Nonlinear Techniques for the Joint Estimation of Cochannel Signals

K. Giridhar, John J. Shynk, *Senior Member, IEEE*, Amit Mathur, Sujai Chari, and Richard P. Gooch, *Member, IEEE*

***Abstract*— Cochannel interference occurs when two or more signals overlap in frequency and are present concurrently. Unlike in spread-spectrum multiple-access systems where the different users necessarily share the same channel, cochannel interference is a severe hindrance to frequency- and time-division multiple-access communications, and is typically minimized by interference rejection/suppression techniques. In this paper, rather than using interference suppression, we are interested in the *joint* estimation of the information-bearing narrow-band cochannel signals. Novel joint estimators are proposed that employ a single-input demodulator with oversampling to compensate for timing uncertainties. Assuming finite impulse-response channel characteristics, maximum likelihood (ML) and maximum *a posteriori* (MAP) criteria are used to derive cochannel detectors of varying complexities and degrees of performance. In particular, a (suboptimal) *two-stage* joint MAP symbol detector (JMAPSD) is introduced that has a lower complexity than the single-stage estimators while accruing only a marginal loss in error-rate performance at high signal-to-interference ratios. Assuming only reliable estimates of the primary and secondary signal powers, a *blind* adaptive JMAPSD algorithm for *a priori* unknown channels is also derived. The performance of these nonlinear joint estimation algorithms is studied through example computer simulations for two cochannel sources.**

***Index Terms*— Blind adaptation, cochannel signal separation, MAPSD, MLSE.**

I. Introduction

MANY narrow-band communication systems encounter cochannel interference, which is a major impairment to the reliable transmission of information. For example, the performance of telephone systems employing twisted-pair subscriber loops is severely degraded in the presence of cochannel interference called near-end crosstalk [1]. Another example is a cellular radio network employing frequency reuse where one or more secondary signals from nearby cells can interfere with the desired (primary) signal [2]. In addition to cochannel interference, the primary and secondary signals may be corrupted by intersymbol interference (ISI) from long multipath delays and by additive noise. Among these factors, cochannel interference often is the dominant channel impairment. Instead of using interference suppression, as in certain applications like dual-polarized microwave radio [3], [4], we are interested in *jointly* estimating both cochannel signals.

Signal recovery schemes in the presence of ISI and cochannel interference have been developed for various applications, including crosstalk suppression in subscriber loops (see, e.g., [5] and the references therein) and asynchronous multiuser detectors for code division multiple access (CDMA) communications [6]. Other related work include maximum likelihood sequence estimation (MLSE) employed for a multichannel transmission system [7], and joint ML algorithms designed to recover digital signals using antenna arrays [8], [9]. Many other authors have also proposed such multiple-input detectors (see, e.g., [10], [11]).

A single-input cochannel signal separation technique in ISI-free channels for angle-modulated signals was proposed in [12]. To our knowledge, the earliest published work on cochannel signal recovery in the presence of ISI using a single-input receiver was the quasi-linear "demod–remod" technique described in [13]. This demod–remod system is simple to implement; it is inherently a linear approach, with the exception of some final error detection and correction. Recently, linear equalizers exploiting the excess bandwidth or the cyclostationary nature of oversampled digital signals have been proposed for ISI channels [14], [15]. These approaches require only a single-input demodulator, and have been extended to decision-feedback equalization (DFE) [16]. In these techniques, the amount of excess bandwidth required is directly proportional to the number of cochannel signals to be jointly recovered.

In this paper, we are interested in cochannel signal demodulation using a single-input/multiple-output (SIMO) detector. Although the combined channel model and detector constitute a multiple-input/multiple-output (MIMO) system, we refer to the detector as being SIMO because it receives and processes a single (complex) data stream. This is done to distinguish the algorithms from those based on an array of antenna elements (as in [8], for example), which process multiple input data streams (i.e., they incorporate spatial as well as temporal processing). Although the approaches discussed here

Paper approved by E. Eleftheriou, the Editor for Equalization and Coding of the IEEE Communications Society. Manuscript received April 7, 1995; revised June 5, 1996 and October 16, 1996. This work was supported by Applied Signal Technology, Inc., the University of California MICRO Program, and the National Science Foundation under Grant MIP 9308919.

K. Giridhar was with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA. He is now with the Department of Electrical Engineering, Indian Institute of Technology, Madras 600036, India.

J. J. Shynk is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA.

A. Mathur was with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA. He is now with Lucent Technologies, Holmdel, NJ 07733-3030 USA.

S. Chari was with Applied Signal Technology, Inc., Sunnyvale, CA 94086 USA. He is now with Amati Communications, San Jose, CA 95124 USA.

R. P. Gooch is with Applied Signal Technology, Inc., Sunnyvale, CA 94086 USA.

Publisher Item Identifier S 0090-6778(97)02781-5.

may be extended to multiple cochannel sources, we consider a communication system model with one secondary data stream to be jointly recovered along with the primary data stream.

We propose a single-input receiver with oversampling where the sampling rate is chosen to provide sufficient insensitivity to asynchronous sampling, but is otherwise independent of the actual number of cochannel data streams to be recovered. We describe a technique based on MLSE using the Viterbi algorithm (VA) [17], and propose a MAP symbol detection (MAPSD) algorithm [18] based on a Bayesian recursion. Since the cochannel data symbols are jointly recovered, we refer to the corresponding algorithms as joint MLSE (JMLSE) and joint MAPSD (JMAPSD) [19], [20].

JMLSE and the closely related JMAPSD are optimal techniques that can be expected to yield an improved bit-error rate (BER) performance compared to the linear/DFE approaches in [14]–[16]. It is possible to implement the JMAPSD algorithm using a suboptimal *two-stage* configuration (employing a secondary feedback mechanism reminiscent of the demod–remod structure in [13]) to substantially reduce the computational complexity. This feedback also greatly enhances the BER performance for low power separation between the data streams, i.e., low signal-to-interference ratios (SIR's).

Another joint estimation scheme for overcoming cochannel interference was proposed recently in [21] where an adaptive radial basis function (RBF) equalizer was employed to approximate the optimal Bayesian classifier for the assumed cochannel signal model. Although the nonlinear RBF equalizer may outperform linear equalizers, its complexity increases exponentially with the order of the equalizer. Moreover, the Bayesian classifier is strictly suboptimal compared to MLSE and MAPSD algorithms of comparable complexity [22]. Hence, we expect the JMLSE and JMAPSD approaches to yield a superior error rate performance compared to the RBF equalizer.

Finally, we describe a *blind* MAP cochannel symbol detector for *a priori* unknown channels. When implemented as a two-stage structure, this joint blind MAPSD (JBMAPSD) algorithm [23] adaptively computes first the stronger (primary) signal and the corresponding channel estimates. The secondary stage is adapted once the primary coefficients are close to convergence (which is typically about 100–200 samples later). While this two-stage JBMAPSD may be employed when the primary and secondary signals have a reasonable power separation (e.g., SIR $>$ 20 dB), the single-stage JBMAPSD algorithm will be necessary for lower SIR's.

The paper is organized as follows. The fractionally spaced cochannel signal and receiver model is defined in Section II. For *known* channels, the JMLSE and JMAPSD approaches are described in Sections III and IV, respectively. For *a priori unknown* channel characteristics, the joint blind MAP symbol detector is described in Section V. Computer simulation results are presented in Section VI, and conclusions are outlined in Section VII.

II. Cochannel Signal Model

The assumed cochannel system model is shown in Fig. 1. The transmitted low-pass equivalent waveforms can be represented by

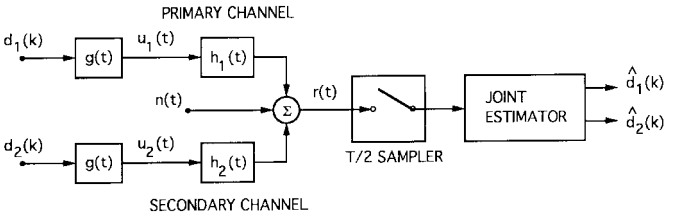

Fig. 1. Cochannel system model.

$$u_m(t) = \sum_k d_m(k) g(t - kT), \qquad m = 1, 2 \tag{1}$$

where T is the symbol duration and $\{d_1(k)\}$ and $\{d_2(k)\}$ are the primary and secondary source symbols, respectively. The pulse function $g(t)$ has a raised-cosine spectrum with a bandwidth B given by $1/2T < B < 1/T$. We propose a $T/2$-spaced implementation, although it has about twice the complexity of a T-spaced implementation, for the following reasons. 1) It eliminates the need for two whitening matched filters. 2) It takes advantage of any excess bandwidth in $g(t)$ (although no critical choice of excess bandwidth is required). 3) We expect the $T/2$-spaced receiver to be nearly insensitive to sampling time offsets, and hence be capable of recovering nonsynchronized cochannel signals more easily than a T-spaced implementation.

Thus, the discrete-time measurement samples of the received signal $r(t)$ at the output of the $T/2$ sampler in Fig. 1 are given by (for $j = 0, 1$)

$$\begin{aligned} r(kT + jT/2) &= \sum_{m=1}^{2} \sum_{p=0}^{L_m} f_{m,p}(kT + jT/2) d_m(k - p) \\ &\quad + n(kT + jT/2) \\ &= r_1(kT + jT/2) + r_2(kT + jT/2) \\ &\quad + n(kT + jT/2) \end{aligned} \tag{2}$$

where the noise sequences $\{n(kT + jT/2)\}$ are assumed to be mutually uncorrelated, white, and Gaussian with zero mean and variance σ_n^2. The delay spreads of the primary and secondary channels are $L_1 T$ and $L_2 T$, respectively. The $L_1 + L_2 + 2$ channel coefficients $\{f_{m,p}(kT + jT/2)\}$ represent the convolution of the impulse responses $\{h_m(t)\}$ of the physical communication medium with that of the transmit filter $g(t)$ sampled at $T/2$ seconds. The cochannel interference component from the mth signal is given by

$$\begin{aligned} &r_m(kT + jT/2) \\ &\quad = \sum_{p=0}^{L_m} f_{m,p}(kT + jT/2) d_m(k - p), \qquad j = 0, 1. \end{aligned} \tag{3}$$

Asynchronous arrival of the two data streams could result in symbol timing offsets. These effects are assumed to be implicitly modeled by the channel coefficients.

For notational convenience and clarity in the sequel, we will use a T-spaced representation for the variables even though the underlying model has two samples per symbol; for example,

we will use $r(k)$ instead of $r(kT+jT/2)$. It should be mentioned that, throughout this work, we implicitly assume that good estimates of the primary and secondary signal powers are available *a priori*. Assuming gain-normalized channels, the signal-to-noise ratio (SNR) at the receiver input is defined by $\text{SNR} \triangleq 10\log(P_1/N_o)$, while in the presence of one interfering signal, the signal-to-interference ratio is defined by $\text{SIR} \triangleq 10\ \log(P_1/P_2)$. The signal powers are $P_j = E[d_j^2(k)]$ and $N_o/2$ is the noise power spectral density. For the computer simulations presented later, we fixed the primary signal power at $P_1 = 1$; with binary signaling (i.e., BPSK), $P_2 = A^2$ where $d_2(k) \in \{+A, -A\}$ and $0 < A \leq 1$.

The goal of the receiver is to accurately estimate the primary and secondary sequences $\{d_1(k)\}$ and $\{d_2(k)\}$ using estimates of the channel coefficient vectors $\boldsymbol{f}_1(k)$ and $\boldsymbol{f}_2(k)$ (where $\boldsymbol{f}_m(k) \triangleq [f_{m,0}(k),\ f_{m,1}(k), \cdots, f_{m,L_m}(k)]^T$). These coefficients are either known (for JMAPSD and JMLSE) or blindly estimated (by JBMAPSD).

III. Joint ML Sequence Estimation

In the single-channel scenario, the aim of MLSE is to determine the one sequence $d_i^k = \{d_i(k), d_i(k-1), \cdots, d_i(0)\}$ out of all possible transmitted symbol sequences such that $p(r^k|d_i^k) \geq p(r^k|d_j^k), \forall j \neq i$, where $r^k = \{r(k), r(k-1), \cdots, r(0)\}$ is the received sequence. When the additive noise components in (2) are independent and Gaussian, the above condition can be replaced by a Euclidean distance criterion given by (for T-spaced MLSE)

$$d_i^k : \sum_{l=0}^{k} |r(l) - \hat{r}_i(l)|^2 \leq \sum_{l=0}^{k} |r(l) - \hat{r}_j(l)|^2, \qquad \forall j \neq i \tag{4}$$

where $\{\hat{r}_i(k)\}$ are the signal estimates generated from d_i^k using the known channel coefficients. (In $T/2$-spaced MLSE, each term in the summation is replaced by two similar terms, one for each sample per symbol.) For the joint detection of two cochannel signals, the objective of JMLSE [19], [20] is to determine the *pair* of sequences $\{d_{1,i}^k, d_{2,j}^k\}$ that minimize the sum of squared errors defined by the error (likelihood) sequence $c_{i,j}^k$, as illustrated in Fig. 2. When the channel has a finite impulse response (FIR), the Viterbi algorithm (VA) is a practical way of implementing optimal (single-channel) MLSE [24]. Assuming a channel memory of L symbols, the VA maintains a decoding trellis with M^L nodes or states (where M is the size of the source alphabet) and an equal number of survivor sequences. Each state is a particular subsequence of L previously transmitted symbols $\{d(k-1), \cdots, d(k-L)\}$ from which the present symbol $d(k)$ could be obtained. For example, the ith state[1] is defined by $d_i^{k-1,L-1} \triangleq \{d_i(k-1), \cdots, d_i(k-L)\}$. It is evident that $d_i^{k-1,L-1}$ can transition to M possible states at time k, and that it could have been reached from M different states at time $k-2$. The VA decisions are usually computed by truncating the survivors after $4L - 5L$ symbols [24].

[1] Although a state is based on L previously transmitted symbols, we use $L-1$ in the superscript of $d_i^{k-1,L-1}$ to be consistent with the notation used for the JMAPSD algorithm described in the next section.

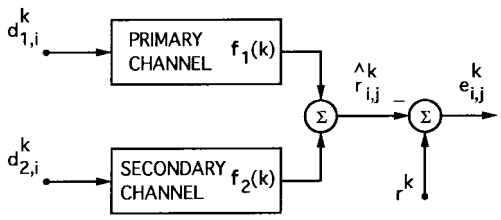

Fig. 2. Joint ML sequence estimation.

The joint VA (JVA) [19], [20], for JMLSE is implemented with a method very similar to that of the standard VA. A joint state $D_i^{k-1,L-1} \triangleq \{d_{1,i}^{k-1,L_1-1}, d_{2,i}^{k-1,L_2-1}\}$ is defined by appending the primary $(d_{1,i}^{k-1,L_1-1})$ and secondary $(d_{2,i}^{k-1,L_2-1})$ states. Hence, the number of states required to implement the optimal JVA is $M^{L_1+L_2}$. Observe that, in this case, each joint state at time $k-1$ can transition to M^2 states at time k, and can be reached by the same number of states from time $k-2$. For high-order signal constellations (e.g., 16-PSK or 64-QAM), complexity reduction techniques originally developed for single-channel MLSE, such as reduced-state sequence estimation (RSSE) [25], may be employed for JMLSE. A discussion of this approach is beyond the scope of this paper, but we have found that RSSE is beneficial at least for the case of minimum-phase channels.

IV. Joint MAP Symbol Detection

The MAP symbol-by-symbol decoding scheme [18] minimizes the probability of a symbol error, i.e.,

$$J = \max_{d(k-L)} p(d(k-L)|r^k) \tag{5}$$

and can provide more reliable decisions than the VA for the same decoding delay. This can be understood from the fact that the probability of symbol detection is equal to the sum of all the sequence probabilities $p(d_i^k|r^k)$ containing that symbol, i.e.,

$$\max_{d(k-L)} p(d(k-L)|r^k) = \max_{d(k-L)} \sum_{\{i:d_i(k-L)=d(k-L)\}} p(d_i^k|r^k). \tag{6}$$

Symbol-error-rate (SER) curves of the *blind* MAPSD algorithm and the blind VA for QPSK signals and fast-fading ISI channels were investigated in [22]. In those simulations, we found that for a decoding delay equal to the memory of the ISI channel (we can consider this as being equivalent to symbol-by-symbol decoding), the blind MAPSD algorithm is superior to the blind VA by nearly 0.5 dB. Thus, we expect that the blind-adaptive JMAPSD algorithm (discussed in Section V) can have an error rate performance superior to that of blind-adaptive JMLSE (based on the VA). However, for *known* time-invariant channels, the advantage of the (nonblind) MAPSD algorithm over the VA is quite small (about 0.1 dB—see [22]). Moreover, once the decoding delay of the VA is increased, its performance approaches that of optimal MLSE, and is superior to that of the symbol-by-symbol MAP detector.

Noting that $p(d_i^k|r^k)$ is proportional to the *a posteriori* MLSE metric $p(r^k|d_i^k)$, the VA saves only the metrics of the survivor sequences, and not of all the sequences containing $d(k-L)$. On the other hand, the MAPSD algorithm maintains a MAP metric $p(d_i^{k,L}|r^k)$ for every *subsequence* of length $L+1$ defined by $d_i^{k,L} \triangleq \{d_i(k), \cdots, d_i(k-L)\}$. Motivated by the FIR nature of the channel, (5) can be rewritten as

$$J = \max_{d(k-L)} \sum_{d(k)} \cdots \sum_{d(k-L+1)} p(d_i^{k,L}|r^k) \tag{7}$$

where the subsequences are such that $d_i(k-L) = d(k-L)$. The recursion to calculate the ith MAP metric in the above summations is given by

$$p(d_i^{k,L}|r^k) = \frac{1}{c} p(r(k)|r^{k-1}, d_i^{k,L}) \cdot \sum_{\left\{j: d_j^{k-1,L} \in d_i^{k,L}\right\}} p(d_j^{k-1,L}|r^{k-1}) \tag{8}$$

where the normalization constant $c = Mp(r(k)|r^{k-1})$. The summation in (8) is performed over the MAP metrics of all possible subsequences $d_j^{k-1,L}$ at time $k-1$ from which $d_i^{k,L}$ could have been obtained. The likelihood $p(r(k)|r^{k-1}, d_i^{k,L})$ (which equals $p(r(k)|d_i^{k,L})$ when the channel is known) is Gaussian, i.e.,

$$\begin{aligned} p(r(k)|d_i^{k,L}) &= \mathcal{N}(r(k); \hat{r}_i(k), \sigma_n^2) \\ &= \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{|r(k) - \hat{r}_i(k)|^2}{2\sigma_n^2}\right] \end{aligned} \tag{9}$$

where $\hat{r}_i(k) = \boldsymbol{h}_i(k)\boldsymbol{f}(k-1)$, $\boldsymbol{h}_i(k) = [d_i(k), \cdots, d_i(k-L)]$ is the data (row) vector, and $\boldsymbol{f}(k-1) = [f_0(k-1), \cdots, f_L(k-1)]^T$ is the coefficient (column) vector. The corresponding detection algorithm was first derived in [18] for known channels, and was extended to blind estimation for unknown channels in [26], [27].

The complexity of the single-channel MAPSD algorithm is roughly the same order as that of MLSE. (Note that although MLSE maintains M^L states, it calculates the same number of likelihoods as does MAPSD with M^{L+1} subsequences.) A suboptimal MAP rule was introduced in [26] to make a decision on the $(k-L)$th symbol (at time k) according to $\hat{d}(k-L) = d_j(k-L)$ where

$$\hat{d}^{k,L} = d_j^{k,L}, \quad j = \arg\max_i p(d_i^{k,L}|r^k). \tag{10}$$

The complexity of the single-channel MAPSD algorithm can be reduced by introducing decision feedback. In this MAP/decision-feedback (MAP/DF) approach [28], [29], a DF filter of length $L_{\text{df}} \leq L+1$ is cascaded with the MAPSD algorithm to truncate the effective channel memory. The size of the MAPSD section is reduced accordingly to $M^{L+1-L_{\text{df}}}$ states. Hence, a performance-complexity tradeoff is possible, ranging from that of the full MAP estimator ($L_{\text{df}} = 0$) to the ideal DFE ($L_{\text{df}} = L$).[2] The MAP/DF approach is similar to the delayed decision-feedback sequence estimator (DDFSE) presented in [30] for the VA. The main difference is that a single feedback filter may be used in the MAPSD algorithm, whereas in DDFSE, each state in the VA trellis employs feedback.

[2]The ideal DFE refers to an L-tap feedback filter which cancels the postcursor ISI; there is no feedforward section to handle the precursor ISI.

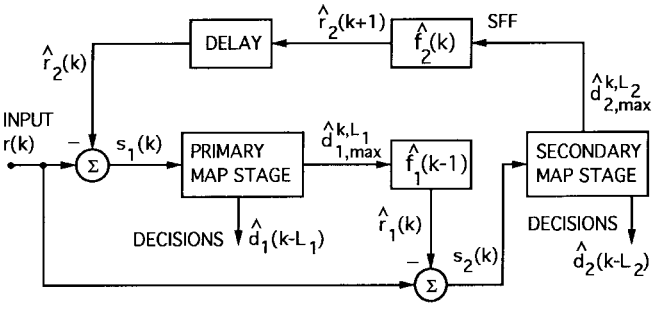

Fig. 3. Two-stage JMAPSD algorithm.

For cochannel symbol detection, an optimal JMAPSD algorithm (of complexity $M^{L_1+L_2+2}$) may be obtained by modifying the single-channel MAPSD algorithm using joint subsequences, i.e., $D_i^{k,L} = \{d_{1,i}^{k,L_1}, d_{2,i}^{k,L_2}\}$. This single-stage JMAPSD algorithm should provide a BER performance comparable to that of JMLSE for the same decoding delay. However, it is preferable to use the JVA when longer decoding delays can be tolerated because the complexity of JMAPSD grows exponentially with the decoding delay, while it is essentially linear for the JVA (once the trellis is constructed). Because of this complexity, in the computer simulations of Section VI, we compare the algorithms to single-stage JMLSE in place of single-stage JMAPSD.

A computational advantage is obtained when the single-stage JMAPSD algorithm is reconfigured as a (suboptimal) *two-stage* algorithm, as illustrated in Fig. 3. The subsequence decisions of the primary MAP stage, denoted by $\hat{d}_{1,\max}^{k,L_1}$ and corresponding to the largest probability metric [given below in (13)], are used to compute the primary signal estimate

$$\hat{r}_1(k) = \sum_{p=0}^{L_1} f_{1,p}(k-1)\hat{d}_{1,\max}(k-p), \tag{11}$$

yielding the residual error signal $s_2(k) = r(k) - \hat{r}_1(k)$. This error becomes the input of the second MAP stage, which models the M^{L_2+1} subsequences of the secondary channel. Hence, the complexity of two-stage JMAPSD is only on the order of $M^{L_1+1} + M^{L_2+1}$. The assumption here is that the SIR is sufficiently large such that the primary MAP metrics converge; thus, cancellation of the primary signal component is nearly complete, and $s_2(k)$ contains only the secondary signal component (plus additive noise). Note that we could also derive a two-stage JMLSE implementation utilizing RSSE. However, since low-delay decisions would be required at the output of the first stage, JMAPSD is preferable because it can yield a lower error rate than the VA for the same delay.

A feedback filter may be used to subtract a partial estimate of the secondary signal from the input. As a result, the two-stage JMAPSD algorithm will perform satisfactorily even under low-to-medium SIR conditions. This secondary feedback filter (SFF) is also shown in Fig. 3 where the partial estimate $\hat{r}_2(k)$ is evaluated using the *last* L_2 suboptimal

decisions of $\hat{d}_{2,\max}^{k-1,L_2}$ [i.e., using all the suboptimal decisions at time $k-1$ with the exception of the MAP symbol decision $\hat{d}_{2,\max}(k-1-L_2)$] as follows

$$\hat{r}_2(k) = \sum_{p=1}^{L_2} f_{2,p}(k-1)\hat{d}_{2,\max}(k-p). \tag{12}$$

Hence, $\hat{r}_2(k)$ is an estimate of the secondary interference from all L_2 previous secondary symbols excluding the current symbol $d_2(k)$. The residual signal $s_1(k) = r(k) - \hat{r}_2(k)$ becomes the input to the primary stage.

The following remarks can be made regarding this two-stage scheme. 1) Although the SFF may introduce some error propagation due to the decision feedback, for low SIR simulations we have found a significant improvement in performance.[3] 2) Due to the uncancelled ISI contribution from the current secondary symbol $d_2(k)$ (even when the decisions entering the computation of $\hat{r}_2(k)$ are correct), the residual $s_1(k)$ of the primary MAP stage conditioned on the input sequence is not Gaussian in general.[4] However, under low-SNR and/or high-SIR conditions, the effect of this additional term on the overall error-rate performance is likely to be minimal. Thus, the conditional likelihoods in the primary MAP stage are approximated by Gaussian functions in order to employ the MAPSD approach. Note that some performance degradation can be expected at low-SIR conditions if the noise power σ_n^2 is reduced. This anomalous behavior at high SNR's has been observed in computer simulations, and is found to be dependent on the uncancelled ISI term $f_{2,0}(k)d_2(k)$ of the secondary channel.

The MAP metrics in each stage are updated independently as in the single-channel MAPSD algorithm, i.e.,

$$p\big(d_{m,i}^{k,L_m}|r^k\big) = \frac{1}{c}p\big(s_m(k)|r^{k-1}, d_{m,i}^{k,L_m}\big) \cdot \sum_{\left\{j:d_{m,j}^{k-1,L_m} \in d_{m,i}^{k,L_m}\right\}} p\big(d_{m,j}^{k-1,L_m}|r^{k-1}\big) \tag{13}$$

where $s_m(k)$ is the input to the mth MAP stage as defined earlier, and the conditional likelihood is assumed to be Gaussian, i.e.,

$$p\big(s_m(k)|r^{k-1}, d_{m,i}^{k,L}\big) = \mathcal{N}\big(s_m(k); \hat{r}_{m,i}(k), \sigma^2\big) \tag{14}$$

where $\hat{r}_{m,i}(k)$ is the signal estimate assuming $d_{m,i}^{k,L_m}$ was sent, and σ^2 depends on the noise power. The summation in (13) is similar to that in (8).

The overall two-stage JMAPSD is summarized in Table I using the T-spaced notation. For a $T/2$-spaced implementation, note that the likelihood function in (9) will be a product of the even and odd sample likelihoods as follows

$$\begin{aligned} p(r(k)|d_i^{k,L_m}) &= \mathcal{N}\big(s_m(kT); \hat{r}_{m,i}(kT), \sigma^2\big) \\ &\quad \cdot \mathcal{N}\big(s_m(kT+T/2); \hat{r}_{m,i}(kT+T/2), \sigma^2\big) \end{aligned} \tag{15}$$

where $\hat{r}_{m,i}(kT+jT/2) = \boldsymbol{h}_{m,i}(k)\boldsymbol{f}_m(kT+jT/2),\ j=0,1$ [similar to (3)]. Referring to (11) and (12), note that for the primary MAP stage, $s_1(kT+jT/2) = r(kT+jT/2) - \hat{r}_2(kT+jT/2)$, while for the secondary MAP stage, $s_2(kT+jT/2) = r(kT+jT/2) - \hat{r}_1(kT+jT/2)$. In general, it may be argued that, depending on the ISI shaping and random timing jitter introduced by the channels, either the odd or the even sample is bound to capture more of the received signal power. Thus, the corresponding likelihood term in (15) may be more reliable (i.e., it may be closer to the true likelihood obtained in the absence of timing offset). However, both the even and odd sample likelihoods in (15) are given the same weight. It may be worthwhile to see if any BER performance advantage is obtained by using some other weighting scheme, but we have not investigated this. In computer simulations, we have achieved good performance by equally employing the even and odd samples.

V. Joint Blind MAPSD (JBMAPSD) Algorithm

The JBMAPSD algorithm is also based on the two-stage structure in Fig. 3 where each stage, in addition to demodulating one cochannel signal, also estimates the corresponding channel coefficients. The adaptation algorithm for each stage resembles the T-spaced, single-channel blind MAPSD algorithm described in [26], [27]. In this blind algorithm, a conditional channel estimate $\widehat{\boldsymbol{f}}_{m,i}^{c}(k)$ is maintained for each subsequence $d_{m,i}^{k,L_m}$ (of length L_m). Once the the MAP metrics are updated, "unconditional" estimates $\widehat{\boldsymbol{f}}_{m,i}(k)$ are obtained from the appropriate predecessor metrics and conditional estimates [22], [27]. We extend each of these steps to the two-stage JBMAPSD algorithm.

The original single-channel blind MAPSD algorithm assumed a first-order autoregressive model for the channel coefficients, and incorporated a Kalman filter to update the channel and error covariance estimates. In order to reduce the complexity, simpler stochastic gradient adaptation was employed in [27] resulting in LMS (least-mean-square) [31] update rules. The LMS update may be viewed as a stochastic gradient descent on a conditional cost function; a direct extension of this method yields the single-stage JBMAPSD algorithm with the conditional cost function given by

$$\begin{aligned} \min_{\widehat{\boldsymbol{f}}_{1,i}(k-1), \widehat{\boldsymbol{f}}_{2,i}(k-1)} & J_i^{\text{opt}}(k) \\ &= \min_{\widehat{\boldsymbol{F}}_i(k-1)} \big|r(k) - \boldsymbol{H}_i(k)\widehat{\boldsymbol{F}}_i(k-1)\big|^2, \\ & \qquad i = 1, \cdots, M^{L_1+L_2+2} \end{aligned} \tag{16}$$

where $\boldsymbol{H}_i(k) = [\boldsymbol{h}_{1,i}(k), \boldsymbol{h}_{2,i}(k)]$ is the joint subsequence, and $\widehat{\boldsymbol{F}}_i(k-1) = [\widehat{\boldsymbol{f}}_{1,i}^T(k-1), \widehat{\boldsymbol{f}}_{2,i}^T(k-1)]^T$ is the joint channel estimate. The superscript $^{\text{opt}}$ indicates that the resulting algorithm is optimal (in the LMS sense) because the minimization is done jointly over both coefficient vectors, yielding a single-stage JBMAPSD algorithm.

[3] It should be mentioned that this two-stage JMAPSD structure (also described in [20]) is slightly different from the one proposed in [19], but is exactly equivalent in function.

[4] It is not Gaussian even if the secondary signal is completely cancelled (except for the current secondary symbol $d_2(k)$, as mentioned above).

TABLE I
SUMMARY OF THE TWO-STAGE JMAPSD ALGORITHM.

Primary MAP Stage

STEP 1: Determine Primary Input

$$s_1(k) = r(k) - \hat{r}_2(k)$$

STEP 2: Update Primary MAP Metrics Using (13)

STEP 3: Compute Primary Decisions Using (10)

STEP 4: Compute Total Estimate of Primary Signal

$$\hat{r}_1(k) = \sum_{p=0}^{L_1} f_{1,p}(k-1)\hat{d}_{1,max}(k-p)$$

Secondary MAP Stage

STEP 5: Determine Secondary Input

$$s_2(k) = r(k) - \hat{r}_1(k)$$

STEP 6: Update Secondary MAP Metrics Using (13)

STEP 7: Compute Secondary Decisions Using (10)

STEP 8: Compute Partial Estimate of Secondary Signal

$$\hat{r}_2(k+1) = \sum_{p=1}^{L_2} f_{2,p}(k)\hat{d}_{2,max}(k+1-p)$$

Return to *STEP 1*

On the other hand, for the two-stage configuration in Fig. 3, two separate cost functions are considered:

$$\begin{aligned} J_{1,i}^{\text{sub}}(k) &= |r_1(k) - \hat{r}_{1,i}(k)|^2 \\ J_{2,j}^{\text{sub}}(k) &= |r_2(k) - \hat{r}_{2,j}(k)|^2 \end{aligned} \tag{17}$$

where $i = 1, \cdots, M^{L_1+1},\ j = 1, \cdots, M^{L_2+1}$,

$$\hat{r}_{m,i}(kT + jT/2) = \boldsymbol{h}_{m,i}(k)\widehat{\boldsymbol{f}}_{m,i}((k-1)T + jT/2), \tag{18}$$

which is based on the channel estimates from the previous instant. Decoupling of the primary and secondary cost functions is clearly suboptimum (hence the superscript $^{\text{sub}}$). Because $r_1(k)$ and $r_2(k)$ are not directly available, they must be approximated using the previous subsequence decisions. Referring to Fig. 3 and assuming that the primary decisions $\hat{d}_{1,\max}^{k,L_1}$ are correct, $r_2(k)$ is replaced by

$$s_2'(k) \equiv s_2(k) = r(k) - \hat{r}_1(k) \tag{19}$$

where $\hat{r}_1(k) = \mathbf{h}_{1,\max}(k)\widehat{\mathbf{f}}_{1,\max}(k-1)$ is determined from the subsequence and channel estimates corresponding to the largest MAP metric. In a similar manner, $r_1(k)$ is replaced by

$$\begin{aligned} s_1'(k) &= [r(k) - \hat{r}_2(k)] - \hat{f}_{2,0}^{\max}(k-1)\hat{d}_{2,\max}(k) \\ &= s_1(k) - \hat{f}_{2,0}^{\max}(k-1)\hat{d}_{2,\max}(k) \end{aligned} \tag{20}$$

where $\hat{f}_{2,0}^{\max}(k-1)$ and $\hat{d}_{2,\max}(k)$ are also determined by the largest MAP metric. Thus, in addition to the partial secondary estimate $\hat{r}_2(k)$ produced by the SFF in Fig. 3, the decision

$\hat{d}_{2,\max}(k)$ obtained after updating the secondary MAP metrics is also fed back and weighted by the first coefficient of $\widehat{\boldsymbol{f}}_{2,\max}(k-1)$. In this way, the accuracy of the gradient value entering the computation of $\widehat{\boldsymbol{f}}^{c}_{1,i}(k)$ is enhanced.

Using the above assumptions, namely, $s'_1(k) \approx r_1(k)$ and $s'_2(k) \approx r_2(k)$, the conditional gradient estimates for the mth MAP stage $(m = 1, 2)$ are given by

$$\frac{\partial J^{\text{sub}}_{m,i}(k)}{\partial \widehat{\boldsymbol{f}}_{m,i}(k-1)} = -2[s'_m(k) - \boldsymbol{h}_{1,i}(k)\widehat{\boldsymbol{f}}_{m,i}(k-1)]\boldsymbol{h}^H_{m,i}(k), \quad i = 1, \cdots, M^{L_m+1}, \tag{21}$$

yielding the following LMS update for the ith conditional channel estimate in the mth stage

$$\widehat{\mathbf{f}}^{c}_{m,i}(k) = \widehat{\mathbf{f}}_{m,i}(k-1) + \mu(k)\mathbf{h}^H_{m,i}(k)[s'_m(k) - \hat{r}_{m,i}(k)]. \tag{22}$$

The corresponding unconditional estimate is

$$\widehat{\boldsymbol{f}}_{m,i}(k) = \sum_{\left\{j: d^{k,L_m}_{m,j} \in d^{k+1,L_m}_{m,i}\right\}} \widehat{\boldsymbol{f}}^{c}_{m,j}(k) \cdot \frac{p\left(d^{k,L_m}_{m,j} | r^k\right)}{\sum_{\left\{n: d^{k,L_m}_{m,n} \in d^{k+1,L_m}_{m,i}\right\}} p\left(d^{k,L_m}_{m,n} | r^k\right)}. \tag{23}$$

These summations are defined in a manner similar to that in (8).

When $T/2$-spaced estimators are employed, separate cost functions are defined for the even sample $r(kT)$ and the odd sample $r(kT+T/2)$. For the primary MAP stage, the even and odd cost functions are given by [analogous to $J^{\text{sub}}_{1,i}(k)$ in (17)]

$$J^{\text{sub}}_{1,i}(kT + jT/2) = |r_1(kT + jT/2) - \hat{r}_{1,i}(kT + jT/2)|^2, \quad j = 0, 1 \tag{24}$$

and the corresponding even and odd conditional gradients are (analogous to (21) for $m = 1$)

$$\frac{\partial J^{\text{sub}}_{1,i}(kT + jT/2)}{\partial \widehat{\boldsymbol{f}}_{1,i}(kT - jT/2)} = -2[s'_1(kT + jT/2) - \boldsymbol{h}_{1,i}(k)\widehat{\boldsymbol{f}}_{1,i}(kT - jT/2)]\boldsymbol{h}^H_{1,i}(k) \tag{25}$$

for $j = 0, 1$, and $i = 1, \cdots, M^{L_1+1}$. Observe that, although the error terms for $j = 0, 1$, are different, the even and odd conditional updates above use the same data vector $\boldsymbol{h}_{1,i}(k)$. Finally, the conditional likelihood computation in (14) is modified for the primary MAP stage to be

$$p(s_1(k) | r^{k-1}_1, d^{k,L}_{1,i}) = \mathcal{N}\left(s_1(kT); \hat{r}_{1,i}(kT), \sigma^2\right) \cdot \mathcal{N}\left(s_1(kT + T/2); \hat{r}_{1,i}(kT + T/2), \sigma^2\right) \tag{26}$$

because the $\{n(kT + jT/2)\}$, $j = 0, 1$, are assumed to be mutually uncorrelated, and where $\hat{r}_{1,i}(kT + jT/2) = \boldsymbol{h}_{1,i}(k)\widehat{\boldsymbol{f}}_{1,i}((k-1)T + jT/2)$, $j = 0, 1$. The form of the unconditional update remains unchanged from (23), except that the even and odd estimates can be computed jointly by stacking them together (since they have the same MAP metrics).

TABLE II
SUMMARY OF THE TWO-STAGE JBMAPSD ALGORITHM.

STEPS 1-3: Similar to *STEPS 1-3* in Table 1
STEP 4: Compute Total Estimate of Primary Signal $\hat{r}_1(k) = \sum_{p=0}^{L_1} \hat{f}^{max}_{1,p}(k-1)\hat{d}_{1,max}(k-p)$
STEPS 5-7: Similar to *STEPS 5-7* in Table 1
STEP 8: Improve the Accuracy of the Primary Signal for *STEP 9a* Using the Latest Decision from the Secondary Stage $s'_1(k) = s_1(k) - \hat{f}^{max}_{2,0}(k-1)\hat{d}_{2,max}(k)$
STEP 9a: Compute Primary Conditional Updates Using (22) *STEP 9b*: Compute Primary Unconditional Updates Using (23)
STEP 10a: Compute Secondary Conditional Updates Using (22) *STEP 10b*: Compute Secondary Unconditional Updates Using (23)
STEP 11: Compute Partial Estimate of Secondary Signal $\hat{r}_2(k+1) = \sum_{p=1}^{L_2} \hat{f}^{max}_{2,p}(k)\hat{d}_{2,max}(k-p+1)$
Return to *STEP 1*

The corresponding expressions for the secondary MAP stage are similar to those in (24)–(26). Complexity reduction techniques, such as the decision-feedback scheme in [29] or metric pruning [22], may be employed by this blind algorithm. The two-stage joint blind MAPSD algorithm is summarized in Table II using the T-spaced notation. Note that we could also derive a two-stage joint estimation algorithm based on the blind MLSE/VA approaches in [32], [33]. However, as discussed previously for the nonblind JMAPSD algorithm, low-delay decisions are required by the second stage so that the MAP approach is preferred (as it can yield a lower error rate for low-decision delays).

VI. COMPUTER SIMULATIONS

For binary signaling (BPSK), the primary and secondary BER's of the simulated algorithms were computed for the channel model in (2). The coefficients in Table III were used in the first set of simulations; their frequency responses are shown in Fig. 4. The optimal $M^{2L} = 2^{12} = 4096$-state JVA (JMLSE) was compared with the suboptimal two-stage

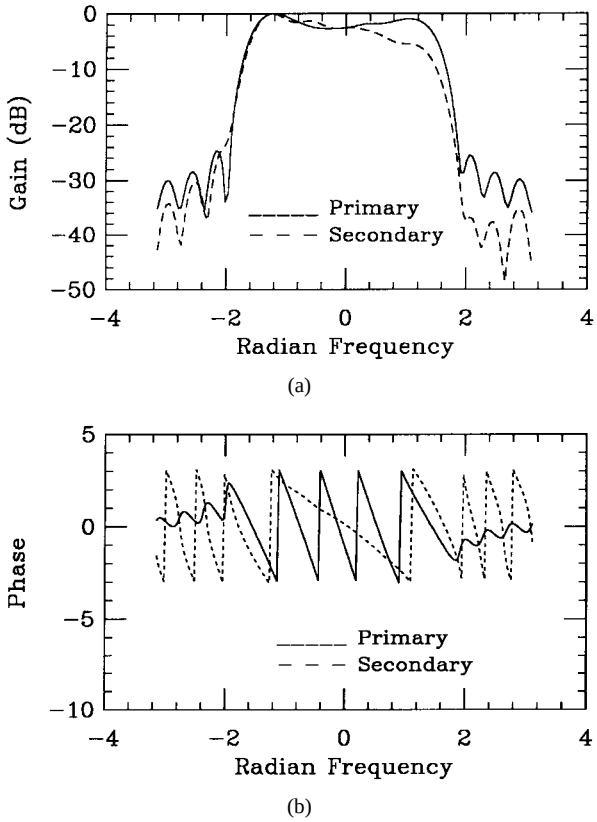

Fig. 4. $T/2$-spaced channels. (a) Magnitude response. (b) Phase response.

JMAPSD$(P_{\rm map}, P_{\rm df}, S_{\rm map}, S_{\rm df})$ algorithm where $P_{\rm map}$ and $P_{\rm df}$ refer to the number of primary (P) channel coefficients modeled by the MAP and DF sections, respectively (similar definitions apply for the secondary (S) channel).[5] (For example, JMAPSD $(5,1,4,3)$ is a $2^5 = 32$-state MAP section cascaded with a one-symbol DF filter for the primary channel, and a 16-state MAP section cascaded with a three-symbol DF filter for the secondary channel.)

The effect of the SFF at various SIR's on the two-stage JMAPSD$(6,1,6,1)$ algorithm is illustrated in Fig. 5 for $\mathrm{SNR} = 27$ dB. Notice that for $\mathrm{SIR} < 6$ dB, the SFF provides more than an order of magnitude improvement in the error rate performance. From the BER curves in Fig. 6(a) for $\mathrm{SIR} = 0$ dB, note that JMLSE provides the best error rate performance. However, observe from Fig. 6(b) ($\mathrm{SIR} = 10$ dB) that the two-stage JMAPSD algorithm (which includes the SFF) provides nearly the same performance as JMLSE (or, equivalently, the single-stage JMAPSD). The MAP/DF approach used in JMAPSD$(5,1,4,3)$ allows for even more computational savings, but at the cost of some performance degradation.

The error rate performance of a simple two-stage joint decision-feedback detector (JDFD) is also included in Fig. 6 for comparison purposes. This JDFD is a direct extension of the single-channel ideal DFE to the cochannel measurement model (i.e., there are no feedforward taps and only $L-1$ feedback taps to cancel the postcursor ISI of a channel with L coefficients). It can be seen that JMAPSD(0,6,0,6) corresponds to this JDFD structure, where the MAP section is absent and the DF section in each stage is a six-tap FIR filter. At the lower SIR, the JDFD requires about 35 dB more SNR to achieve the same BER as the JMAPSD algorithm, and at least 10 dB more at the higher SIR.

[5] As mentioned previously, we will compare the algorithms to single-stage JMLSE since, for low-delay decisions, the performance of JMLSE is similar to that of JMAPSD, and it has a lower complexity.

TABLE III
$T/2$-SPACED CHANNEL COEFFICIENTS

Primary	Secondary
$0.03687 + j0.01069$	$0.00143 + j0.01187$
$-0.04924 + j0.01239$	$0.03958 + j0.02610$
$-0.07221 - j0.00453$	$0.10806 + j0.01603$
$0.07563 - j0.03614$	$0.13480 - j0.01580$
$0.14937 - j0.02998$	$0.06418 - j0.02735$
$-0.06092 + j0.05171$	$-0.03438 + j0.00071$
$-0.26619 + j0.11931$	$-0.04526 + j0.02663$
$-0.09562 + j0.02440$	$0.01293 + j0.01065$
$0.25408 - j0.20513$	$0.02962 - j0.01447$
$0.32938 - j0.34410$	$-0.00848 - j0.00619$
$0.13014 - j0.26904$	$-0.02183 + j0.01307$
$-0.00265 - j0.11041$	$0.00265 + j0.00778$
$0.01960 - j0.02915$	$0.01088 - j0.00544$
$0.03387 - j0.01374$	$0.00000 + j0.00000$

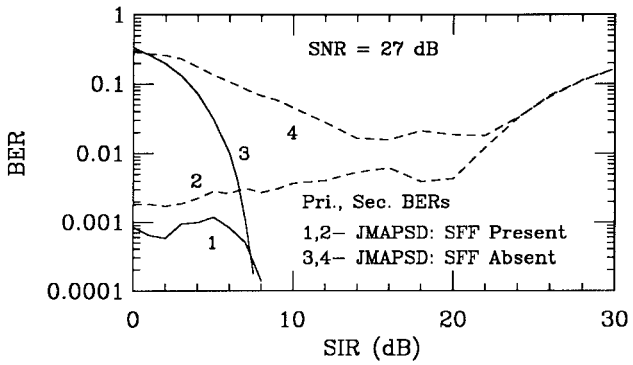

Fig. 5. Effect of SFF on the BER's of two-stage JMAPSD.

The various joint detectors were also compared for an artificial near minimum-phase channel with the coefficients in Table IV. The corresponding BER curves are shown in Fig. 7 for two values of SIR (= 0 and 10 dB). Since the channels have a span of only $4T$ seconds (i.e., eight $T/2$-spaced coefficients), $2^6 = 64$ states are required to implement JVA/JMLSE, while the two-stage JMAPSD algorithm has $2^4 = 16$ states in each of its stages. Since both the primary and secondary channels have most of their ISI contribution from the first two symbols (i.e., from the first four $T/2$-spaced coefficients), JMAPSD(2, 2, 2, 2) was also implemented. From these results, the following observations can be made. 1) For low SIR's, JMLSE yields the best BER performance; observe, however, that for SIR = 10 dB in Fig. 7(b), the BER curves of JMAPSD(4, 0, 4, 0) are virtually indistinguishable from those of JMLSE. 2) As mentioned earlier, the two-stage JMAPSD structure may suffer from uncancelled secondary power due to the symbol

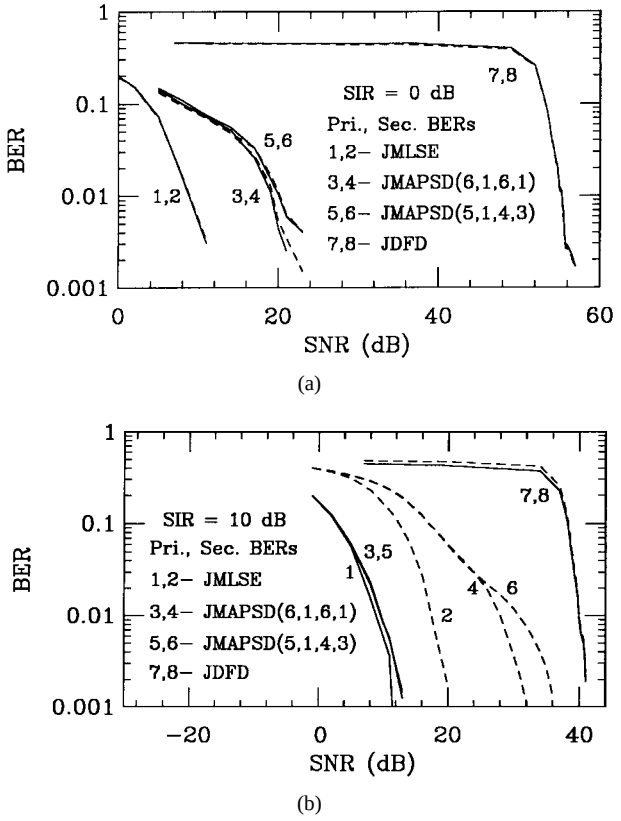

Fig. 6. BER's of the joint detectors for the channels in Table III. (a) SIR $= 0$ dB. (b) SIR $= 10$ dB.

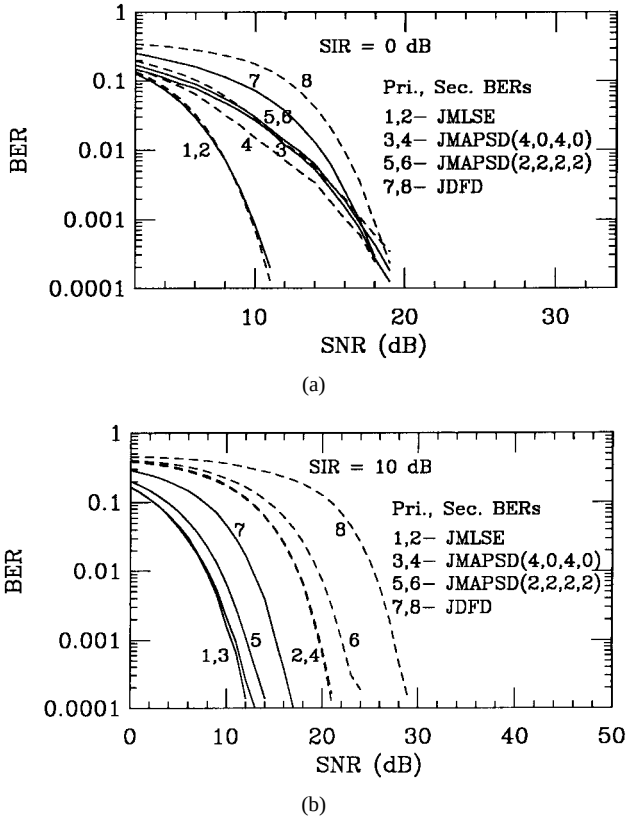

Fig. 7. BER's for the near minimum phase channels in Table IV. (a) SIR $= 0$ dB. (b) SIR $= 10$ dB.

TABLE IV
$T/2$-SPACED COEFFICIENTS OF THE NEAR MINIMUM-PHASE CHANNELS.

Primary	Secondary
$-0.35 + j0.20$	$0.10 + j0.25$
$-0.30 + j0.55$	$0.15 - j0.35$
$-0.05 - j0.60$	$-0.45 + j0.20$
$0.30 - j0.15$	$0.10 - j0.55$
$0.15 - j0.35$	$0.25 + j0.10$
$-0.20 + j0.25$	$-0.30 + j0.25$
$0.15 + j0.10$	$0.15 - j0.05$
$-0.05 + j0.05$	$0.05 + j0.10$

TABLE V
$T/2$-SPACED CHANNEL COEFFICIENTS FOR THE JBMAPSD ALGORITHM

Primary	Secondary
$0.139 - j0.068$	$0.378 - j0.175$
$0.607 - j0.284$	$0.748 - j0.372$
$0.772 - j0.427$	$0.688 - j0.442$
$0.547 - j0.420$	$0.397 - j0.364$
$0.264 - j0.275$	$0.158 - j0.170$
$0.071 - j0.066$	$0.005 + j0.010$

$d_2(k)$. This is evident in Fig. 7(a) where, at high SNR's, its performance is slightly poorer than that of the simpler JDFD [i.e., JMAPSD(0, 6, 0, 6)]. 3) The JDFD algorithm can yield adequate performance under such minimum-phase channel conditions (say, within 4–5 dB of the more complex JMLSE and JMAPSD structures).

The joint blind MAP symbol detector (JBMAPSD) was simulated for the artificial channel coefficients listed in Table V. The even and odd filter coefficients were updated using the corresponding gradients in (25). In order to reduce the misadjustment error at convergence, the step size $\mu(k)$ in (22) was allowed to decay at a rate of $\beta = 0.995$ for the primary stage, i.e., $\mu(k) = \beta^k \mu_o$ with $\mu_o = 0.5$; for the secondary stage, $\mu_o = 0.8$ and $\beta = 0.9999$. Because the algorithm was studied for reasonably large SIR's (i.e., low secondary signal powers), a higher gain was chosen for the secondary stage to improve its convergence rate.

Fig. 8(a) and (b) shows the evolution of the probability metrics of the primary and secondary MAP stages for one run of the algorithm with SNR=25 dB and SIR=15 dB. Since the channels have six $T/2$-spaced coefficients, there were $2^3 = 8$ metrics. Observe that, in the primary MAP stage, one of the metrics converges to unity in less than 100 iterations. Although the metric trajectories are noisier in the secondary MAP stage, there is still only one metric that dominates the rest. Fig. 8(c) shows the trajectories for the corresponding ensemble-averaged coefficient errors. These were generated by averaging the squared error between the actual channel $\boldsymbol{f}_m$ and the estimate $\widehat{\boldsymbol{f}}_{m,\max}(k)$ corresponding to the largest metric at each iteration [i.e., $|\boldsymbol{f}_m - \widehat{\boldsymbol{f}}_{m,\max}(k)|^2/(L_m+1)$].

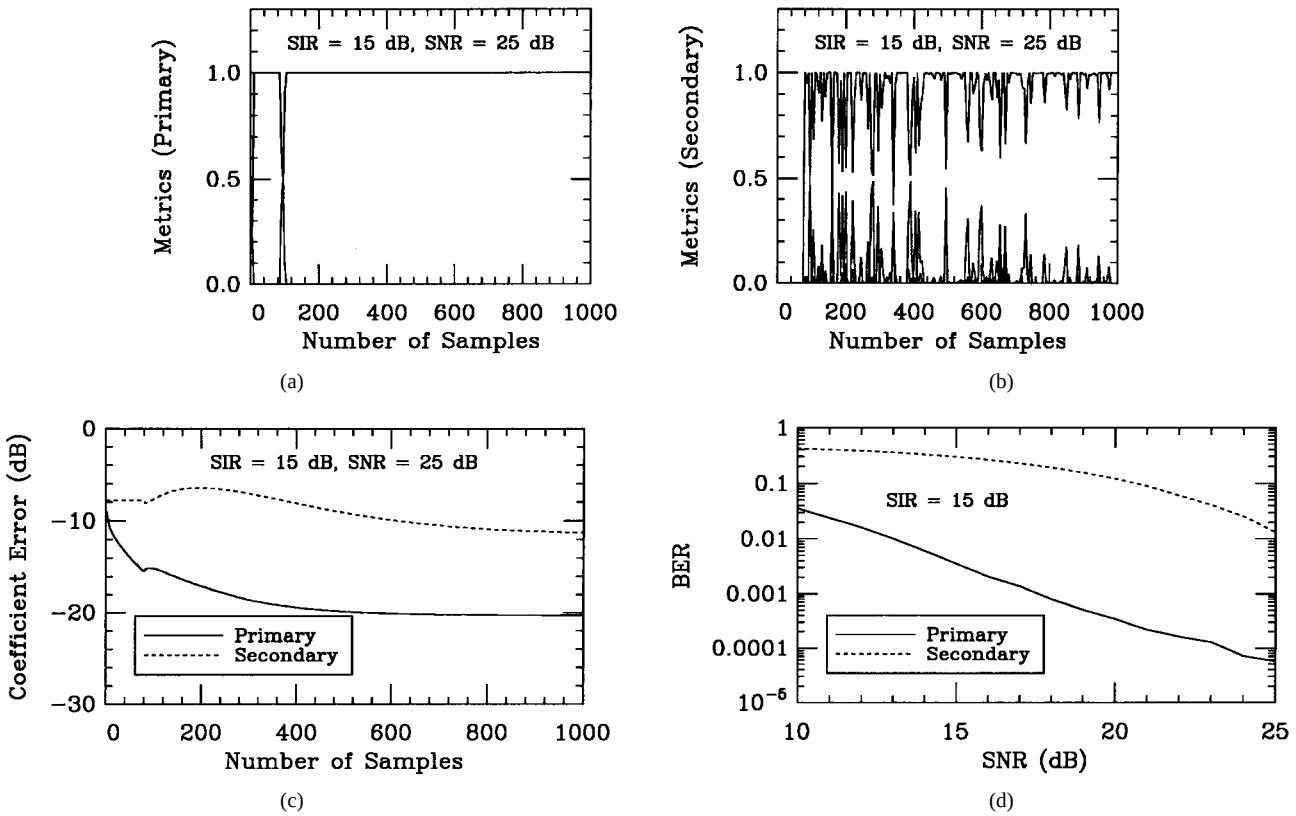

Fig. 8. Performance results of the JBMAPSD algorithm for the channels in Table V. (a) Primary metric trajectories. (b) Secondary metric trajectories. (c) Coefficient error trajectories. (d) BER curves.

Finally, BER curves are shown in Fig. 8(d) with results as expected. For higher SNR's (not shown), the curves tend to level off with a lower limit determined by the error rate achieved without additive noise. There is a difference of 12–13 dB between the primary and secondary curves. Hence, for the same relative SNR's (e.g., primary error rate at SNR=10 dB and secondary error rate at SNR=10+SIR=25 dB), the secondary stage has a slightly better performance. This is expected because the converged primary coefficients yield almost complete cancellation of the primary signal from the input to the secondary stage of the algorithm, whereas the effect of the most recent secondary symbol is not cancelled from the primary stage input. Another reason for this is that the secondary decisions are less reliable when the primary SNR $= 10$ dB since the effective SNR for the secondary signal $= 10 - 15 = -5$ dB.

In summary, we expect the performance of the two-stage JMAPSD algorithm to degrade for the following scenarios: 1) at low SNR when the SIR is close to 0 dB, and 2) when $f_{2,0}(k)$ of the secondary channel (corresponding to the weaker signal) is large relative to the other coefficients. In addition, the two-stage blind algorithm JBMAPSD will not converge satisfactorily at low SIR's, especially if the secondary channel coefficients are similar to those of the primary channel (e.g., if they are nearly scaled versions of each other).

VII. Conclusion

Nonlinear techniques for the joint estimation of narrowband cochannel signals have been presented. These techniques use a single-input receiver which directly provides Nyquist-rate samples to the detector, obviating the need for whitening matched filters. When the channel coefficients are known, single-stage JMAPSD and JMLSE are optimal techniques providing the lowest possible BER's. However, for high SIR conditions, the suboptimal two-stage JMAPSD algorithm can provide a performance approaching that of (single-stage) JMLSE, but at a much lower complexity.

We have also presented a blind adaptive algorithm for the recovery of cochannel data streams in the presence of ISI. This joint blind MAP symbol detection (JBMAPSD) algorithm also employs a single-input receiver with a two-stage structure where the first stage estimates the strongest signal and the second stage estimates the weaker signal. The corresponding channel coefficients are estimated using LMS gradient updates. By employing feedback of past decisions, the effect of cochannel interference is reduced at the input to each stage.

Simulation results for a primary signal and one cochannel interferer demonstrate the rapid convergence properties that are possible with the two-stage JBMAPSD algorithm. The BER performance curves indicate that the scheme performs well for relatively high SIR's. However, the overall performance of the blind algorithm depends on the condition that the primary channel coefficients converge, despite some residual interference from the secondary signal. Hence, the two-stage implementation may not perform well for low SIR's, whereas the single-stage JBMAPSD (or blind JMLSE) algorithm should be able to provide better performance (depending on the specific channels encountered).

In the presence of multiple cochannel signals, a multistage JMAPSD (and JBMAPSD) algorithm may be derived by a direct extension of this work. Unlike linear MMSE-based receivers, the nonlinear detectors proposed here are sensitive only to the relative signal powers of the cochannel sources. Three or more cochannel signals could be detected in the same bandwidth, provided their signal powers are not equal. However, it is likely that the effect of cascading several MAP sections will lead to a higher incidence of error propagation due to the SFF and decision feedback. Also, the corresponding single-stage algorithm for multiple cochannel signals may be computationally expensive. As a result, these algorithms may be better suited for cochannel interference mitigation in mobile radio systems employing frequency reuse where there is usually only one significant interferer, rather than in digital subscriber loop applications where there are many cochannel sources.

REFERENCES

[1] J. C. Campbell, A. J. Gibbs, and B. M. Smith, "The cyclostationary nature of crosstalk interference from digital signals in multipair cable—Part I: Fundamentals and Part II: Applications," *IEEE Trans. Commun.*, vol. COM-31, pp. 629–649, May 1983.

[2] D. C. Cox, "Universal digital portable radio communications," *Proc. IEEE*, vol. 75, pp. 436–477, Apr. 1987.

[3] T. S. Chu, "Restoring the orthogonality of two polarizations in radio communication systems, I," *Bell Syst. Tech. J.*, vol. 50, pp. 3063–3069, Nov. 1971.

[4] Special issue on digital communications by radio, *IEEE J. Select. Areas Commun.*, vol. SAC-5, Apr. 1987.

[5] B. R. Petersen and D. D. Falconer, "Minimum mean square equalization in cyclostationary and stationary interference—Analysis and subscriber loop calculations," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 931–940, Aug. 1991.

[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] W. Van Etten, "Maximum likelihood receiver for multiple channel transmission systems," *IEEE Trans. Commun.*, vol. COM-24, pp. 276–283, Feb. 1976.

[8] B. Ottersten, R. Roy, and T. Kailath, "Signal waveform estimation in sensor array processing," in *Proc. 23rd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1989, pp. 787–791.

[9] G. Xu, Y. Cho, A. Paulraj, and T. Kailath, "Maximum likelihood detection of co-channel communication signals via exploitation of spatial diversity," in *Proc. 26th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1992, pp. 1142–1146.

[10] H. E. Nichols, A. A. Giordano, and J. G. Proakis, "MLD and MSE algorithms for adaptive detection of digital signals in the presence of interchannel interference," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 563–575, Sept. 1977.

[11] K. Ishida, I. Oka, and I. Endo, "Cochannel interference cancelling effects of MSE processing and Viterbi decoding for on-board processing satellites," *IEEE Trans. Commun.*, vol. COM-34, pp. 1049–1053, Oct. 1986.

[12] Y. Bar-Ness and H. Bunin, "Co-channel interference suppression and signal separation method," in *Proc. IEEE Int. Conf. Commun.*, Philadelphia, PA, June 1988, pp. 1077–1081.

[13] R. P. Gooch and B. J. Sublett, "Demodulation of co-channel QAM signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Glasgow, Scotland, May 1989, pp. 1392–1395.

[14] A. Duel-Hallen, "Equalizers for multiple input/multiple output channels and PAM systems with cyclostationary input sequences," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 630–639, Apr. 1992.

[15] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 614–629, Apr. 1992.

[16] M. Abdulrahman and D. D. Falconer, "Cyclostationary crosstalk suppression by decision feedback equalization in digital subscriber loops," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 640–649, Apr. 1992.

[17] 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.

[18] K. Abend and B. D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, vol. 58, pp. 779–785, May 1970.

[19] K. Giridhar, S. Chari, J. J. Shynk, and R. P. Gooch, "Joint demodulation of cochannel signals using MLSE and MAPSD algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. IV, Minneapolis, MN, Apr. 1993, pp. 160–163.

[20] K. Giridhar, S. Chari, J. J. Shynk, R. P. Gooch, and D. J. Artman, "Joint estimation algorithms for cochannel signal demodulation," in *Proc. IEEE Int. Conf. Commun.*, Geneva, Switzerland, May 1993, pp. 1497–1501.

[21] S. Chen and B. Mulgrew, "Overcoming co-channel interference using an adaptive radial basis function equalizer," *Signal Processing*, vol. 28, pp. 91–107, July 1992.

[22] K. Giridhar, "Nonlinear algorithms for channel equalization and MAP symbol detection," Ph.D. dissertation, Dep. Elect. Comput. Eng., Univ. California, Santa Barbara, Sept. 1993.

[23] K. Giridhar, A. Mathur, and J. J. Shynk, "A blind adaptive MAP algorithm for the recovery of cochannel signals," in *Proc. IEEE Mil. Commun. Conf.*, Fort Monmouth, NJ, Oct. 1994, pp. 133–138.

[24] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[25] V. M. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, pp. 13–20, Jan. 1988.

[26] R. A. Iltis, J. J. Shynk, and K. Giridhar, "Recursive Bayesian algorithms for blind equalization," in *Proc. 25th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 1991, pp. 710–715.

[27] ———, "Bayesian algorithms for blind equalization using parallel adaptive filtering," *IEEE Trans. Commun.*, vol. 42, pp. 1017–1032, Feb./Mar./Apr. 1994.

[28] K. Giridhar, J. J. Shynk, and R. A. Iltis, "A modified Bayesian algorithm with decision feedback for blind adaptive equalization," in *Proc. IFAC Int. Symp. Adaptive Syst. in Contr. and Signal Processing*, Grenoble, France, July 1992, pp. 737–742.

[29] ———, "Bayesian/decision-feedback algorithm for blind adaptive equalization," *Opt. Eng.*, vol. 31, pp. 1211–1223, June 1992.

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

[31] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[32] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Commun.*, vol. 42, pp. 1000–1011, Feb./Mar./Apr. 1994.

[33] M. Ghosh and C. L. Weber, "Maximum-likelihood blind equalization," *Opt. Eng.*, vol. 31, pp. 1224–1228, June 1992.

K. Giridhar received the B.Sc. degree in applied sciences from the PSG College of Technology, Coimbatore, India, in 1985, the M.E. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, in 1989, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1993.

He was a Member of Research Staff at the Central Research Laboratory of Bharat Electronics, Bangalore, during 1989–1990, and a Research Assistant at UCSB from 1990 to 1993. During 1993–1994, he was a Post-Doctoral Research Affiliate with the Information Systems Laboratory, Stanford University. In 1994, he joined the Indian Institute of Technology, Madras, as a Senior Project Officer in a DECT-based wireless local loop project, and since 1995 has been an Assistant Professor in the Electrical Engineering Department at I.I.T. Madras, India. His present research interests are in the areas of modulation and estimation theory, and in adaptive signal processing applications for wireline and wireless communications. He is also involved in telecommunication product development activities, including fast modems, statistical speech multiplexing, and digital FM receivers.

John J. Shynk (S'78-M'86-SM'91) received the B.S. degree in systems engineering from Boston University, Boston, MA, in 1979, the M.S. degree in electrical engineering and in statistics, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1980, 1985, and 1987, respectively.

From 1979 to 1982, he was a Member of Technical Staff in the Data Communications Performance Group at AT&T Bell Laboratories, Holmdel, NJ, where he formulated performance models for voiceband data communications. He was a Research Assistant from 1982 to 1986 in the Department of Electrical Engineering at Stanford University where he worked on frequency-domain implementations of adaptive IIR filter algorithms. From 1985 to 1986, he was also an Instructor at Stanford University, teaching courses on digital signal processing and adaptive systems. Since 1987, he has been with the Department of Electrical and Computer Engineering at the University of California, Santa Barbara, where he is currently a Professor. His research interests include adaptive signal processing, adaptive beamforming, wireless communications, direction-of-arrival estimation, and neural networks. He served as an Associate Editor for adaptive filtering for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and is currently an Editor for adaptive signal processing for the *International Journal of Adaptive Control and Signal Processing*, and an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He was Technical Program Chair of the 1992 International Joint Conference on Neural Networks.

Dr. Shynk is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

Amit Mathur was born in Budayun, India, in 1964. He received the B.Tech. (Hon) degree in electrical engineering from the Indian Institute of Technology, Kharagpur, India, in 1986, the M.S. degree in electrical engineering, and the M.A. degree in mathematics from the University of Hawaii in 1990 and 1992, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1996.

He is currently employed by Lucent Technologies in Holmdel, NJ. His main interests are in mathematical analysis, and in the applications of adaptive algorithms for signal processing and communication systems.

Sujai Chari received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, in 1990 and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1991.

He spent four years at Applied Signal Technology, Sunnyvale, CA, where he extensively studied the problem of demodulating communication signals in the presence of severe cochannel interference. He developed and implemented cochannel demodulation and interference mitigation algorithms for various applications in wireless mobile communications (GSM, AMPS, D-AMPS) as well as voiceband and microwave modems. He is presently working at Amati Communications, San Jose, CA, as a systems engineer focusing on the design and implementation of DMT (discrete multitone) high performance ADSL modems.

Richard P. Gooch (S'78–A'79–M'82) received the B.S. degree from the University of Illinois at Urbana-Champaign in 1978 and the Ph.D. degree from Stanford University, Stanford, CA, in 1983, both in electrical engineering.

From 1983 to 1985 he was a Research Associate and Lecturer at Stanford University. Since 1985, he has been with Applied Signal Technology, Sunnyvale, CA, where he is head of the Advanced Techniques Department. He is involved with the development of innovative algorithms and products in the areas of modems, wireless communications, and adaptive beamforming.