# DS-CDMA System with Joint Channel Estimation and MAP Detection in Time-Selective Fading Channels

Shiauhe Tsai, *Member, IEEE*, Tan F. Wong, *Member, IEEE*, and James S. Lehnert, *Fellow, IEEE*

***Abstract*—In this paper, maximum *a posteriori* (MAP) detection is applied to a direct-sequence code-division multiple-access (DS-CDMA) system jointly with identification and estimation of time-selective fading channels. By sampling the outputs of the matched filter and combining antenna array elements, strong and time-varying multiple-access interference (MAI) is characterized and suppressed instantaneously. The decision statistics for MAP detection are obtained from the conditional probability density function of prediction error. The prediction is accomplished by approximating the fading channel with a constrained nonlinear state model. Unknown parameters such as auto-regressive (AR) process coefficients, noise covariance matrices, and the antenna array vector are estimated based on received sample vectors only. Also, differential modulation is applied to eliminate the need for pilot insertion. Through computer simulations, near-optimum bit error rates (BERs) are found.**

***Index Terms*—Antenna arrays, code-division multiple-access, Kalman filtering, MAP estimation, parameter estimation, pseudonoise coded communication, time-varying channels.**

## I. Introduction

AMONG various technologies for mobile cellular communications, direct-sequence code-division multiple-access (DS-CDMA) is well-known for its large capacity that results from several advantages [1], such as soft handoffs, a high frequency-reuse factor, and efficient use of voice activity. However, because of asynchronism in DS-CDMA systems, signals from different users cannot be kept orthogonal and multiple-access interference (MAI) arises. One solution to this problem is multi-user detection, which is covered by [2] and the references cited therein. The other is the conventional single-user detection with power control, which is adopted by the current North American DS-CDMA digital cellular system (IS-95 [3]). Here, we consider the approach of single-user detection.

In addition to the MAI, the uncertain properties of mobile radio channels pose a serious challenge for wireless DS-CDMA systems. Especially, fast-moving transceivers with a high carrier frequency can make the fading channel characteristics change rapidly over time. For example, if a 2-GHz carrier modulated by 10-kb/s data is transmitted to a mobile receiver moving at 100 km/h, the autocorrelation of the received envelope will have zero crossings every 20 symbols, according to the mobile radio channel model in [4] and [5], and other models [6], [7] for time-varying, uplink antenna array channels. In such situations, the channel can neither be assumed stationary, nor slowly varying. Hence, estimation and dynamic tracking of the channel variation becomes necessary.

To estimate and mitigate effects from fast-varying channels, one commonly used technique is pilot symbol-assisted modulation (PSAM) [8]. In PSAM systems, the receiver derives unknown channel coefficients by interpolating distortions on the pilots and then rotates decision boundaries accordingly. Analytical bit-error-rate (BER) expressions for PSAM were given in [8] for different SNRs and Doppler spreads. In addition to PSAM, a linear predictive receiver, which jointly estimates the fading channel and detects the data symbol, was proposed in [9] for a continuous phase modulation (CPM) signal over frequency-flat Rayleigh fading channels. The basic idea of [9] has been extended to numerous narrowband systems. For single-user spread-spectrum systems, a joint channel state estimator and maximum likelihood (ML) detector with code delay estimation was developed in [10]. The channel estimation was based on a standard Kalman filter (KF) with given parameters, and the pseudonoise (PN) code delay estimation was based on an extended Kalman filter (EKF) because of the nonlinearity in the observed output. In the area of multiuser detection, synchronous CDMA systems using a standard KF to track fast channel variations were proposed in [11]–[13]. Given crosscorrelations among all users, the receiver first performs decorrelation under different symbol hypotheses. The likelihood of hypotheses is then computed from the prediction error and its covariance provided by the standard KF, where parameters of the associated linear state model were known *a priori*.

A decision-directed MAP detection, assuming known channel statistics, was proposed in [14] for the single-user receiver of CDMA systems with aperiodic random sequences.[1] Unlike the multiuser receiver, crosscorrelations among users are random variables, and the MAI must be characterized statistically. While the MAI is suppressed through time-domain noise whitening by oversampling the sequence matched filter (MF) output [14], the frequency-flat, fast-varying fading

Manuscript received February 14, 2000; revised November 1, 2000. This work was supported in part by the US government DARPA GloMo Project AO F383, AFRL Contract F30602–97–C-0314. This paper was presented in part at the IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, Sept. 21–24, 1999.

S. Tsai is with Ericsson Wireless Communications, San Diego, CA 92121 USA.

T. F. Wong is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA.

J. S. Lehnert is with the Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-1285 USA.

Publisher Item Identifier S 0733-8716(01)00962-3.

[1]The IS-95 system uses an m-sequence with an extremely long period to approximate a random sequence. Some reasons for adopting random signature sequences can be found in [15]

channel is estimated by a moving average (MA) linear predictor with pilots to avoid cycle slipping. Results in [14] have shown a BER that is close to ideal coherent BPSK. However, the MA parameters were assumed from *a priori* spectral analysis, causing a major drawback for the method in [14]. In this paper, no *a priori* knowledge of the channel statistics is assumed, and regardless of strong interference, information about the fast-varying, antenna-array channel is derived only from received samples. Given a training sequence, the proposed algorithm only assumes knowledge of the spreading sequence and timing of the desired user. The real-time channel identification is based on a complex state model with undetermined parameters. Such modeling can approximate nonsymmetric Doppler spectra and more general time-varying characteristics. Because nonlinearities exist in both state and measurement equations, the principle of the EKF is applied as part of the solution to the problem of real-time identification. In addition, considering two quantities not included in the EKF (the noise variance and the phase array constraints) a complete solution is proposed without significantly increasing the computational complexity. As a result, the estimation of average signal and noise power is integrated to the identification algorithm such that no SNR information is required. At the same time, MAI suppression similar to minimum mean-square-error (mmse) beamforming is accomplished through the combination of MAI characterization and array vector estimation. Furthermore, using differentially coherent detection, pilot symbols are removed to increase the spectral efficiency. Simulation results on BER are compared with the ideal and the fixed-design PSAM in an MAI-free channel as benchmarks. The performance degradation due to estimation error and the removal of pilots is found to be small.

This paper is organized as follows. In Section II, the block diagram of a DS-CDMA system and the corresponding continuous-time and discrete-time models are established. Using the discrete-time model, the maximum *a posteriori* (MAP) detection is derived in Section III. A joint channel estimation and MAP detection scheme is then developed in Section IV with detailed discussions on parameter identification. Simulation results for the system performance are presented in Section V, and conclusions are drawn in Section VI.

## II. System Model for Aperiodic DS-CDMA

The block diagram of the aperiodic DS-CDMA system with an antenna array is shown in Fig. 1(a). In the remaining part of this section, we describe the continuous-time baseband model and the discrete-time receiver output model, where the former reflects practical system operation, and the latter accommodates digital processing.

### A. Continuous-Time Model

The continuous-time baseband system model is illustrated in Fig. 1(b). The modulation by a carrier is not shown in Fig. 1(b) because the phase error is incorporated into the complex channel coefficient. There are $K+1$ users transmitting to the receiver with the desired user denoted as the 0th user. At each time interval $[lT,(l+1)T]$, the desired user generates a symbol $d_l$ from

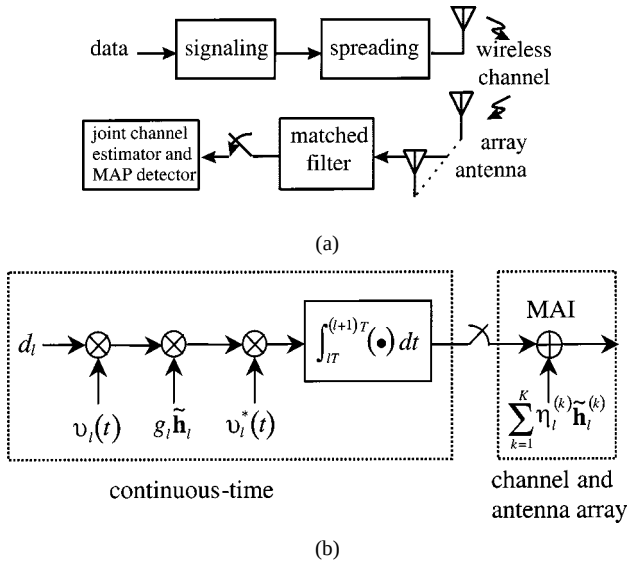



Fig. 1. System model. (a) Block diagram of CDMA system with antenna array. (b) Continuous-time and discrete-time model.

a BPSK (or QPSK with minor changes) symbol set $\mathcal{S}$ with cardinality $|\mathcal{S}|$. The Cartesian product of $n$ symbols is denoted by $\mathcal{S}^n$. Each possible symbol from a particular signal set is sent with equal probability. The input information bit $b_l$ is related to the transmitted symbols by $b_l = f(d_l, d_{l+1}) = (1 - d_l d_{l+1})/2$ for differentially encoded BPSK, and $b_l = f(d_l, d_{l+1}) = (1-d_l)/2$ for coherent BPSK. The symbol is spread by the waveform

$$\upsilon_l(t) = \sum_{n=0}^{N-1} \frac{1}{\sqrt{N}} c_{lN+n}\psi(t - lT - nT_c)$$

where $\{c_n\}$ is the quadriphase random signature sequence with $|c_n|^2 = 1$, $\psi(t)$ is the time-limited chip waveform with duration $[0, T_c)$, and $\int_0^{T_c} |\psi(t)|^2\, dt = 1$. The spreading gain is denoted by $N$, and the spreading waveform is normalized such that $\int_{lT}^{(l+1)T} |\upsilon_l(t)|^2\, dt = 1$. Likewise, the spreading waveform of each interfering user $k$ is given by

$$\upsilon_l^{(k)}(t) = \sum_{n=0}^{N-1} \frac{1}{\sqrt{N}} c_{lN+n}^{(k)}\psi(t - lT - nT_c - \tau^{(k)})$$

where $\tau^{(k)} \in [0, T_c)$ is the relative chip delay of the $k$th user. In the following, all quantities involving interfering users are denoted with the superscript $^{(k)}$, where $k \in \{1, \ldots, K\}$, and those involving the desired user are denoted without the superscript.

After spreading, each user transmits to the receiver through a wireless channel independently from the other users. In the following, we assume a single-path, time-selective fading channel for simplicity. This assumption does not preclude the applicability of the techniques developed here to multipath fading channels. Note that if multiple paths are mutually independent and resolvable by a Rake receiver [1], we can use the algorithm by concatenating the received vectors at different Rake fingers to form a single large vector. This generalization to multipath channels is straightforward. The single-path channel is equivalent to a time-varying, multiplicative distortion. For the example in Section I, the autocorrelation can be greater than 0.997 within one symbol duration, although there are zero-crossings every 20

symbols. Therefore, the distortion on each symbol can be simplified to the multiplication by a complex constant $g_l$. In Section IV, we show that the channel effects on interfering users can be estimated symbol-by-symbol, regardless of whether they are time-varying or stationary.

The signal at the receiver in complex representation is given by

$$\mathbf{r}(t)=\sum_{l=0}^{\infty} g_l d_l v_l(t)\mathbf{h}_l+\sum_{k=1}^{K}\sum_{l=-\infty}^{\infty} g_l^{(k)} v_l^{(k)}(t)\mathbf{h}_l^{(k)}+\mathbf{u}(t) \tag{1}$$

in which

$$\mathbf{u}(t)=[u_1(t) \quad \cdots \quad u_M(t)]^T$$

where $M$ is the number of array elements, and $\{u_i(t)\}$ are i.i.d. complex additive white Gaussian noise (AWGN) processes with in-phase and quadrature power spectral densities equal to $N_0/2$. In addition, the array vector is given by

$$\mathbf{h}_l=[1 \quad e^{i2\pi(\Delta/\lambda)\sin\theta} \quad \cdots \quad e^{i2\pi(M-1)(\Delta/\lambda)\sin\theta}]^T$$

where $\lambda$, $\theta$, and $\Delta$ are the carrier wavelength, the incidence angle, and the antenna separation, respectively. For the beamforming purpose, we assume a linear-phase array with closely spaced elements. Notice that interferers' data symbols are omitted in (1) because they are multiplied by random spreading sequences, and hence, the statistical properties of the products do not depend on the data.

### *B. Discrete-Time Model*

The received signal is processed as depicted in Fig. 1. The receiver performs the matched filtering during the $l$th symbol interval and samples the output at time $(l+1)T$ to form a statistic, given by

$$\mathbf{r}_l=\int_{lT}^{(l+1)T}\mathbf{r}(t)v_l^*(t)\,dt.$$

The statistic $\mathbf{r}_l$ can be expressed as

$$\mathbf{r}_l=d_l g_l\mathbf{h}_l+\sum_{k=1}^{K} g_l^{(k)}\,\eta_l^{(k)}\mathbf{h}_l^{(k)}+\mathbf{u}_l \tag{2}$$

where the random vector $\mathbf{u}_l$ accounts for the effect of the complex AWGN with mean $\mathbf{0}$ and covariance $N_0\mathbf{I}_M$. The sum of random vectors $\sum_{k=1}^{K} g_l^{(k)}\eta_l^{(k)}\mathbf{h}_l^{(k)}$ accounts for effects from the total MAI, in which the $k$th interfering user contributes

$$\eta_l^{(k)}=\left\{\sum_{n=0}^{N-1} c_{lN+n}^{(k)}\,c_{lN+n}^{*}\hat{R}(\tau^{(k)})+\sum_{n=0}^{N-1} c_{lN+n+1}^{(k)}\,c_{lN+n}^{*}R(\tau^{(k)})\right\}$$

where for $0\le\tau<T_c$

$$\hat{R}(\tau)=\int_{\tau}^{T_c}\psi(t-\tau)\psi^*(t)\,dt \quad \text{and}$$

$$R(\tau)=\int_{0}^{\tau}\psi(t+T_c-\tau)\psi^*(t)\,dt.$$

The parameter $g_l^{(k)}$ denotes the approximated channel coefficient for the $k$th user during the time interval $[lT,(l+1)T]$. For notational convenience, we define

$$\mathbf{G}_l=\{g_l^{(1)},\ldots,g_l^{(K)}\}.$$

The set of channel coefficients $\mathbf{G}_l$ affecting interfering users makes the MAI characterization different from the standard Gaussian approximation. For example, interfering signals might also pass through fast-varying fading channels, and the overall interference of each user $g_l^{(k)}\eta_l^{(k)}$ becomes the product of two random variables. In [11]–[13], the term $\eta_l^{(k)}$ is given by symbol hypotheses and known crosscorrelations so that the observations of all users can be simplified as jointly Gaussian. This is not achievable for a single-user receiver because crosscorrelations among different users are random variables, which can only be characterized statistically. In contrast to the multiuser case, if we condition on the set of channel coefficients $\mathbf{G}_l$, the linear combination of $g_l^{(k)}\eta_l^{(k)}$ can be approximated as conditionally Gaussian. Mathematically rigorous arguments can be found in [16] and [17]: if the set of channel coefficients $\mathbf{G}_l$, the ratio $K/N$, and the delays $\{\tau^{(k)}\}$ of interfering users are given, then the MAI of the MF output is asymptotically zero-mean Gaussian as $N\to\infty$. Given that $N$ is large enough, the MAI-plus-AWGN vector at each sampling time $(l+1)T$ can be characterized as

$$\mathbf{n}_l=\mathbf{u}_l+\sum_{k=1}^{K} g_l^{(k)}\eta_l^{(k)}\mathbf{h}_l^{(k)} \tag{3}$$

where the random vector $\mathbf{n}_l$ under the given condition $\mathbf{G}_l$ is denoted as

$$\mathbf{n}_l|\mathbf{G}_l\sim N(\mathbf{0}_M,\mathbf{U}_l)$$

and

$$\mathbf{U}_l=N_0\mathbf{I}_M+E\left\{\sum_{k=1}^{K}|g_l^{(k)}\eta_l^{(k)}|^2\mathbf{h}_l^{(k)}\mathbf{h}_l^{(k)^H}\Big|\mathbf{G}_l\right\} \tag{4}$$

where $\mathbf{I}_M$ denotes the $M\times M$ identity matrix. If the covariance matrix $\mathbf{U}_l$ is estimated during the period of time when the number of users and channel conditions remain unchanged, the conditions required above can be satisfied. The implementation of this conditional Gaussian approximation (CGA) is described in Sections III and IV.

## III. MAP Detection

In terms of minimizing error probability, MAP detection is an optimal criterion for symbol-by-symbol detection [18]. By characterizing probabilities of each symbol, MAP detection also delivers reliability information about its decision. Strictly speaking, the optimum MAP detector requires all observations as sufficient statistics. However, to meet the practical restriction on complexity, MAP detection is usually achieved by using the up-to-date observation with a fixed delay. The derivation of the coherent $D$-lag MAP algorithm is given in [14]. The $D$-lag differentially coherent MAP detection rule is

$$\hat{b}_{l-D}=\arg\max_{b_{l-D}}\Pr\left(b_{l-D}|\mathbf{R}_l,\boldsymbol{\Gamma}_l\right) \tag{5}$$

where $\boldsymbol{\Gamma}_l = \{\mathbf{G}_0, \mathbf{G}_1, \ldots, \mathbf{G}_l\}$ and $\mathbf{R}_l = \{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_l\}$ denote interferers' channel coefficients and MF output vectors up to time $l$, respectively. For the following derivation, some sequences of symbol hypotheses are denoted as

$$\begin{aligned}\mathbf{d}_l &= \{d_l, d_{l-1}, \ldots, d_0\}\\ \mathbf{d}_l(D) &= \{d_l, d_{l-1}, \ldots, d_{l-D}\}\\ \overline{\mathbf{d}}_l &= \{d_{l-D+1}, d_{l-D}\}, \quad \text{and}\\ \mathbf{d}'_l &= \{\mathbf{d}_l(D-2), \mathbf{d}_{l-D-1}\}.\end{aligned}$$

Then, according to Bayes' rule, the probability mass function (pmf) in (5) can be written as

$$\begin{aligned}&\Pr\,(b_{l-D}|\mathbf{R}_l, \boldsymbol{\Gamma}_l)\\ &\quad= \frac{\sum_{f(\overline{\mathbf{d}}_l)=b_{l-D}} \sum_{\mathbf{d}'_l \in \mathcal{S}^{l-1}} p(\overline{\mathbf{d}}_l, \mathbf{d}'_l, \mathbf{R}_l, \boldsymbol{\Gamma}_l)}{p(\mathbf{R}_l, \boldsymbol{\Gamma}_l)}\\ &\quad= \frac{\sum_{f(\overline{\mathbf{d}}_l)=b_{l-D}} \sum_{\mathbf{d}'_l \in \mathcal{S}^{l-1}} p(\mathbf{R}_l|\overline{\mathbf{d}}_l, \mathbf{d}'_l, \boldsymbol{\Gamma}_l)\,\Pr\,(\overline{\mathbf{d}}_l, \mathbf{d}'_l)}{p(\mathbf{R}_l|\boldsymbol{\Gamma}_l)}\end{aligned}$$

in which $p(\mathbf{R}_l|\boldsymbol{\Gamma}_l)$ is independent of the information bit hypothesis $b_{l-D}$, and $\Pr\,(\overline{\mathbf{d}}_l, \mathbf{d}'_l)$ is constant over all hypotheses.[2] As a result, the MAP detection rule is equivalent to the maximization problem

$$\hat{b}_{l-D} = \arg\,\max_{b_{l-D}} \sum_{f(\overline{\mathbf{d}}_l)=b_{l-D}} \sum_{\mathbf{d}'_l \in \mathcal{S}^{l-1}} p(\mathbf{R}_l|\overline{\mathbf{d}}_l, \mathbf{d}'_l, \boldsymbol{\Gamma}_l). \tag{6}$$

Nevertheless, the number of terms summed in (6) increases exponentially with time. To reduce the complexity, we adopt the ML sequence feedback technique stated next. Suppose the actual symbol sequence $\tilde{\mathbf{d}}_{l-D-1}$ is given, then (6) becomes

$$\begin{aligned}\hat{b}_{l-D} = \arg\,\max_{b_{l-D}} &\sum_{f(\overline{\mathbf{d}}_l)=b_{l-D}}\\ &\cdot \sum_{\mathbf{d}_l(D-2) \in \mathcal{S}^{D-1}} p(\mathbf{R}_l|\mathbf{d}_l(D), \tilde{\mathbf{d}}_{l-D-1}, \boldsymbol{\Gamma}_l).\end{aligned} \tag{7}$$

As shown above, (7) only sums over the past $D+1$ symbol hypotheses $\mathbf{d}_l(D) \in \mathcal{S}^{D+1}$; hence, the complexity of the MAP detection in (6) is reduced and fixed to $|\mathcal{S}|^{D+1}$. In practice, we approximate $\tilde{\mathbf{d}}_{l-D-1}$ by feeding back the ML symbol sequences $\hat{\mathbf{d}}_{l-D-1}$, leading to the same $\mathbf{d}_l(D-1)$. Moreover, with these feedback sequences, (7) can be recursively computed by the following steps:

Step 1) At time $l$, update the conditional probabilities of all hypotheses. That is, $\forall\ \mathbf{d}_l(D) \in \mathcal{S}^{D+1}$

$$\begin{aligned}&p(\mathbf{R}_l|\mathbf{d}_l(D), \hat{\mathbf{d}}_{l-D-1}, \boldsymbol{\Gamma}_l)\\ &\quad= p(\mathbf{r}_l|\mathbf{R}_{l-1}, \mathbf{d}_l(D), \hat{\mathbf{d}}_{l-D-1}, \mathbf{G}_l)\cdot\\ &\qquad p(\mathbf{R}_{l-1}|\mathbf{d}_{l-1}(D-1), \hat{\mathbf{d}}_{l-D-1}, \boldsymbol{\Gamma}_{l-1}).\end{aligned} \tag{8}$$

[2]Definitions of $f(\bullet)$, $\mathcal{S}^{l-1}$, and $|\mathcal{S}|$ can be found in Section II-A.

Step 2) Make the $D$-lag MAP decision according to (7), and select the survival set of conditional probabilities. $\forall\ \mathbf{d}_l(D-1) \in \mathcal{S}^D$

$$\begin{aligned}&p(\mathbf{R}_l|\mathbf{d}_l(D-1), \hat{\mathbf{d}}_{l-D}, \boldsymbol{\Gamma}_l)\\ &\quad\leftarrow \max_{d_{l-D} \in \mathcal{S}} p(\mathbf{R}_l|\mathbf{d}_l(D-1), d_{l-D}, \hat{\mathbf{d}}_{l-D-1}, \boldsymbol{\Gamma}_l).\end{aligned} \tag{9}$$

The righthand side of (9) indicates that the feedback is not based on the decision of $b_{l-D}$ but the likelihood that symbol $d_{l-D}$ leads to the same state $\mathbf{d}_l(D-1)$. The selection of the survival sequence in the proposed algorithm is similar to that in the suboptimal soft-output algorithm [19] or the max-log-MAP algorithm [20]. What remains is the computation of the conditional probability density function (pdf)

$$p(\mathbf{r}_l|\mathbf{R}_{l-1}, \mathbf{d}_l(D), \hat{\mathbf{d}}_{l-D-1}, \mathbf{G}_l) \tag{10}$$

for the MAP detection.

From (2) and (3) in Section II and the pdf in (10), the conditional characterization of the receiver output random vector can be decomposed into two parts: characterizing $d_l g_l \mathbf{h}_l$ given $\{\mathbf{R}_l, \mathbf{d}(D), \hat{\mathbf{d}}_{l-D-1}\}$ and characterizing $\mathbf{r}_l$ given $\mathbf{G}_l$. By definition, $d_l g_l \mathbf{h}_l$ is independent of $\mathbf{G}_l$, and if the current interferers' channel coefficients are given, $\mathbf{n}_l$ is independent of the desired user's symbols and past received vectors. From (3), we have already characterized $\mathbf{n}_l$ using the CGA. The term, $d_l g_l \mathbf{h}_l$, however, needs further interpretation. First, when conditioned on each symbol hypothesis $\mathbf{d}_l(D)$, the transmitted symbol $d_l$ can be viewed as given. Second, the array vector can be treated as a set of unknown constants because of its slow variation. For example, if the transmitter is 300 m from the base station (BS) and is moving at 100 km/hr, the incidence angle will change at most 0.5° for 0.1 s (1000 symbols at 10 kb/s). Because of the angular spread, the separation between antenna elements must be kept small to avoid any reduction in correlation. Since the antenna array is used for the purpose of beamforming, nonperfectly correlated array elements will significantly reduce dimensions that are available to form directional nulls toward interference. To ensure nearly perfect correlations among array elements, the array size $(M-1)\Delta$ should not exceed $\lambda$ if the angular spread is 1° [21]. For the above example, the maximum change of phase difference among array elements will be $2\pi \sin(0.5°) = 5.4831 \times 10^{-2}$ for 1000 symbols. Therefore, $\mathbf{h}_l$ is assumed to be quasistationary. However, it is generally more difficult to ascertain the channel coefficient $g_l$ for fast fading. To make a systematic channel characterization possible, we compromise between accuracy and tractability by using a model-based estimation. In fact, the goal of such modeling is not to achieve an exact channel description, but instead to achieve a good system performance with a reasonable amount of computation. Furthermore, because $\mathbf{r}_l$ and $\mathbf{R}_l$ are jointly Gaussian conditioned on $\mathbf{G}_l$, the pdf in (10) is completely characterized by the conditional mean

$$\begin{aligned}\hat{\mathbf{r}}_{l|l-1} &= E\{\mathbf{r}_l|\mathbf{R}_{l-1}, \mathbf{d}_l(D), \hat{\mathbf{d}}_{l-D-1}, \mathbf{G}_l\}\\ &= d_l E\{g_l \mathbf{h}_l|\mathbf{R}_{l-1}, \mathbf{d}_l(D), \hat{\mathbf{d}}_{l-D-1}\}\end{aligned} \tag{11}$$

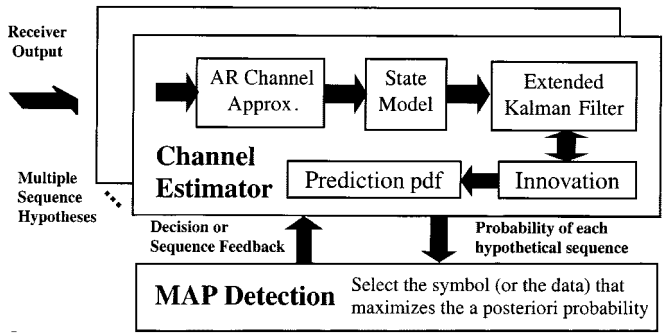



Fig. 2. Joint channel estimation and MAP detection.

and the conditional covariance

$$\boldsymbol{\Sigma}_{l|l-1} = E\{(\mathbf{r}_l - \hat{\mathbf{r}}_{l|l-1})(\mathbf{r}_l - \hat{\mathbf{r}}_{l|l-1})^H \cdot \Big| \mathbf{R}_{l-1}, \mathbf{d}_l(D), \hat{\mathbf{d}}_{l-D-1}, \mathbf{G}_l\}. \tag{12}$$

As shown in (11), a Bayesian channel estimator [22] is required to characterize the pdf in (10).

In summary, the computation of the pdf in (10) involves the following two parts of the channel estimation. First, since $\mathbf{r}_l$ and $\mathbf{R}_{l-1}$ are jointly Gaussian given $\mathbf{d}_l$ and $\mathbf{G}_l$, the conditional mean in (11) is equal to the mmse prediction of $\mathbf{r}_l$. Second, by (11), the conditional covariance is equivalent to the error covariance of the mmse predictor. Since the set $\mathbf{G}_l$ is unknown, an estimator of the error covariance in (12) is necessary. In the next section, estimators for these two parts are established. For notational convenience, we omit the given conditions in the following and simply write the three quantities associated with the pdf in (10) as $\mathbf{r}_{l|l-1}$, $g_{l|l-1}\mathbf{h}_{l|l-1}$, and $\mathbf{n}_l$, respectively.

## IV. Joint Channel Estimation and MAP Detection

Joint channel estimation and MAP detection is outlined in Fig. 2. The recursive computation of the pdfs needed in MAP symbol detection has already been shown. We show in this section how to implement the estimator required in Section III. The goal is to remove the assumption on *a priori* channel knowledge and to implement the MAP detection with a real-time model identification. To achieve this goal, joint channel estimation using an EKF [9], [10][3] is investigated in this section.

### A. Auto-Regressive Channel Approximation and Nonlinear State Model

The fading channel coefficients $\{g_l\}$ are modeled by a $p$th order auto-regressive (AR) process. Let the first element of the channel state vector be

$$g_l = \sum_{i=1}^{p} a_{i,l-1} g_{l-i} + w_l = \mathbf{a}_{l-1}^T \mathbf{g}_{l-1} + w_l \tag{13}$$

[3]The use of Kalman filtering for channel estimation and ML sequence detection was first suggested, although not applied, in [9]. Kalman filtering was not applied because some recursion quantities for state vectors, such as conditional mean and estimation error covariance, were redundant in the problem formulated in [9]. Roughly at the same time, the use of the EKF with joint ML detection was published in [10].

where $\mathbf{g}_l = [g_l, \ldots, g_{l-p+1}]^T$ denotes the channel state vector at time $l$. We assume the initial condition of the channel $g_0$ is Gaussian distributed and independent of the zero-mean white Gaussian model noise $w_l$ with variance $\sigma_{w_l}^2$. The vector $\mathbf{a}_l = [a_{1,l}, \ldots, a_{p,l}]^T$ represents undetermined complex AR model coefficients. Such modeling and its identification algorithm developed here extend the work in [10]–[13] to the current scenario. In the previous work, the *a priori* given real AR coefficient can only apply to cases where the Doppler spread is fixed and the $z$-transform of the channel has a single real pole or a complex conjugate pair of poles. This limits the application to strictly time-invariant, symmetric Doppler spectra. In reality, empirical data support nonsymmetric Doppler spectra, such as COST1, COST2, or one-sided exponential spectrum [23], [24]. The complex AR-channel modeling and the adaptive estimation of coefficients can accommodate more general spectra with slowly varying Doppler spread. The ability to track slowly varying Doppler spread is shown in the simulation.

Combining (13) and (2), the time-varying fading process and the received array vectors can be described by a nonlinear state model with the state equation

$$\begin{aligned}\mathbf{x}_l &= \begin{bmatrix} \mathbf{g}_l \\ \mathbf{a}_l \\ \mathbf{h}_l \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{l-1}\mathbf{g}_{l-1} \\ \mathbf{a}_{l-1} \\ \mathbf{h}_{l-1} \end{bmatrix} + w_l \begin{bmatrix} 1 \\ \mathbf{0}_{2p+M-1} \end{bmatrix}_{(2p+M)\times 1} \\ &= \mathbf{F}(\mathbf{x}_{l-1}) + w_l \mathbf{e}_1 \end{aligned}$$

where

$$\mathbf{F}_{l-1} = \left[ \frac{\mathbf{a}_{l-1}^T}{\mathbf{I}_{p-1} | \mathbf{0}_{p-1}} \right]_{p\times p}$$

and $\mathbf{O}_p$ denotes the $p \times 1$ zero-vector. The measurement equation is given by

$$\begin{aligned}\mathbf{r}_l &= d_l[\mathbf{h}_l \quad \mathbf{O}_{M\times(p-1)}]\mathbf{g}_l + \mathbf{n}_l \\ &= d_l \mathbf{H}(\mathbf{x}_l) + \mathbf{n}_l \end{aligned}$$

with the nonlinear, nondifferentiable constraint

$$\begin{aligned}\mathbf{h}_l &= [h_{0,l} \quad h_{1,l} \quad \cdots \quad h_{M-1,l}]^T \\ h_{m,l} &= h_{1,l}^m, \quad \text{for } m = 0, \ldots, M-1, \quad \text{and} \\ |h_{0,l}| &= |h_{1,l}| = \cdots = |h_{M-1,l}| = 1. \end{aligned} \tag{14}$$

To estimate the unknown model parameters, the principle of EKF is applied as part of the solution.

### B. Linear Perturbation Model and EKF

We define the vector operator by

$$\nabla_{\mathbf{x}} = \left[ \frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \cdots \quad \frac{\partial}{\partial x_{2p+M}} \right]$$

and define the first-order derivatives by

$$\boldsymbol{\Phi}_l = \nabla_{\mathbf{x}} \mathbf{F}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_l} = \left[ \begin{array}{c|c|c} \mathbf{F}_l & \frac{\mathbf{g}_l^T}{\mathbf{O}_{p-1\times p}} & \mathbf{O}_{p\times M} \\ \hline \mathbf{O}_{p\times p} & \mathbf{I}_p & \mathbf{O}_{p\times M} \\ \hline \mathbf{O}_{M\times p} & \mathbf{O}_{M\times p} & \mathbf{I}_M \end{array} \right]$$

and

$$\boldsymbol{\Psi}_l = \nabla_{\mathbf{x}}\mathbf{H}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_l} = [\mathbf{h}_l|\mathbf{O}_{M\times(2p-1)}|g_l\mathbf{I}_M]$$

where $\mathbf{O}$ is the zero-matrix. By Taylor's expansion, if the second-order error is negligible, the extended state model can be approximated by the linear perturbation model

$$\begin{aligned}\Delta\mathbf{x}_{l|l-1} &= (\mathbf{x}_l - \hat{\mathbf{x}}_{l|l-1}) \\ &\approx \nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x})|_{\mathbf{x}=\hat{\mathbf{x}}_{l-1|l-1}}(\mathbf{x}_{l-1} - \hat{\mathbf{x}}_{l-1|l-1}) + w_l\mathbf{e}_1 \\ &= \hat{\boldsymbol{\Phi}}_{l-1|l-1}\Delta\mathbf{x}_{l-1|l-1} + w_l\mathbf{e}_1 \end{aligned} \tag{15}$$

and

$$\begin{aligned}\boldsymbol{\varepsilon}_{l|l-1} &= \mathbf{r}_l - \hat{\mathbf{r}}_{l|l-1} \\ &\approx d_l\nabla_{\mathbf{x}}\mathbf{H}(\mathbf{x})|_{\mathbf{x}=\hat{\mathbf{x}}_{l|l-1}}(\mathbf{x}_l - \hat{\mathbf{x}}_{l|l-1}) + \mathbf{n}_l \\ &= d_l\hat{\boldsymbol{\Psi}}_{l|l-1}\Delta\mathbf{x}_{l|l-1} + \mathbf{n}_l. \end{aligned} \tag{16}$$

By a recursive mmse procedure similar to the standard KF, the EKF algorithm is obtained as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{l|l-1} &= \mathbf{F}(\hat{\mathbf{x}}_{l-1|l-1}) \\ \boldsymbol{\Sigma}_{l|l-1} &= \hat{\boldsymbol{\Psi}}_{l|l-1}\mathbf{P}_{l|l-1}\hat{\boldsymbol{\Psi}}^H_{l|l-1} + \mathbf{U}_l \\ \mathbf{K}_l &= d_l\mathbf{P}_{l|l-1}\hat{\boldsymbol{\Psi}}^H_{l|l-1}\boldsymbol{\Sigma}^{-1}_{l|l-1} \end{aligned} \tag{17}$$

$$\begin{aligned}\boldsymbol{\varepsilon}_{l|l} &= \mathbf{r}_l - d_l\mathbf{H}(\hat{\mathbf{x}}_{l|l-1}) = \mathbf{r}_l - \hat{\mathbf{r}}_{l|l-1} \quad \text{and} \\ \hat{\mathbf{x}}_{l|l} &= \hat{\mathbf{x}}_{l|l-1} + \mathbf{K}_l\boldsymbol{\varepsilon}_{l|l} \end{aligned} \tag{18}$$

where $\mathbf{P}_{l|l-1}$ is the estimation error covariance with a computation provided later.

### C. Estimation of Conditional MAI-plus-AWGN Covariance $\mathbf{U}_l$

In the expression for the Kalman gain matrix, the conditional MAI-plus-AWGN covariance $\mathbf{U}_l$ plays a critical role in MAI suppression. Since the Kalman gain matrix is actually a Wiener solution to the estimation of $\Delta\mathbf{x}_{l|l-1}$ from $\boldsymbol{\varepsilon}_{l|l}$, the inclusion of $\mathbf{U}_l$ in the matrix inversion acts as an mmse notch filter to the interference. In the context of an antenna array, this is equivalent to nulling out the strong directional interference by mmse beamforming. Based on oversampling the output of the matched filter, an estimation scheme for $\mathbf{U}_l$ has been proposed in [25] for slowly varying fading channels. From the characterization (4), the time-selective fading channel requires that the estimation of $\mathbf{U}_l = E\{\mathbf{n}_l\mathbf{n}_l^H|\mathbf{G}_l\}$ be done within the snapshot of each symbol time. This can be accomplished by taking $J$ samples between the two MF output peaks $\mathbf{r}_l$ and $\mathbf{r}_{l+1}$, as shown in Fig. 3. The vector $\mathbf{r}_l^{(j)}$ denotes the off-peak sample taken at time $(l+1)T + jn_0T_c$, where $n_0 \geq 2$, and $j = 1, \ldots, J$. Since the MAI-plus-AWGN components in the terms $\mathbf{r}_l^{(j)}$ are asymptotically uncorrelated as $N$ gets large, they can be approximated as jointly independent by the CGA. The magnitude of the desired user component is suppressed by the processing gain. Therefore, the conditional covariance $\mathbf{U}_l$ can be estimated by averaging outer products of the off-peak samples, i.e.,

$$\mathbf{U}_l \approx \frac{1}{J}\sum_{j=1}^{J}\mathbf{r}_l^{(j)}\mathbf{r}_l^{(j)H} \qquad 1 \leq J \leq \left\lfloor\frac{N}{n_0}\right\rfloor - 1. \tag{19}$$

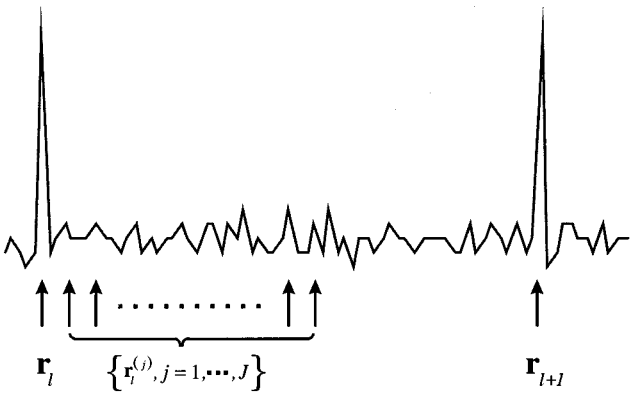



Fig. 3. Sampling of MF outputs.

Since the estimation is done in one symbol time, during which the channel condition is almost constant, (19) can be viewed as the sample covariance of independent identically distributed (i.i.d.) random trials of $\mathbf{n}_l$ conditional on $\mathbf{G}_l$. In addition, abrupt changes might occur between $\mathbf{r}_l$ and $\mathbf{r}_{l+1}$ if an interferer starts transmission during the period $[(l+1)T, (l+1)T + n_0JT_c]$. Nevertheless, this can be monitored by the squared norm of the off-peak sample vectors. If large, abrupt changes are observed, the estimate of $\mathbf{U}_l$ can be replaced by the previous estimate $\mathbf{U}_{l-1}$.

### D. Estimation of AR Model Noise Variance

In [10]–[13], the model noise variance $\sigma^2_{w_l}$ was given *a priori*. However, in the channel identification problem, it is unreasonable to assume the model noise variance is known, while the AR parameters are not. The above EKF algorithm can only estimate the AR parameters because the model noise variance is not related to the Taylor expansion of the state equation. Given the second-order AR coefficients, a simple expression for the normalized model noise variance can be found in [11]. However, it requires a perfect estimate of the average signal power.[4] A less stringent power estimate is also required in [14]. To get an estimate of $\sigma^2_{w_l}$ without a power estimate, we observe that the model noise variance approximately fits the following relations. First, if the difference between the true state vector and its estimate $\mathbf{x}_{l|l-1}$ is negligible, i.e.,

$$\mathbf{x}_l \approx \mathbf{F}(\hat{\mathbf{x}}_{l-1|l-1}) + \mathbf{K}_l\boldsymbol{\varepsilon}_{l|l-1}$$

then

$$\Delta\mathbf{x}_{l|l-1} = \mathbf{x}_l - \hat{\mathbf{x}}_{l|l-1} \approx \mathbf{K}_l\boldsymbol{\varepsilon}_{l|l-1}.$$

Also, if the model parameters are correct, the variance of the model noise $w_l$ dominates the prediction error. Therefore

$$\sigma^2_{w_l} = \text{tr}\,\{\sigma^2_{w_l}\mathbf{e}_1\mathbf{e}_1^H\} \approx \mathbf{E}\{\Delta\mathbf{x}^H_{l|l-1}\Delta\mathbf{x}_{l|l-1}\}. \tag{20}$$

The intuition of this approximation comes from the feedback of innovations used in [26]. There are two reasons for not using the modified EKF in [26]. First, the Kalman gain matrix must

[4]Using notation in this paper, [11, (42)] should be $\hat{\sigma}^2_w = ([(1+a_2)^2 - a_1^2](1-a_1))/(1+a_2)E\{|g_l|^2\}$, where the average power of the received signal $E\{|g_l|^2\}$ is normalized to 1 *a priori* by the assumption of perfect power estimation.

also be included in the state vector, increasing its dimension from $2p+M$ to $2p+M+pM$. In addition, the Kalman gain matrix is considered as a fixed set of parameters in [26]. Since the MAI-plus-AWGN covariance is time-varying and estimated at each sampling time, the assumption of a fixed Kalman gain matrix cannot reflect the instantaneous interference property. Therefore, an estimator of $\sigma_{w_l}^2$ for the $\mathbf{P}_{l|l-1}$ recursion [provided later in (22)] is introduced to complete the computation of $\mathbf{K}_l$ in (17). Using a weighted average to account for the changing dynamics of the channel, the estimation of $\sigma_{w_l}^2$ is

$$\hat{\sigma}_{w,l}^2 = \frac{\sum_{i=1}^{l} \rho^{l-i} \boldsymbol{\varepsilon}_{l|l-1}^H \mathbf{K}_l^H \mathbf{K}_l \boldsymbol{\varepsilon}_{l|l-1}}{\sum_{i=1}^{l} \rho^{l-i}} = \frac{\gamma_l}{q_l}$$

in which

$$\gamma_l = \rho\gamma_{l-1} + \boldsymbol{\varepsilon}_{l|l-1}^H \mathbf{K}_l^H \mathbf{K}_l \boldsymbol{\varepsilon}_{l|l-1}, \gamma_0 = 0$$
$$q_l = \rho q_{l-1} + 1, \quad \text{and} \quad q_0 = 0$$

where $\rho$ is the forgetting factor. The advantage of this estimator is that the estimation of the average received power is integrated into the algorithm. Unlike most MAP receivers, since the MAI-plus-AWGN covariance estimator is obtained by averaging outer products of off-peak samples, the proposed algorithm does not need any SNR information from the automatic gain control (AGC) unit. One drawback of the above estimator is that it might result in biased estimates of AR coefficients [26]; however, the bit error rate (BER) performance of MAP detection is consistent over the entire simulated range of signal-to-noise ratio (SNR) (see simulations in Section V). In fact, as long as the prediction error of the true data sequence is consistently smaller than those of other hypotheses, the MAP detection can still make correct decisions in spite of the biased AR coefficients.

### *E. Recursion of Estimation Error Covariance*

The recursion of the state-error covariance is divided into two steps in the EKF process. The purpose for doing so is to ensure the invertibility of the innovation covariance matrices $\boldsymbol{\Sigma}_{l|l}$. By (15)–(18), the recursion can be expressed as follows:

$$\mathbf{P}_{l|l} = (\mathbf{I}_{2p+M} - \mathbf{K}_l \hat{\boldsymbol{\Psi}}_{l|l-1}) \mathbf{P}_{l|l-1} \quad \text{and} \tag{21}$$
$$\mathbf{P}_{l+1|l} = \hat{\boldsymbol{\Phi}}_{l|l} \mathbf{P}_{l|l} \hat{\boldsymbol{\Phi}}_{l|l}^H + \tilde{\mathbf{B}}_l \tag{22}$$

where

$$\mathbf{P}_{l|l} = \left[\begin{array}{c|c} p \times p \text{ matrix} & \\ \hline & \tilde{\mathbf{P}}_{l|l} \end{array}\right]_{(2p+M)\times(2p+M)}$$
$$= E\{\Delta\mathbf{x}_{l|l} \Delta\mathbf{x}_{l|l}^H\} \quad \text{and}$$
$$\tilde{\mathbf{B}}_l = \left[\begin{array}{c|c|c} \hat{\sigma}_{w,l}^2 & & \\ \hline & \mathbf{O}_{p-1} & \\ \hline & & -\xi\tilde{\mathbf{P}}_{l|l}^2 \end{array}\right]$$

in which $\tilde{\mathbf{P}}_{l|l}$ is the submatrix taken from $\mathbf{P}_{l|l}$ corresponding to

$$E\left\{\begin{bmatrix} \Delta\hat{\mathbf{a}}_{l|l} \\ \Delta\mathbf{h}_{l|l} \end{bmatrix} [\Delta\hat{\mathbf{a}}_{l|l}^H \quad \Delta\mathbf{h}_{l|l}^H]\right\}.$$

The matrix $\tilde{\mathbf{B}}_l$ is chosen in this way for numerical reasons [26]. Usually, $\xi$ is chosen to be small ($10^{-8}$ in our simulations).

### *F. Constraints on the Array Vector* $\mathbf{h}_l$

In addition, the constraint on the array vector $\mathbf{h}_l$ given in (14) must be considered. However, because (14) is not a differentiable equality, the constrained mmse solution to (17) forms a difficult mathematical problem. To work around this difficulty, we normalize the estimated array vector $\hat{\mathbf{h}}_{l|l}$ after each EKF update of $\hat{\mathbf{x}}_{l|l-1}$ to $\hat{\mathbf{x}}_{l|l}$. That is, after (17), we perform

$$\hat{\mathbf{g}}_{l|l} \leftarrow \|\hat{\mathbf{h}}_{l|l}\| \frac{\hat{\mathbf{g}}_{l|l}}{\sqrt{M}} \quad \text{and} \quad \hat{\mathbf{h}}_{l|l} \leftarrow \sqrt{M} \frac{\hat{\mathbf{h}}_{l|l}}{\|\hat{\mathbf{h}}_{l|l}\|}. \tag{23}$$

The intuition behind this normalization is to prevent the magnitude of the estimated array vector $\hat{\mathbf{h}}_l$ from being unconstrained. If the magnitude of $\hat{\mathbf{h}}_{l|l}$ were unconstrained, $\hat{g}_{l|l}$ and $\hat{\mathbf{h}}_{l|l}$ could drift without bound because the mmse criterion applies to the product of the quantities and not to the individual quantities. Because of the assumed noise structure in (15) and the presumed approximation in (20), the estimate of $g_l$ will become extremely noisy when its magnitude drops significantly. Since the product $\hat{g}_{l|l}\hat{\mathbf{h}}_{l|l}$ is kept close to the true $g_l\mathbf{h}_l$ by the EKF, we leave out the constraint (14) but instead use the common scalar $\hat{g}_{l|l}$ for equalizing magnitudes of array elements. By this simple, albeit suboptimal, normalization, the channel estimate $\hat{g}_{l|l}$ will stay at a reasonable level so that the EKF can work properly. Note that with (23), submatrices in $\mathbf{P}_{l|l}$ corresponding to $E\{|\Delta g_{l|l}|^2\}$ and $E\{\Delta\mathbf{h}_{l|l}\Delta\mathbf{h}_{l|l}^H\}$ must be scaled before computing (22). The derivation for the scaling procedure is straightforward and, hence, omitted.

### *G. Joint MAP Detection Using Innovations*

From the previous section, the channel output prediction $\hat{\mathbf{r}}_{l|l-1}$ is an mmse estimate of $\mathbf{r}_l$ conditioned on the past channel output $\mathbf{R}_{l-1}$. Because $\mathbf{r}_l$ and $\mathbf{R}_{l-1}$ are conditionally jointly Gaussian, $\hat{\mathbf{r}}_{l|l-1}$ is also the ML estimate of $\mathbf{r}_l|\mathbf{R}_{l-1}$. Accordingly, the random vector $\mathbf{r}_l$ conditioned on $\mathbf{R}_{l-1}$ has the pdf given by

$$p(\mathbf{r}_l|\mathbf{R}_{l-1}) = \frac{1}{\pi^M|\boldsymbol{\Sigma}_{l|l-1}|} \exp[-(\mathbf{r}_l - \hat{\mathbf{r}}_{l|l-1})^H \boldsymbol{\Sigma}_{l|l-1}^{-1} (\mathbf{r}_l - \hat{\mathbf{r}}_{l|l-1})]. \tag{24}$$

The above equation is equivalent to the pdf in (10) defined in (8) of Section III. Therefore, the recursive MAP detection formulated in (8) and (9) can be completed as follows:

Step 1) At time $l$, $\forall\, \mathbf{d}_l(D) \in \mathcal{S}^{D+1}$, perform the following:
- EKF: (17), (18), and (21).
- Normalization: (23) and

$$\mathbf{P}_{l|l} \leftarrow \text{Scaling } \mathbf{P}_{l|l}$$

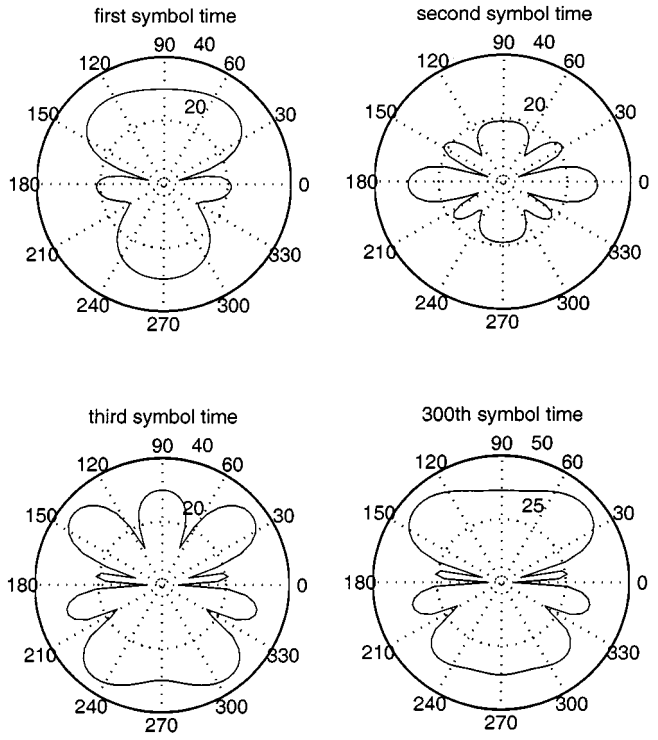



Fig. 4. Beamforming process of the five-element array gain pattern.

- Conditional pdf computation: (24) and (8)
- Estimation error covariance update: (22)

Step 2) D-lag MAP detection and decision feedback: (7) and (9).

Finally, although different from the channel interpolation in PSAM, the estimation still requires pilots if coherent detection is applied. Notice that during a deep fade, a sequence of phase-reversed symbol hypotheses can have smaller prediction errors. Once a wrong decision is made, a cycle slipping may persist in subsequent detections if those detections are performed without the assistance of pilots. With the MAP detection proposed in Section III, differential encoding eliminates the use of pilots that are needed for the prevention of cycle slipping in coherent systems. The performance of pilot-assisted and differentially encoded systems will be compared in the next section.

## V. Simulation Results

In this section, simulation results of the proposed algorithm are presented. The Rayleigh fading channel is produced by the Jakes simulator [4] for simplicity. In fact, since the Jakes simulator produces a random process with a nonrational power spectrum, it can also verify the validity of the AR channel approximation. The processing gain is 128 (21 dB) and the normalized Doppler spread (Doppler shift multiplied by the symbol duration) is chosen to be 0.01. For the AR channel model, we assume first- and second-order models for simplicity. The forgetting factor $\rho$ is set at 0.95 in order to track the model noise variance $\sigma_{w_l}^2$ effectively. The separation of MF output sampling ($n_0$) is four, and the number of samples ($J$) for estimating the MAI-plus-AWGN covariance is 30. The decision delay $D$ is equal to six.

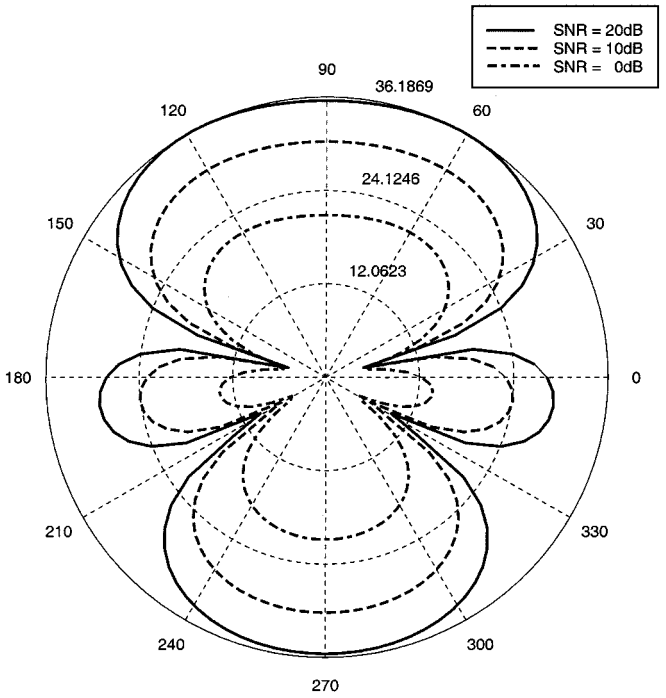



Fig. 5. Directional pattern of the three-element array formed by first row of Kalman gain matrix.

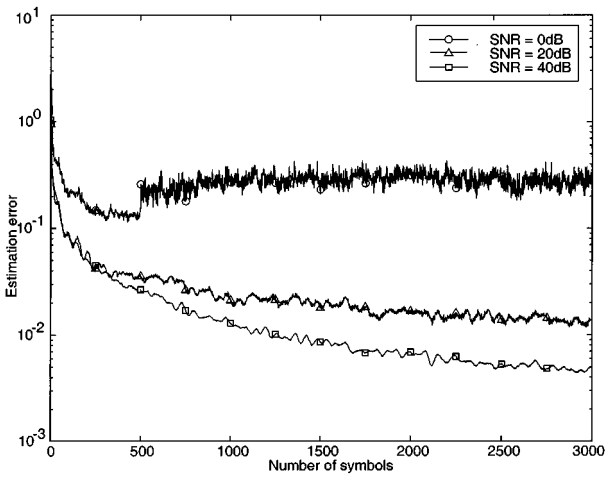



Fig. 6. MF output prediction error averaged over 3 antenna elements through 300 runs.

The directional interference rejection capability of the channel estimator described in Section IV is presented in Figs. 4 and 5, which show the squared inner product of the first row of the Kalman gain matrix and the phase array vectors corresponding to different incidence angles. The first row of the Kalman gain matrix represents the mmse estimator of $\Delta g_{l|l-1}$ from the observation $\varepsilon_{l|l-1}$. Because the MAI-plus-AWGN vector $\mathbf{n}_l$ in $\varepsilon_{l|l-1}$ does not contain information for estimating the desired user's channel, the first row ofthe KF should be designed to suppress strong MAI by nulling it directionally. In contrast, the channel prediction error at the desired user's direction should be extracted so that the estimator can keep adapting to the latest channel variations. The effect of such extraction and suppression can be clearly demonstrated by polar plots as in Figs. 4 and 5, where the gain is expressed

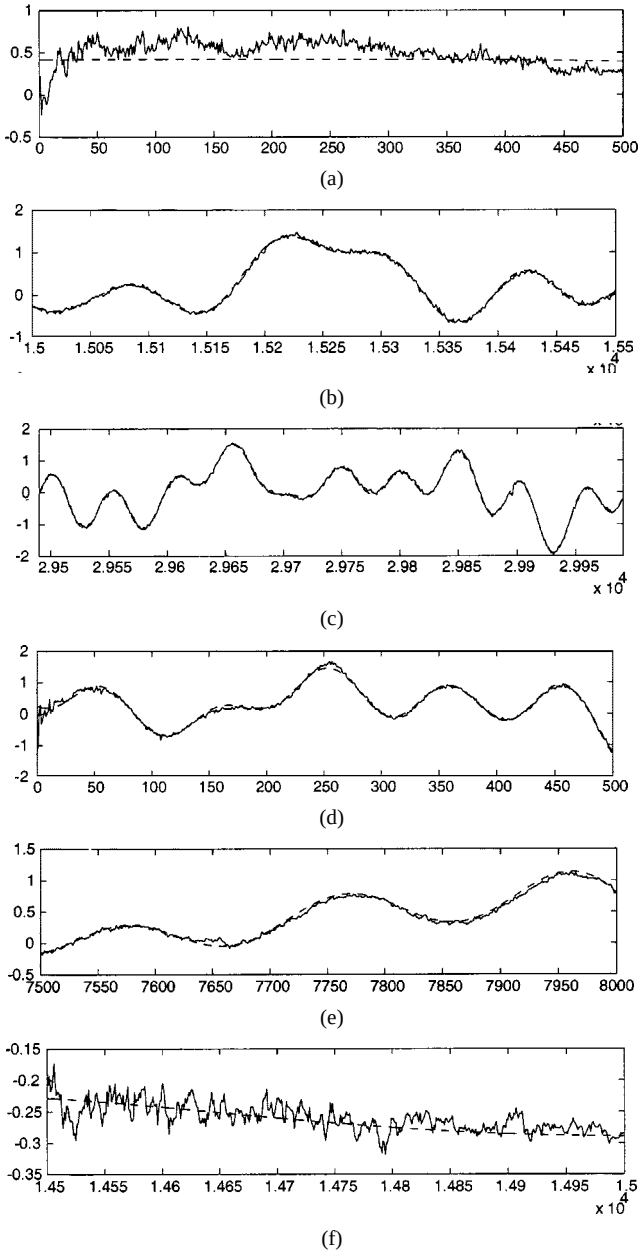



Fig. 7. Dashed line: actual. Solid line: estimated. Tracking of fading channel with slowly varying doppler spread. Real parts of the actual and estimated fading channel coefficients are shown.

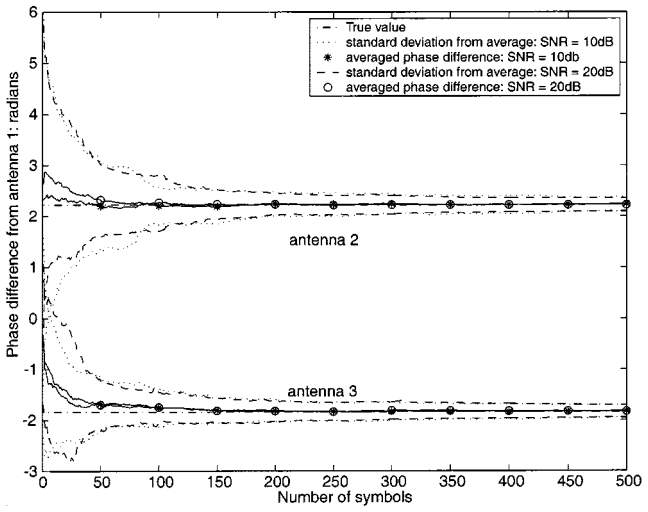
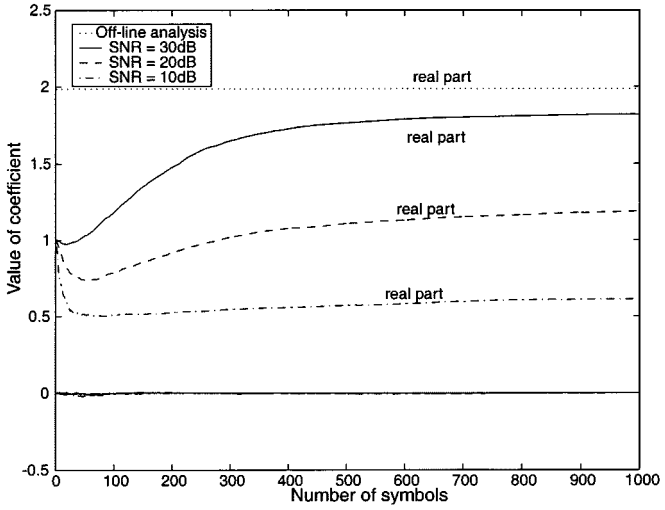



Fig. 8. Estimated incidence angle of desired user: relative phase difference to the first antenna element (averaged over 100 runs).

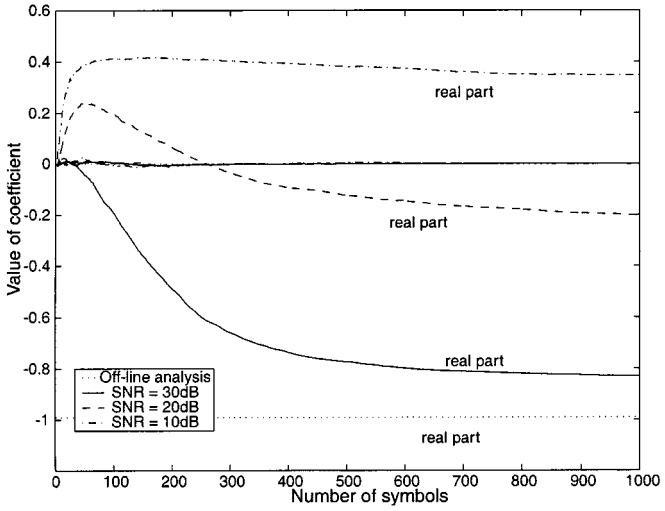
radially in decibels, with the minimum gain subtracted. The plots in Fig. 4 show the beamforming process of a five-element array. The incidence angle of the desired user is 45°, and those of the interferers are 15°, 0°, and $-30°$. All interfering users are 25 dB stronger than the desired user. Notice that in the first two symbols, because the initial condition of the array vector is an all-ones vector, the direction of the desired signal is roughly pointing at 0°. Consequently, the null of the antenna pattern at 0° is not obvious for the first two symbols. Nevertheless, nulls at 15° and $-30°$ appear immediately once the algorithm starts, showing the instantaneous MAI suppression ability. As the adaptation progresses, the direction of the array vector moves to the right direction at 45° and results in an mmse gain pattern. The gain pattern of a three-element array is shown in Fig. 5, which is averaged over the first 250-symbol period at different SNR settings. The incidence angle of the desired user remains the same, and those of interferers are 15° and $-30°$. Interferers are set to be 30 dB stronger than the desired user in this case. Because the number of users exceeds the degrees of freedom ($M-1$), there is some loss in the direction of the desired user.

The remaining simulation results are all based on the setup in Fig. 5. The second result is to show the performance of the channel estimator. The channel coefficient prediction errors at different $E_b/N_0$ are plotted in Fig. 6, which indicates the convergence rate of the channel estimator. The length of the training sequence is set to be 500 symbols to ensure the proper identification of model parameters, although the convergence shown in Fig. 6 indicates possibilities for shorter training. The sudden increment of estimation error after the training period at 0 dB reflects the effect of error propagation. As can be seen from the two curves at 20 and 40 dB, this effect has negligible impact on the performance of the channel estimation.

The ability to track slowly varying Doppler spread is illustrated in Fig. 7, where channel coefficients and their estimates with both increasing and decreasing Doppler spreads are plotted. At the symbol rate of 10 kHz and the SNR of 20 dB, Figs. 7(a)–7(c) and Figs. 7(d)–(f) correspond to Doppler spreads from 0 up to 0.01 in 3 s and from 0.01 down to 0 in 1.5 s, respectively. The capability of finding the desired user's incidence angle is shown in Fig. 8 by the difference of estimated phases between array elements, which converges in 250 symbols. From results shown in Fig. 7 and 8, the quasistationary assumption made in the previous sections is justified because the adaptation is sufficient for such variations. Also, the estimated AR model coefficients are compared with an analysis result in Fig. 9 that is not done in real-time. The non-real-time analysis uses the forward-backward approach [27] on the noise-free fading process. The curves showing the real parts are labeled in Fig. 9, and those showing the imaginary parts are overlapped on the horizontal line corresponding to zero. At low SNR, we can observe the bias in AR coefficients, but the corresponding BER performance is consistent with that at high SNR (see Fig. 10).




(a)




(b)

Fig. 9. Comparison of estimated AR coefficient and off-line analysis (averaged over 100 runs). (a) First order coefficient $\hat{a}_{1,l}$: real and imaginary parts. (b) Second order coefficient $\hat{a}_{2,l}$: real and imaginary parts.

Finally, we present the BER performance of the EKF-MAP detection scheme. For comparison, performance curves of both ideal coherent BPSK and ideal PSAM with no interference are plotted. By [1], the average BER of ideal coherent BPSK communication in a flat fading channel is

$$\overline{P}_e = \frac{1}{2}\left[1 - \sqrt{\frac{M\overline{\gamma}_b}{1+M\overline{\gamma}_b}}\right] \tag{25}$$

where $\overline{\gamma}_b = E\{|g_l|^2/N_0\}$. Since exact values of the channel coefficients are assumed in (25), the BER of (25) is unreachable in practice. For ideal PSAM, the BER is obtained from solving a normal equation, which involves known Doppler spread, known SNR, and the 0th-order Bessel function of the first kind [8]. For practical applications, fixed-design PSAM optimized for the worst-case Doppler and a prespecified SNR is proposed in [8]. Nevertheless, ideal PSAM can still serve as a benchmark. For the tested coherent EKF-MAP algorithm, the pilot insertion rate is 1/7, and the AR model order is chosen to be 1. As Fig. 10 shows, even in the presence of strong interference, the coherent

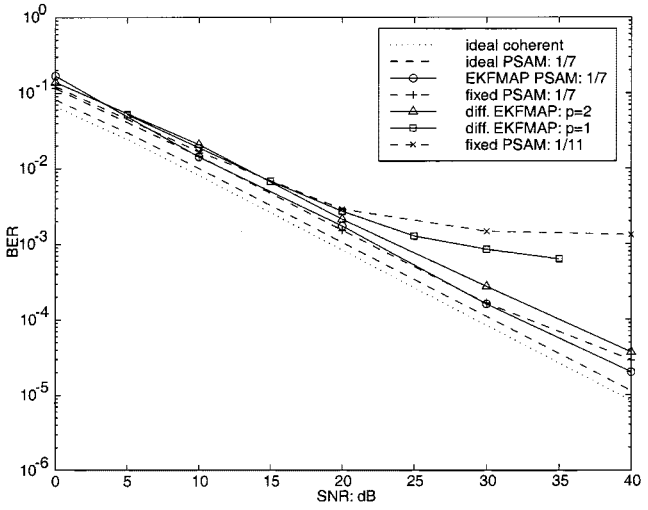



Fig. 10. BER performance of coherent and differentially coherent EKF-MAP algorithm.

EKF-MAP algorithm can attain a BER performance 3 dB from the ideal coherent BPSK and 1dB from the ideal PSAM. BER curves of the differential EKF-MAP algorithm with both the first- and the second-order models are also plotted. Compared with the coherent EKF-MAP algorithm, the proposed differential algorithm can achieve a higher spectral efficiency with a small performance degradation. This is unachievable for fixed PSAM [8] with pilot rate 1/11, even in an interference-free environment.

## VI. Conclusion

A joint channel estimation and MAP detection algorithm for DS-CDMA systems in a time-selective fading channel is developed in this paper. Using the MF output characteristics and the extra spatial dimensions provided by an antenna array, strong interference can be extracted and suppressed. In order to provide a real-time estimation, the fast-varying channel is approximated by an AR model with a total number of parameters that is much smaller than the previously proposed MA model. Also, we include the unknown antenna array vector and apply extended Kalman filtering techniques to identify the unknown channel parameters. From the statistical distribution of the prediction error (innovations), we have proposed that the differential MAP detection be performed jointly with channel estimation so that a pilot-free transmission becomes possible. With no *a priori* information, this approach achieves a near-ideal coherent system performance.

## References

[1] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[3] *Mobile Station-Base Station Compatibility Standard for Dual Mode Wideband Spread Spectrum Cellular System*, TIA/EIA/IS-95 Interim Stand., July 1993.

[4] W. C. Jakes, *Microwave Mobile Communications*. Piscataway, NJ: IEEE Press, 1994.

[5] W. C. Y. Lee, *Mobile Communications Engineering: Theory and Applications*, 2nd ed. New York, NY: McGraw-Hill, 1997.

[6] G. Raleigh, S. N. Diggavi, A. F. Naguib, and A. Paulraj, "Characterization of fast fading vector channels for multi-antenna communication systems," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Los Alamitos, CA, 1994, pp. 853–857.
[7] A. Stéphenne and B. Champagne, "A new multi-path vector channel simulator for the performance evaluation of antenna array systems," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, vol. 3. Piscataway, NJ, 1997, pp. 1125–1229.
[8] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.
[9] J. H. Lodge and M. L. Moher, "Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 38, pp. 787–794, June 1990.
[10] R. A. Iltis, "Joint estimation of PN code delay and multipath using the extended Kalman filter," *IEEE Trans. Commun.*, vol. 38, pp. 1677–1685, Oct. 1990.
[11] H.-Y. Wu and A. Duel-Hallen, "Performance comparison of multiuser detectors with channel estimation for flat Rayleigh fading CDMA channels," *Wireless Pers. Commun.*, vol. 6, no. 1–2, pp. 137–160, Jan. 1998.
[12] S. Vasudevan and M. K. Varanasi, "Achieving near-optimum asymptotic efficiency and fading resistance over the time-varying Rayleigh-faded CDMA channel," *IEEE Trans. Commun.*, vol. 44, pp. 1130–1143, Sept. 1996.
[13] Z. Zvonar and M. Stojanovic, "Performance of antenna diversity multiuser receivers in CDMA channels with imperfect fading estimation," *Wireless Pers. Commun.*, vol. 3, no. 1–2, pp. 91–110, 1996.
[14] T. F. Wong, Q. Zhang, and J. S. Lehnert, "Decision feedback MAP receiver for time-selective fading CDMA channels," *IEEE Trans. Commun.*, vol. 48, pp. 829–840, May 2000.
[15] S. Vembu and A. J. Viterbi, "Two different philosophies in CDMA—A comparison," in *Proc. IEEE Veh. Technol. Conf.*, Atlanta, GA, Apr. 28–May 1, 1996, pp. 869–873.
[16] T. M. Lok and J. S. Lehnert, "Error probabilities for generalized quadriphase DS/SSMA communications with random signature sequences," *IEEE Trans. Commun.*, vol. 44, pp. 876–885, Oct. 1996.
[17] ——, "An asymptotic analysis of DS/SSMA communication systems with random polyphase signature sequences," *IEEE Trans. Inform. Theory*, vol. 42, pp. 129–136, Jan. 1996.
[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] Y. Li, B. Vucetic, and Y. Sato, "Optimum soft-output detection for channels with intersymbol interference," *IEEE Trans. Information Theory*, vol. 41, pp. 704–713, May 1995.
[20] P. Robertson, P. Hoeher, and E. Villebrun, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo-decoding," *Euro. Trans. Telecommun. Related Areas*, vol. 8, no. 2, pp. 119–125, Mar./Apr. 1997.
[21] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel models for mobile radio systems with smart antenna," *Proc. Inst. Elect. Eng. Radar, Sonar Navig.*, vol. 145, no. 1, pp. 32–41, Feb. 1998.
[22] G. Lee, S. B. Gelfand, and M. P. Fitz, "Bayesian techniques for blind deconvolution," *IEEE Trans. Commun.*, vol. 44, pp. 826–835, July 1996.
[23] A. Larsson, "Power control in mobile communications—A statistical analysis of the radio channel variability," Licentiate and Doctoral dissertation, Dept. Appl. Electron., Lund Univ., Lund, Sweden, Oct. 1997.
[24] P. M. Crespo and J. Jiménez, "Computer simulation of radio channels using a harmonic decomposition technique," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 414–419, Aug. 1995.
[25] T. F. Wong, T. M. Lok, J. S. Lehnert, and M. D. Zoltowski, "A linear receiver for direct-sequence spread-spectrum multiple-access systems with antenna arrays and blind adaptation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 659–676, Mar. 1998.
[26] L. Ljung, "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," *IEEE Trans. Automatic Control*, vol. AC-24, pp. 36–50, Feb. 1979.
[27] ——, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

**Shiauhe Tsai** (M'00) received the B.S. degree from Catholic Fu-Jen University, Taipei, Taiwan, in 1991, and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1995 and 2000, respectively.

From 1991 to 1994, he served as Second Lieutenant in the R.O.C. Army and worked as a full-time lab administrator in Catholic Fu-Jen University. From 1995 to 2000, he was a Teaching and Research Assistant with the School of Electrical and Computer Engineering, Purdue University, and participated in the Small Unit Operations and the Global Mobile Computing programs of the Defense Advanced Research Agency. He also helped to establish the Spread Spectrum and Satellite Communications Laboratory at Purdue University. He is currently a Senior Engineer at Ericsson Wireless Communications, San Diego, CA. His current research interests include spread-spectrum and satellite communications, error control coding, digital signal processing, and wireless fading channels.

**Tan F. Wong** (S'96–M'97) received the B.Sc. degree (with first-class honors) in electronic engineering from the Chinese University of Hong Kong in 1991 and the M.S.E.E. and Ph. D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1992 and 1997, respectively.

He was a Research Engineer working on the high-speed wireless networks project in the Department of Electronics, Macquarie University, Sydney, Australia. He also served as a Postdoctoral Research Associate with the School of Electrical and Computer Engineering, Purdue University. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Florida, Gainesville. His research interests include spread-spectrum communication systems, multiuser communications, and wireless cellular networks.

**James S. Lehnert** (S'83–M'84–SM'95–F'00) received the B.S. (with highest honors), M.S., and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1978, 1981, and 1984, respectively.

From 1978 to 1984, he was a Research Assistant with the Coordinated Science Laboratory, University of Illinois. He was a University of Illinois Fellow from 1978 to 1979 and an IBM Pre-Doctoral Fellow from 1982 to 1984. He has held summer positions at Motorola Communications, Schaumburg, IL, in the Data Systems Research Laboratory and Harris Corporation, Melbourne, FL, in the Advanced Technology Department. He is currently a Professor with the Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN.

Dr. Lehnert has served as Editor for Spread Spectrum for the IEEE TRANSACTIONS ON COMMUNICATIONS and as Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.