection schemes on channel 1.

number of symbol errors on the path per the number of the corresponding input sequences) with Euclidean distance $\delta_{j,k}$ that terminate in the tail e'_j. Then, the symbol error probability for MUDFSE can be computed as

$$P_s \leq \sum_{j=1}^{2 \times 3^{L-J-1}} \sum_{k} \sum_{\boldsymbol{t}} B_{j,k} D^{\delta_{j,k} + \gamma(\boldsymbol{e}'_j, \delta_{j,k}, \boldsymbol{t})}$$
$$\cdot P_{\boldsymbol{t}} \Big|_{D^x = Q(x/\sqrt{N_0})} \qquad (36)$$

where we use the fact that the residual interference $\gamma(\boldsymbol{e}, \boldsymbol{t})$ depends on the error sequence $\boldsymbol{e}$ only through the tail of the error sequence $\boldsymbol{e}'$ and the distance of the error sequence $\delta(\boldsymbol{e})$.

Note from Fig. 1 that the error-state pairs that are negative of each other are indistinguishable on the basis of branch values. Thus, they can be combined as in [13]. It follows that the number of paths that terminate in the tail $\boldsymbol{e}'$ is the same as the number of paths that terminate in the tail $-\boldsymbol{e}'$. Moreover, since $\gamma(\boldsymbol{e}'_j, \delta_{j,k}, \boldsymbol{t}) = \gamma(-\boldsymbol{e}'_j, \delta_{j,k}, -\boldsymbol{t})$, (36) simplifies to

$$P_s \leq 2 \sum_{j=1}^{3^{L-J-1}} \sum_{k} \sum_{\boldsymbol{t}} B_{j,k} D^{\delta_{j,k} + \gamma(\boldsymbol{e}'_j, \delta_{j,k}, \boldsymbol{t})}$$
$$\cdot P_{\boldsymbol{t}} \Big|_{D^x = Q(x/\sqrt{N_0})}. \qquad (37)$$

In general, the reduced error-state diagram for binary signaling comprises $(3^L - 1)/2$ nonzero nodes with 3^{L-J-1} terminating nodes.

In order to compute (37), we assume that the sequence of tentative decision errors $\boldsymbol{t}$ is an i.i.d. sequence which is independent of the sequence of main decision errors $\boldsymbol{e}$ and has

distribution

$$t_n = \begin{cases} 0 & 1-p \\ +1 & \frac{1}{2} p \\ -1 & \frac{1}{2} p \end{cases} \qquad (38)$$

where p is the probability of tentative decision error which, in the case of a single-stage MUDFSE algorithm, is the symbol error probability of a conventional matched-filter detector.

As the noise is correlated, tentative decision errors are correlated with each other as well as with main decision errors. Our assumptions are thus optimistic because errors in the tentative detector will tend to occur in bursts, inducing errors in the main detector. Nevertheless, independence can be assumed in case noise correlations are small.

VII. PERFORMANCE RESULTS

In this section, we compare the performance of the various detection algorithms discussed in this paper via simulation and analysis using several example channels. First, we consider binary signaling over static time-dispersive AWGN channels. The receiver is assumed to have perfect knowledge of the symbol timing and the impulse response of the channel. We run each simulation for a count of 1000 errors.

Channel 1 is given by $\boldsymbol{s'} = (0.2, -0.25, 1.0, -0.25, 0.2)$ where $\boldsymbol{s'}$ represents the normalized channel correlations. The ISI channel has memory $L = 2$. Fig. 2 shows the BER performance of the various detection schemes on this channel. The memory order J is chosen to be one for UDFSE and MUDFSE. The bounds in Fig. 2 are computed using (34) by averaging over all error sequences $\boldsymbol{e}$ with squared distance $\delta^2(\boldsymbol{e}) \leq 10.0$. For MUDFSE, the probability of tentative decision error [p in (38)] is found from simulation. The
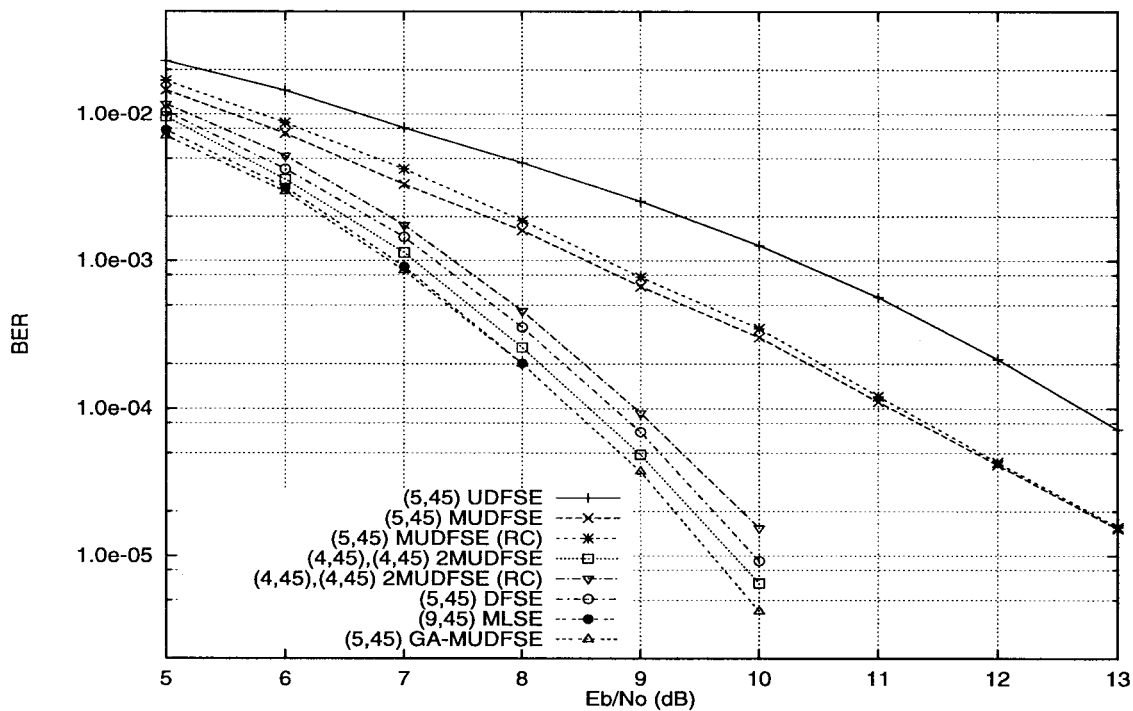
Fig. 3.   BER performance of various detection schemes on channel 2.

bounds are tight, especially for moderate to high SNR as can be expected of union bounds (the curves are virtually indistinguishable from those for simulation in the case of UDFSE and GA-MUDFSE at high SNR). Note that the bounds are approximate as they do not account for decision error propagation. However, the effect of error propagation is usually small in schemes that employ conditional decision feedback at medium to high SNR, as noted in [3]. The semi-analytic bound for MUDFSE, shown in Fig. 2, seems to diverge somewhat from the simulation curve, although it can be expected to cross over the simulation curve at high SNR due to dependence between the tentative and main decisions.

For channel 1, MUDFSE provides some gain over UDFSE in the low to medium SNR range. But at high SNR, its performance is worse than that of UDFSE. The performance differential depends on how effectively the anticausal interference components affecting UDFSE (Section IV-A) are canceled in the modified algorithm by means of tentative decisions based on matched filter outputs. GA-MUDFSE which cancels the anticausal interference ideally, performs slightly better than MLSE, as shown in Fig. 2. This is due to the reasons discussed in Section V-A.

Fig. 3 shows the BER performance (simulated) of various detection schemes over an ISI channel (2) with memory $L = 9$. Channel 2 is given by $f = (0.861, 0.258, -0.100, -0.274, 0.130, 0.100, -0.038, 0.112, -0.114, -0.228)$. The channel is minimum-phase and is arbitrarily chosen. The various schemes are indexed with the memory order and the decision lag (J, G). Except for DFSE, all algorithms operate on symbol-spaced samples obtained from conventional matched filtering. Ideal noise whitening is assumed for DFSE. All schemes shown in Fig. 3, except for MLSE, require similar overall complexity and storage. Fig. 3 shows that

(5, 45) MUDFSE gains $1.0 - 1.5$ dB over (5, 45) UDFSE in the SNR range shown. The two-stage scheme [(4, 45), (4, 45)] 2MUDFSE closely approaches MLSE performance and obtains a gain of 4 dB over UDFSE at an error rate of 10^{-4}. The single-stage and two-stage MUDFSE (reduced computation) schemes lose less than 0.5 dB due to bias approximation (23).

The performance of (5, 45) DFSE is also close to MLSE for channel 2, as shown in Fig. 3. We have neglected the effects of nonideal noise-whitening (due to fixed delay constraint) on the performance of DFSE. The delay incurred from anticausal (noise-whitening) filtering needed in DFSE can be compared to the delay of a multistage MUDFSE scheme. As channel 2 is minimum phase, the performance achieved by DFSE is the best possible for any channel phase. The performance of UDFSE and MUDFSE, however, is independent of channel phase.

Fig. 4 shows the BER performance (simulated) over another ISI channel (3) with memory $L = 9$. Channel 3 is given by $f = (0.707, -0.153, 0.468, -0.306, 0.285, -0.104, 0.244, 0.068, -0.070, 0.053)$. Channel 3 is more correlated than channel 2. Note that the (5, 45) MUDFSE schemes do a little worse than (5, 45) UDFSE. This is partly because conventional hard decisions are unreliable (BER $\approx 1/5$). The two-stage MUDFSE schemes, therefore, obtain better performance. Due to heavy correlation between tentative and main decisions, however, the two-stage schemes are unable to provide as much improvement over UDFSE as obtained for channel 2.

Next, we simulate a BPSK modulated asynchronous DS-CDMA system with eight users whose signature waveforms are derived from Gold sequences of length 31. The relative delays of users are fixed for the simulation and are in an increasing order. The multiuser channel is static and
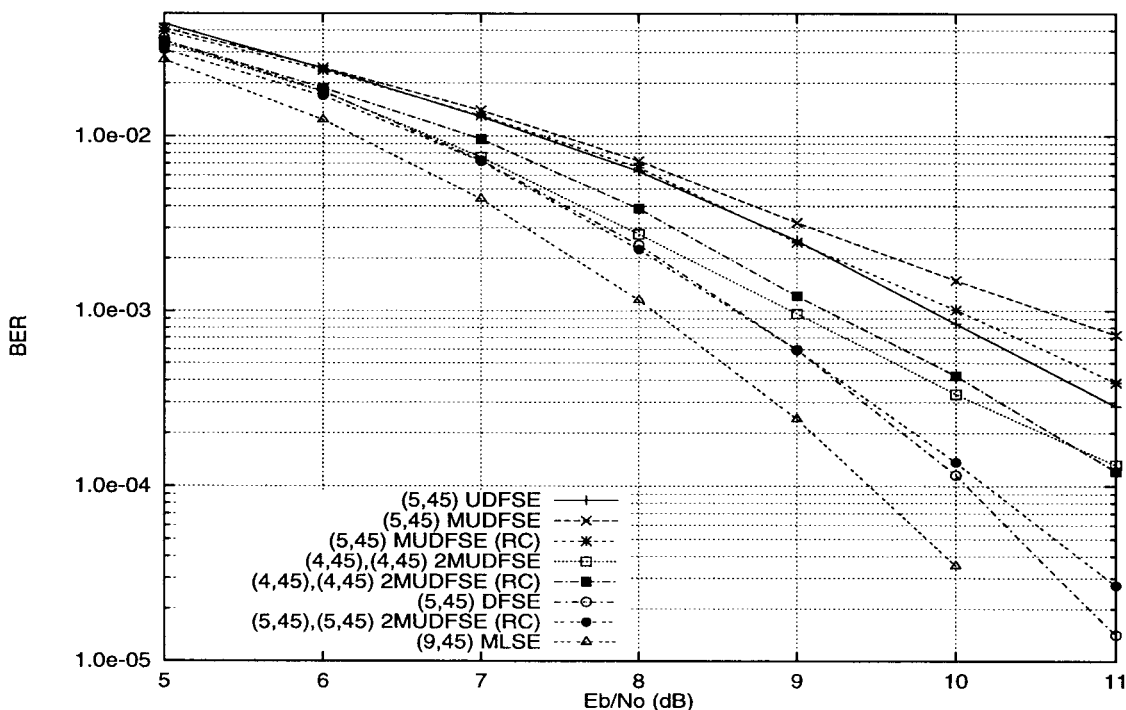
Fig. 4.   BER performance of various detection schemes on channel 3.

has the same spectrum as given in [11].[5] Fig. 5 shows the BER of user 1 for various detection schemes when all users have identical SNR. Each simulation was run for a count of 500 errors. It is evident that even with ideal power control, the performance of the conventional single-user detector is significantly worse than optimum MLSE. (2, 28) UDFSE provides some improvement over the conventional detector. The four-state (unwhitened) M-algorithm that operates directly on the matched filter output is about 2 dB worse than the M-algorithm that operates on the equivalent whitened minimum-phase channel. The linear decorrelator that nulls out all interference loses about 0.5 - 1.0 dB as compared to MLSE due to noise enhancement. (2, 28) DFSE also obtains near MLSE performance. DFSE, the linear decorrelator and the (whitened) M-algorithm, however, require multiuser channel inversion and/or factorization which has complexity quadratic in the number of users. Moreover, the M-algorithm requires sorting of survivor paths at each iteration which is not needed for DFSE or UDFSE as they are trellis based.

The single-stage MUDFSE algorithms which require a bank of matched filters only, obtain the best performance on this channel (next to MLSE). (2, 28) MUDFSE (with four states only) closely approaches the performance of MLSE which requires 128 states in the Viterbi algorithm. With bias approximation, (2, 28) MUDFSE (RC) obtains a gain of 4.5 dB over UDFSE at an error rate of 0.1%.

Fig. 6 shows the BER of user 1 versus the SNR of the rest of the users. The SNR of user 1 is held constant at 7.0 dB. It can be seen that the conventional detector, UDFSE and the unwhitened M-algorithm do not perform well in a near–far

situation. The linear decorrelator, DFSE, and the whitened M-algorithm perform well. The MUDFSE schemes, however, outperform the other methods and converge to MLSE in high SNR of interfering users on this channel.

The improvement afforded by the MUDFSE algorithm over UDFSE depends on several factors. Among these factors is the amount of ISI, the reliability of tentative decisions, and the choice of the memory order. On channels with small sampled correlations, the conventional matched filter detector makes relatively reliable decisions. The reliability of tentative decisions can be improved further by means of an auxiliary stage. If the memory order is not too small, the error performance of the modified detector (28) is dominated by error event distances in the main stage rather than the residual interference. Close to MLSE performance is achieved in this case. Unlike DFSE, error event distances in MUDFSE do not diminish with the memory order, as was discussed in Section V-A. Thus, MUDFSE generally outperforms DFSE (with similar complexity) on channels with small correlations since the residual interference is reliably removed. An example is the multiuser DS-CDMA channel where well-designed spreading codes or long spreading codes in a not-so-highly loaded system provide low channel crosscorrelations.

The MUDFSE algorithm's performance may be worse than that of UDFSE (as well as DFSE) for channels with large sampled correlations (highly dispersive ISI channels or heavily loaded multiuser channels). This is not due mainly to the unreliability of tentative decisions. An auxiliary stage can render relatively reliable tentative decisions. Due to the strong dependence between error events in the auxiliary and the main stage, however, the main stage can not significantly improve upon the decisions obtained in the auxiliary stage.

[5] The simulation results reported in [11] are off by 3 dB (worse) due to a mistake in normalization of symbol energies.
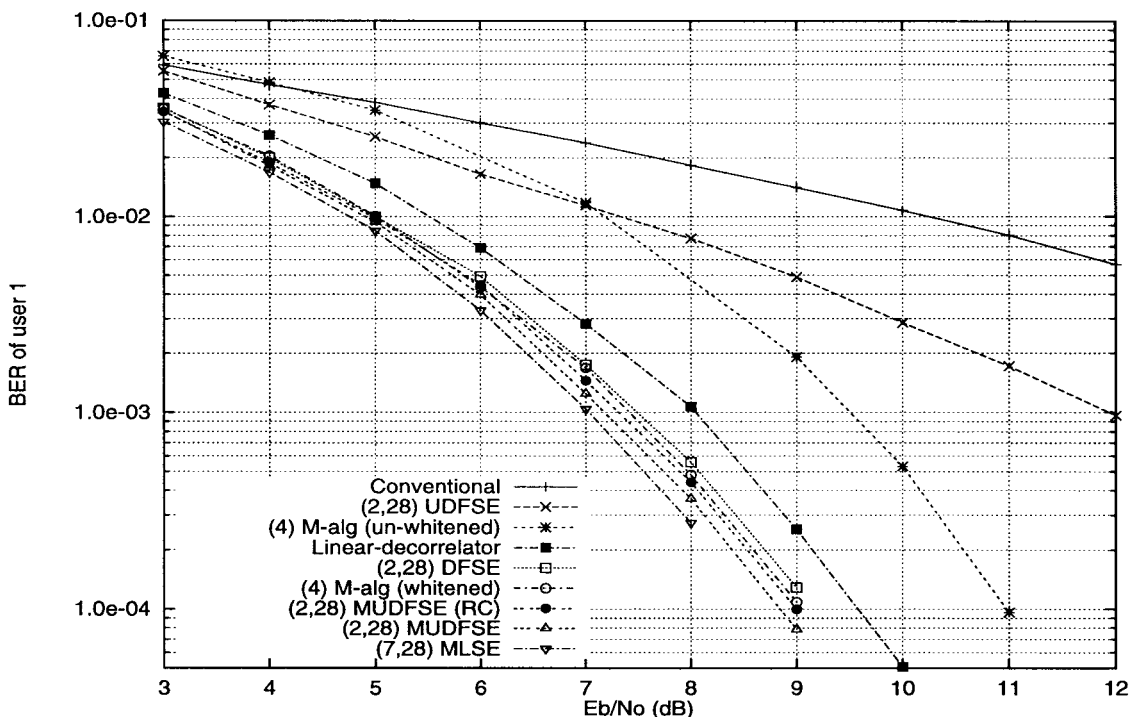
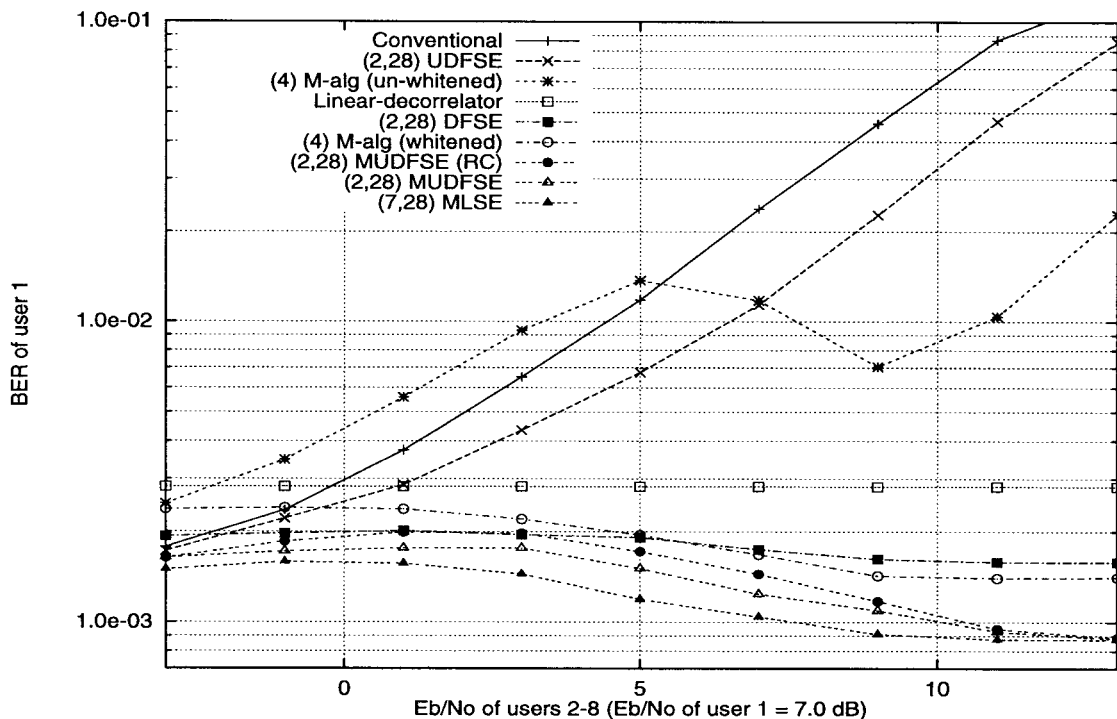Fig. 5. BER performance of various detection schemes over an eight-user asynchronous DS-CDMA channel.

Fig. 6. Near–far performance of various detection schemes over an eight-user asynchronous DS-CDMA channel.

VIII. CONCLUSIONS

In this paper, we show that decision feedback sequence estimation based on conventional matched filter outputs (a reduced-state version of the Ungerboeck MLSE formulation) is affected by untreated interference components. We derive a modified algorithm which uses tentative decisions to cancel the anticausal interference components. The algorithm, which operates on matched filter outputs, provides excellent performance/complexity tradeoffs on channels with relatively small sampled correlations. The algorithm is insensitive to channel phase and does not require noise-whitening. These attributes make it attractive for bidirectional equalization in the GSM system and multiuser detection in asynchronous DS-CDMA systems.

The algorithm can be improved by using soft tentative decisions instead of hard decisions. An adaptive form of the algorithm should be considered for time-varying media. The algorithm can be modified to provide soft outputs for coded systems. An approximate error probability analysis is conducted in this paper which provides some insight into the proposed algorithm. An in-depth analysis is needed, however, to better identify the class of channels for which gains are afforded by the modified algorithm and to quantify the gains for a given channel with given complexity.

REFERENCES

[1] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.

[2] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Trans. Commun. Technol.*, vol. COM-22, pp. 624–636, May 1974.

[3] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.

[4] V. M. Eyuboglu and S. U. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, pp. 13–20, Jan. 1988.

[5] K. Wesolowski, R. Krenz, and K. Das, "Efficient receiver structure for GSM mobile radio," *Int. J. Wireless Inform. Networks*, vol. 3, no. 2, pp. 117–122, Apr. 1996.

[6] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[7] L. Wei and L. K. Rasmussen, "A near ideal noise whitening filter for an asynchronous time-varying CDMA system," *IEEE Trans. Commun.*, vol. 44, pp. 1355–1361, Oct. 1996.

[8] A. Duel and C. Heegard, "Delayed decision feedback sequence estimation for QAM and trellis coded systems," in *Proc. 20th Conf. Information Science and Systems*, Mar. 1986, pp. 367–372.

[9] J. B. Anderson and S. Mohan, "Sequential coding algorithms: A survey cost analysis," *IEEE Trans. Commun.*, vol. COM-32, pp. 169–176, Feb. 1984.

[10] L. Wei, L. K. Rasmussen, and R. Wyrwas, "Near optimum tree-search detection schemes for bit-synchronous multiuser CDMA systems over Gaussian and two-path Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 691–700, June 1997.

[11] A. Hafeez and W. E. Stark, "Decision feedback sequence estimation for unwhitened ISI channels," in *Proc. 35th Annual Allerton Conf. on Communication, Control, and Computing*, Sept. 1997, pp. 493–502.

[12] G. E. Bottomley and S. Chennakeshu, "Unification of MLSE receivers and extension to time-varying channels," *IEEE Trans. Commun.*, vol. 46, pp. 464–472, Apr. 1998.

[13] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.
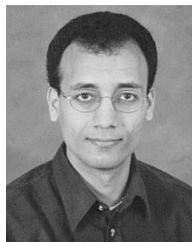
Abdulrauf Hafeez was born on January 4, 1969 in Karachi, Pakistan. He received the B.S. degree in electrical engineering from NED University of Engineering and Technology, Karachi, in 1991 and the M.S. degree in electrical engineering from the University of Michigan, Ann Arbor, in 1994.

He was with Alcatel (Pakistan) Limited as a site engineer working on the deployment of a fiber-optic transmission networks. In 1997 he was at Ericsson Incorporated, Raleigh, North Carolina, where he was engaged in research on baseband techniques for narrow-band mobile communication systems. He is currently pursuing the Ph.D. degree at the University of Michigan. His research interests include equalization, multiuser detection and coding for land-line and mobile communication systems and spread spectrum techniques. He has had several conference publications in these areas. He has a patent pending in the area of cochannel interference cancellation for cellular and PCS systems.
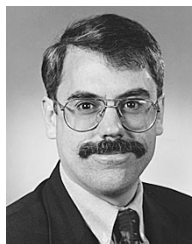
Wayne E. Stark (S'77–M'78–SM'94–F'98) received the B.S. (highest hons.), M.S., and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana–Champaign in 1978, 1979, and 1982, respectively.

Since September, 1982 he has been a faculty member at the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, where he is currently a Professor. He is the principal investigator for an Army Research Office Multidisciplinary University Research Initiative (MURI) project on low energy mobile communications. His research interests are in the areas of coding and communication theory, especially for spread-spectrum and wireless communication networks.

From 1984–1989 Dr. Stark was the Editor for communication theory of the IEEE TRANSACTIONS ON COMMUNICATIONS in the area of spread-spectrum communications. He was involved in the planning and organization of the 1986 International Symposium on Information Theory which was held in Ann Arbor, Michigan. He was selected by the National Science Foundation as a 1985 Presidential Young Investigator. He is a member of Eta Kappa Nu, Phi Kappa Phi, and Tau Beta Pi.