A Blind Particle Filtering Detector of Signals Transmitted Over Flat Fading Channels

Yufei Huang, *Member, IEEE,* and Petar M. Djurić, *Senior Member, IEEE*

***Abstract*—A new particle filtering detector (PFD) is proposed for blind signal detection over flat Rayleigh fading channels whose model coefficients are unknown. The detector employs a hybrid importance function and a mixture Kalman filter. It also incorporates an auxiliary particle filtering strategy with a smoothing kernel in the resampling step. Further, by considering practical information of communication systems and the physical interpretation of the adopted second-order autoregressive (AR) channel model, a fully blind particle filtering implementation is developed. The structure of the proposed PFD can be easily adapted to other system requirements. Simulations are provided that demonstrate the performance of the new PFD.**

***Index Terms*—Autoregressive moving average, blind signal detection, least mean square, particle filtering detector, recursive least square.**

I. Introduction

With the ever-growing frequency bandwidth used for wireless transmission of voice and data, the detection of transmitted symbols through frequency flat Rayleigh fading channels is increasingly challenging. This problem has been treated as detection of Gaussian signals in Gaussian noise, and under the maximum *a posteriori* (MAP) criterion, its optimum solution is known to be a quadratic receiver: a form of the discrete Wiener filter [1].

In the past two decades, the maximum-likelihood sequence detection (MLSD) in both flat and frequency-selective fading channels has drawn much research interest [2]–[4]. With known channel state information (CSI), the MLSD is implemented using the Viterbi algorithm, and the solution is optimum [2], [5]. When the CSI is unknown to the receiver, it must be extracted in order to carry out signal detection. It is known that when linear parametric structures such as autoregressive (AR) and autoregressive moving average (ARMA) models with known parameters are used to represent fading channels, the optimal solution can be obtained by a bank of Kalman filters [6], [7]. However, this optimal implementation requires a separate run of a Kalman filter for each possible sequence, and that leads to a structure of exponential complexity. To reduce the complexity, suboptimal algorithms have been proposed. These algorithms usually consist of a separate channel estimator followed by a per-survive sequence detector [3], [4]. The channel estimator often employs linear prediction, least mean square (LMS), or recursive least square (RLS) algorithms. In comparison with optimal implementations, these suboptimal solutions suffer from a number of drawbacks. First, since no channel model is assumed, the tracking ability degrades especially for fast fading channels. Second, due to time correlation of the fading channel, the surviving branches in the per-survive algorithm may not guarantee larger metrics than the discarded branches in later stages. As a result, the branch associated with the optimal solution could be discarded in an early stage, leaving only suboptimal solutions for the rest of the process. This implies that an error committed at an early stage propagates.

Recently, novel particle filtering detectors (PFDs) for sequential detection with linear complexity have been proposed [8]–[11]. Using the idea of sequential Monte Carlo sampling, these detectors can approximate the optimum solution directly without compromising the system model. Additionally, another salient feature of these detectors is that the decision made at time t does not depend on any decisions made previously, and thus, no error is propagated in their implementation. Furthermore, they are fully blind detectors and allow for both Gaussian and non-Gaussian ambient noise as well as parallel implementations.

Commonly, in the implementation of the PFDs, state-space models are adopted, and a typical assumption is that the model coefficients are known to the detectors. In practice, they have to be estimated in advance, and to obtain good estimate, a separate long interval of training data is required. In addition, since the model parameters are assumed fixed, these algorithms cannot adapt to fast changes of the channel statistics. In [12], a hybrid algorithm is proposed to integrate the estimation of the model parameters within the PFD. This novel detector employs an RLS algorithm for estimation of the unknown parameters. To avoid ambiguity, pilot symbols are required in the implementation.

In this paper, a new PFD is proposed for blind signal detection under unknown channel model coefficients. The PFD employs a newly introduced hybrid importance function [13] that, together with mixture Kalman filtering (MKF) [14]–[17], reduces significantly the computational complexity of a generic implementation of the particle filtering algorithm. An AR(2) process is adopted to model the fading channels, where the modeling imposes a direct link between the AR coefficients and the underlying fading channel. The physical information of the communication systems also enables us to resolve inherent ambiguities that may arise in the implementation of the PFD.

Manuscript received August 29, 2002; revised July 7, 2003. This work was supported by the National Science Foundation under Grants CCR-9903120 and CCR-0082607. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Joseph Tabrikian.

Y. Huang is with the Department of Electrical Engineering, University of Texas at San Antonio, San Antonio, TX 78249-0669 USA (e-mail: yfhuang@utsa.edu).

P. M. Djurić is with the Department of Electrical and Computer Engineering, State University of New York, Stony Brook, NY, 11794 USA (e-mail: djuric@ece.sunysb.edu).

Digital Object Identifier 10.1109/TSP.2004.828902

Consequently, a fully blind implementation of the PFD is possible. Finally, a resampling technique is employed that incorporates auxiliary particle filtering and a smoothing kernel, and it is combined with the hybrid importance function to further improve the efficiency and effectiveness of the proposed PFD.

The paper is organized as follows. In Section II, the mathematical formulation of the problem is given. In Section III, the PFD is proposed where the implementation procedure, initial sampling, and resampling methods are described. In Section IV, the extensions of the detector to cases of semiblind implementation and time-varying AR coefficients are discussed. In Section V, simulation results are provided that show the performance of the proposed detector. Concluding remarks are included in Section VI. There are also two appendices that provide a derivation and a proof of claims made in the paper.

II. Problem Formulation

We consider detection of digital signals transmitted through flat Rayleigh fading channels. At the transmitter, a modulated M-ary PSK symbol sequence s_t is first passed through a pulse shaping filter to form the baseband signal $s(\tau)$, which is then transmitted across a flat Rayleigh fading channel. At the receiver, the received baseband signal $y(\tau)$ is fed into a matched filter and then sampled with a symbol rate $1/T$. The resulting sampled sequence y_t can be expressed as

$$y_t = h_t s_t + e_t \quad t = 1, 2, 3, \cdots \tag{1}$$

where h_t and e_t are the complex fading coefficient and additive ambient noise. The noise e_t is assumed complex Gaussian with zero mean and variance σ^2. Since the fading channel is a Rayleigh process, the stochastic characteristics of the fading coefficient h_t depend on the maximum Doppler spread

$$f_d = \frac{v}{\lambda} \tag{2}$$

where v denotes the speed of the mobile, and λ is the carrier wavelength. When v is constant, h_t is modeled by the Jakes' model as a stationary, circular complex Gaussian process with zero mean and autocorrelation function [18], [19]

$$r_h(m) = E\left\{h_n h_{n-m}^*\right\} = P J_0(2\pi f_d T m) \tag{3}$$

where P denotes the power of the fading process, and $J_0(\cdot)$ is the zero-order Bessel function of the first kind. It is, however, not feasible to directly apply the Jakes' model in our computation for it leads to intractable solutions. Alternatively, an AR process can often be used to approximate the Jakes' model with satisfactory accuracy [20], [21]. Particularly, in this paper, an AR(2) process as used in [19] and [22] is adopted, i.e.,

$$h_t = -a_1 h_{t-1} - a_2 h_{t-2} + v_t \tag{4}$$

where a_1 and a_2 are the model coefficients, and $v_t \sim \mathcal{CN}(0, \sigma_v^2)$, where σ_v^2 is the variance of the driving noise. This variance is chosen such that the average power of h_t matches the power of the fading process P, which leads to

$$\sigma_v^2 = \frac{P(1+a_2)\left((1-a_2)^2 - a_1^2\right)}{(1-a_2)}. \tag{5}$$

The coefficients a_1 and a_2 are closely related to the physical characteristics of the underlying fading process that will be discussed in detail in Section III-B. Here, of our interests, is the detection of the transmitted symbol s_t without knowing the instantaneous value of h_t. Commonly, a_1 and a_2 are unknown, and for detection, they have to be estimated. Their estimation is done separately by, for example, using pilot signals. For accurate estimation of the parameters, a long sequence of pilot signals is needed. It is important to note that this implementation does not have the ability to adapt to the changes of the channel statistics. When changes occur, the sequence of pilot signals must be retransmitted. In our work, we assume no knowledge of a_1 and a_2.

In presenting the problem in a mathematical form, we use the following state-space representation:

$$\begin{cases} a_{1,t} = a_{1,t-1}, \quad a_{2,t} = a_{2,t-1} \\ \mathbf{h}_t = \mathbf{D}\mathbf{h}_{t-1} + \mathbf{g}v_t \end{cases}, \quad \text{state equations}$$
$$y_t = \mathbf{g}^{\mathsf{T}}\mathbf{h}_t s_t + e_t, \quad \text{observation equation} \tag{6}$$

where $\mathbf{h}_t = [h_t \; h_{t-1}]^{\mathsf{T}}$, $\mathbf{g} = [1 \; 0]^{\mathsf{T}}$, and

$$\mathbf{D} = \begin{bmatrix} -a_{1,t} & -a_{2,t} \\ 1 & 0 \end{bmatrix}.$$

Notice that under the stationary channel assumptions, a_1 and a_2 are modeled as static parameters. In a nonstationary channel, as long as the dynamics of a_1 and a_2 can be represented as a Markovian model, the solution can be obtained similarly, as proposed in the sequel. This will be discussed in more detail in a later section. If $\mathbf{a}_t = [a_{1,t}, a_{2,t}]^{\mathsf{T}}$, at any instant of time t, the unknowns are s_t, h_t, and $\mathbf{a}_t$, and our main objective is to detect the transmitted symbol s_t sequentially without sending pilot signals. Note that h_t and $\mathbf{a}_t$ are nuisance parameters. We assume that the power of the channel P and the noise variance σ^2 are known to the receiver, but the proposed algorithm can be easily extended to include them as unknowns. Notice that phase ambiguity exits that is an inherent problem for detection of PSK-modulated signals over Rayleigh fading [19]. To combat phase ambiguity, differential encoding is applied [8], [19], [23].

III. Particle Filtering Detector

A. Particle Filtering Solution with Hybrid Importance Functions and MKF

Particle filtering is a sequential Monte Carlo sampling method built on the Bayesian paradigm [24], [25]. From a Bayesian perspective, at time t, the posterior distribution $p(s_t|y_{0:t})$ is the main entity of interest. However, due to the nonlinearity of the model (6), the analytical expression of $p(s_t|y_{0:t})$ cannot be obtained. Alternatively, particle filtering can be applied to approximate $p(s_t|y_{0:t})$ by stochastic samples generated using a sequential importance sampling strategy.

Now, we explain the proposed PFD. First, we note that given $\mathbf{a}_t$ and s_t, (6) is linear and Gaussian in h_t. Therefore, the MKF can be used to integrate out h_t and obtain $p(s_t, \mathbf{a}_t|y_{0:t})$ [14]. Then, our objective is to generate samples from the distribution $p(s_t, \mathbf{a}_t|y_{0:t})$. For this purpose, define $\mathbf{x}_t^{(j)} = \{\mathbf{a}_t^{(j)}, \mathbf{s}_t^{(j)}\}$, and suppose that at time $t-1$, we have collected N sets of samples $\mathbf{x}_{0:t-1}^{(j)} = \{\mathbf{x}_0^{(j)}, \cdots, \mathbf{x}_{t-1}^{(j)}\}$ with weights

$w_{t-1}^{(j)}, j = 1, \cdots, N$. In particular, the weighted samples, or particles, $\{\mathbf{x}_{0:t-1}^{(j)}, w_{t-1}^{(j)}\}_{j=1}^{N}$ approximate $p(\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t-1})$. When a new observation y_t arrives, sampling from $p(\mathbf{x}_{0:t}|y_{0:t})$ is carried out in a sequential manner based on $\{\mathbf{x}_{0:t-1}^{(j)}, w_{t-1}^{(j)}\}_{j=1}^{N}$, as in Chart 1.

Chart 1. Particle filter

- For $j = 1, \cdots, N$
  - Sample $\mathbf{x}_t^{(j)}$ from an importance function $q(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(j)}, y_{0:t})$, and set $\mathbf{x}_{0:t}^{(j)} = \{\mathbf{x}_{0:t-1}^{(j)}, \mathbf{x}_t^{(j)}\}$.
  - Calculate the weight of $\mathbf{x}_t^{(j)}$ by

$$\bar{w}_t^{(j)} = w_{t-1}^{(j)} \frac{p\left(\mathbf{x}_{0:t}^{(j)}|\mathbf{y}_{0:t}\right)}{p\left(\mathbf{x}_{0:t-1}^{(j)}|y_{0:t-1}\right) q\left(\mathbf{x}_t^{(j)}|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}\right)} \quad (7)$$

- For $j = 1, \cdots, N$, normalize the weights by

$$w_t^{(j)} = \frac{\bar{w}_t^{(j)}}{\sum_{j=1}^{N} \bar{w}_t^{(j)}}. \quad (8)$$

In Chart 1, $q(\mathbf{x}_t^{(j)}|\mathbf{x}_{0:t-1}^{(j)}, y_{0:t})$ is an importance function that must be specified. The choice of the importance function is essential because it determines the efficiency as well as the complexity of the particle filtering algorithm. Two standard choices of the importance function are the posterior and the prior importance functions. The posterior importance function is considered optimal because it minimizes the variance of the importance weights. Here, we observe that due to the presence of a_1 and a_2, the calculation of the posterior importance function leads to intractable computations of the weights. Hence, one would usually resort to using the prior importance function. However, since the prior importance function employs no information from observations in proposing new samples, its use is often ineffective and leads to poor filtering performance. Here, we adopt a hybrid importance function [13], which is expressed as

$$\begin{aligned} &q\left(\mathbf{x}_t|\mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}\right) \\ &\quad = p\left(s_t|\mathbf{a}_t, \mathbf{x}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}\right) p\left(\mathbf{a}_t|\mathbf{a}_{t-1}^{(j)}\right) \quad (9) \\ &\quad = p\left(s_t|\mathbf{a}_t^{(j)}, \mathbf{a}_{0:t-1}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}\right) \delta\left(a_{1,t-1} - a_{1,t-1}^{(j)}\right) \\ &\qquad \times \delta\left(a_{2,t-1} - a_{2,t-1}^{(j)}\right) \quad (10) \end{aligned}$$

where $\mathbf{a}_{0:t}^{(j)}$ and $\mathbf{s}_{0:t}^{(j)}$ are defined in the same way as $\mathbf{x}_{0:t}^{(j)}$ and $\mathbf{a}_t^{(j)} = \mathbf{a}_{t-1}^{(j)}$, and $\delta(\cdot)$ is the Dirac delta function. The last equality is obtained based on the state equations $a_{1,t} = a_{1,t-1}$ and $a_{2,t} = a_{2,t-1}$. The corresponding unnormalized weight is computed by

$$\begin{aligned} \bar{w}_t^{(j)} &= w_{t-1}^{(j)} p\left(y_t|\mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1}\right) \\ &= w_{t-1}^{(j)} \sum_{A_i \in \mathcal{A}} p\left(y_t|s_t = A_i, \mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1}\right) \quad (11) \end{aligned}$$

where $\mathcal{A} = \{A_1, \cdots, A_M\}$ is the alphabet set of s_t. Note that as suggested by its name, the hybrid importance function (9) is a combination of the posterior and the prior importance functions. Intuitively, due to the use of observations, the hybrid importance function is more effective than the prior importance function. Moreover, compared with the posterior importance function, it is implementable since the sampling from (9) and the computation of the weight in (11) can be readily carried out.

Now, we discuss the sampling of s_t and $\mathbf{a}_t$ from (9) and the calculation of the weight (11). First, we observe that no sampling for $\mathbf{a}_t$ is needed, i.e., $\mathbf{a}_t^{(j)} = \mathbf{a}_{t-1}^{(j)}$. Although it simplifies the sampling process, the absence of sampling introduces lack of diversity on $\mathbf{a}_t$. To address this problem, kernel smoothing techniques can be used during the resampling procedure, which will be discussed in Section III-C. As for s_t, since it is discrete, its sampling only requires the evaluation of the importance function on $\mathcal{A}$. In particular, with a uniform prior on s_t, the sampling distribution becomes

$$\begin{aligned} &p\left(s_t = A_i|\mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, \mathbf{y}_{0:t}\right) \\ &\qquad = \frac{p\left(y_t|s_t = A_i, \mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1}\right)}{\sum_{d_i \in \mathcal{A}} p\left(y_t|s_t = d_i, \mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1}\right)}. \quad (12) \end{aligned}$$

Now, from (11) and (12), we see that both the sampling of s_t and the calculation of the weight $w_t^{(j)}$ can be achieved by evaluating $p(y_t|s_t, \mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1})\ \forall s_t \in \mathcal{A}$. This distribution is the likelihood function and can be obtained using the results from the predictive and update steps of the Kalman filter (note that h_t is marginalized out). We show in Appendix A that

$$p\left(y_t|s_t, \mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1}\right) = \mathcal{N}\left(m_t^{(j)}, c_t^{(j)}\right) \quad (13)$$

where $m_t^{(j)}$ and $c_t^{(j)}$ are the mean and variance, respectively, which are derived in Appendix A.

We have identified every element required in the implementation of the particle filtering algorithm. The resulting weighted samples $\{s_t^{(j)}, w_t^{(j)}\}_{j=1}^{N}$ approximate $p(s_t|y_{0:t})$, and the minimum mean square error (MMSE) estimate of s_t can be easily calculated according to

$$\hat{s}_{\text{tMMSE}} = \sum_{j=1}^{N} s_t^{(j)} w_t^{(j)}. \quad (14)$$

In summary, the overall structure of the PFD at time t is presented in Fig. 1. The parallelism of the algorithm is clearly demonstrated. For each trajectory, an independent filter is employed whose interior functional diagram is displayed in Fig. 2.

The performance of the PFD can be further improved by the delayed weight method [15]. In a delayed weight implementation, the MMSE estimate of s_t is computed by

$$\hat{s}_{\text{tMMSE}} = \sum_{j=1}^{N} s_t^{(j)} w_{t+d}^{(j)} \quad (15)$$

where d is a positive integer. In essence, the delayed weight implementation is the same as the implementation of the generic PFD, except that the decision on s_t is delayed for d steps, which requires extra memory for storing the samples of $s_{t:t+d-1}$.

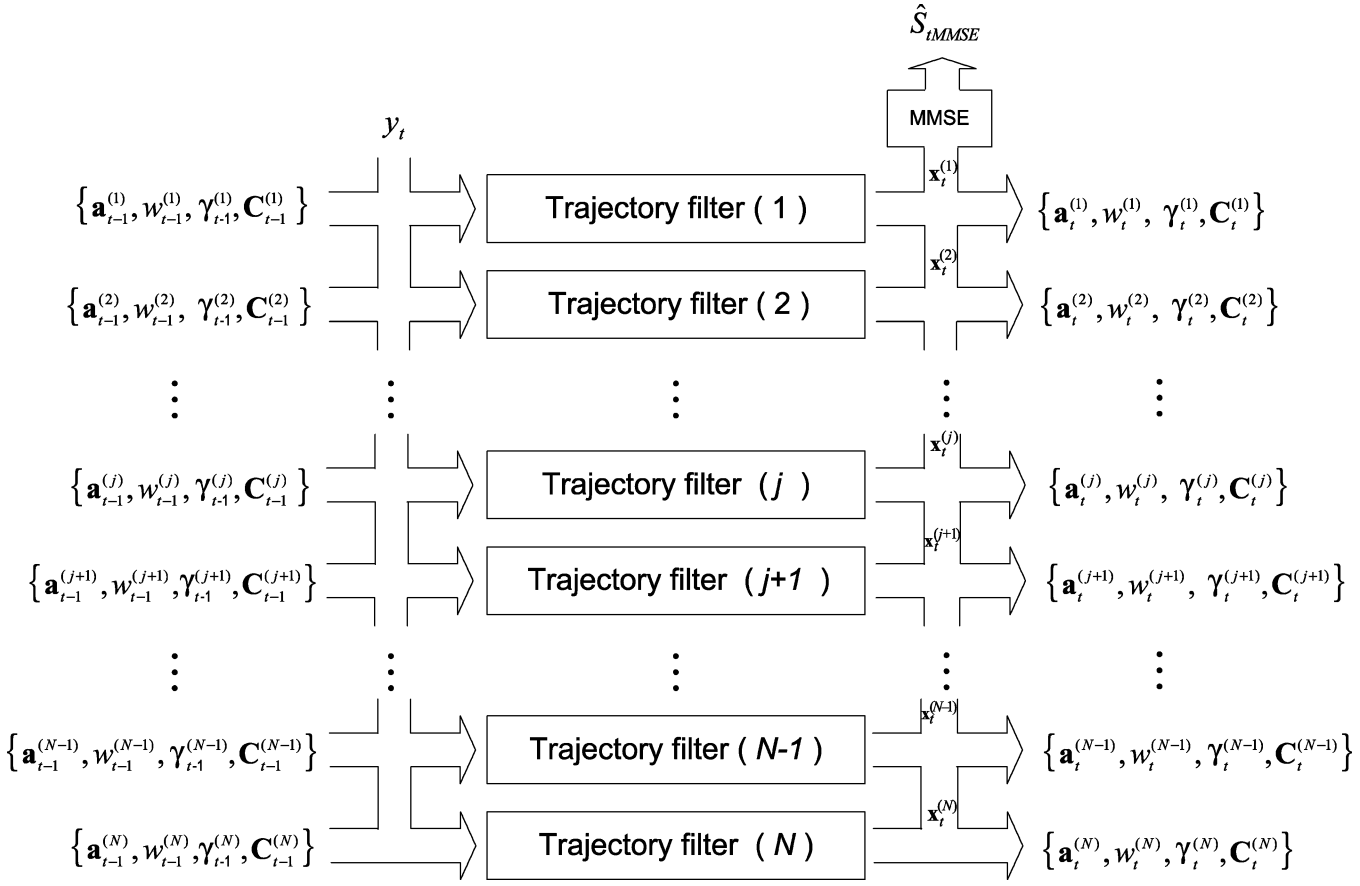


Fig. 1. Structural plot of the particle filtering detector at time t. γ and $\mathbf{C}$ are defined in Appendix A.

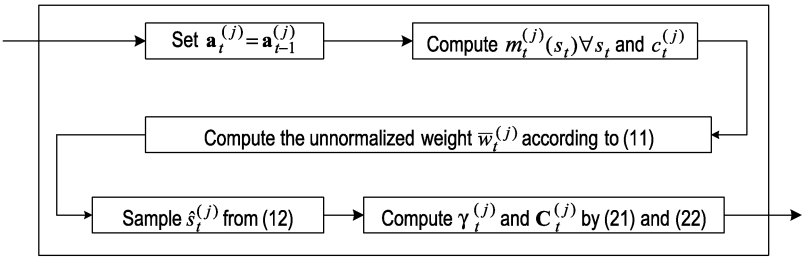


Fig. 2. Interior structure of the jth trajectory filter of the particle filtering detector at time t. γ and $\mathbf{C}$ are defined in Appendix A.

B. Initial Sampling of the AR(2) Coefficients

Before starting the PFD implementation, the initial N samples of $\mathbf{a}_0$ are drawn from a predefined prior distribution. Usually, if no prior knowledge is available, a uniform distribution defined on the whole parameter space is used. However, in our problem, the physical communication systems provide useful prior information, and based on it, the sample space of the uniform distribution can be confined to enhance the performance and the efficiency of the proposed PFD.

First, to ensure stability and minimum phase of the AR(2) process, the parameter space of $\mathbf{a}_0$ is defined within a triangular region depicted in Fig. 3, and the sampling from a uniform distribution defined on the region can be carried out as in [26]. Unfortunately, ambiguity exists when estimating a_1 and a_2 using samples from the triangle region, and it is an inherent problem in blind detection. For instance, when the transmitted symbols are binary phase-shift keying (BPSK) modulated, and if $\bar{a}_1$ and $\bar{a}_2$ are one set of estimates, then $-\bar{a}_1$ and $\bar{a}_2$ would be a set of estimates that would produce the same value of the likelihood function. In other words, there are sets of ambiguity pairs that are symmetric with respect to the a_2 axis. This ambiguity could dramatically deteriorate the bit error rate (BER) performance. For instance, if the detector with estimates $\bar{a}_1$ and $\bar{a}_2$ produces a BER of 0, then the detector with ambiguous estimates would theoretically produce a BER of 0.5. The proof of existence of ambiguity in a BPSK modulated system is shown in Appendix B.

To combat the ambiguity in a fully blind implementation, further restrictions on the sampling space need to be imposed. This can be achieved by considering the relationship between the AR coefficients and the physical parameters of the underlying

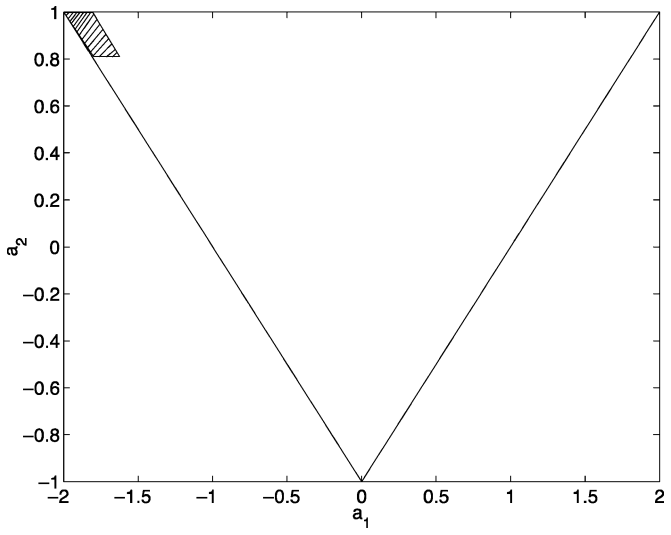


Fig. 3. Plot of the sample space for the coefficients a_1 and a_2 of the AR(2) channel model. The area inside the triangle corresponds to the coefficients that induce stable and minimum-phase process. The shaded area is a constrained region derived from a practical system.

fading channels. It is shown in [19] that the AR(2) coefficients are chosen by

$$a_1 = -2r_d \cos\left(\frac{2\pi\Omega_d}{\sqrt{2}}\right) \quad \text{and} \quad a_2 = r_d^2 \tag{16}$$

where r_d is the pole radius of the AR(2) model, and

$$\Omega_d = f_d T = \frac{v}{\lambda T} \tag{17}$$

which is the normalized maximal Doppler frequency. The variable r_d determines the steepness of the power spectrum of the AR process to closely approximate the Jakes' model, and therefore, r_d is often taken between [0.9, 0.999]. Next, an upper limit on Ω_d is required. Since the carrier frequency is usually predefined in the standards, the upper limit of Ω_d can be obtained from (17) by specifying the upper limits on the mobile velocity and the symbol time (or the lower limits on the symbol rate). From a practical perspective, this upper limit can easily be obtained from the prior knowledge of real-world communication systems. For example, for a system with a carrier frequency of 2 GHz, if we know that the vehicle speeds are less than 75 mi/h and that for all transmissions the symbol rates are greater than 3600 Hz, then Ω_d must be less than 0.062.

Once we determine the bounds on r_d and Ω_d, we can obtain a refined region for a_1 and a_2 from (16), and this region is automatically in the triangle region of a stable AR process. The initial samples of a_1 and a_2 from the refined region can be obtained by first sampling r_d and Ω_d uniformly from the bounded regions and then computed from (16). In Fig. 3, with the shaded lines, we also plot the region corresponding to $r_d \in [0.9, 0.999]$ and $\Omega_d \in [0, 0.1]$. Note that samples of a_1 and a_2 are no longer uniformly distributed in the region, and they reflect better the prior information about these coefficients than the uniform prior. Particularly, we see that no ambiguity exists in this region and that it is much more restricted compared with the triangular region. Therefore, by having such restricted sampling space, we not only resolve the phase ambiguity but greatly increase the efficiency of the detector as well.

C. Resampling Procedure

A resampling procedure [24], [25] must be incorporated in the PFD algorithm to enhance its performance. Recall that a_1 and a_2 are *static* parameters and that no sampling is involved for a_1 and a_2 throughout the implementation. With conventional resampling with time, the particles of a_1 and a_2 degenerate into very few different values. The inability to rejuvenate a_1 and a_2 with the arrival of new observations makes the accuracy of the final estimates greatly dependent on the initial samples. To overcome this drawback, one can rejuvenate the impoverished particles by a Markov chain Monte Carlo move such as a Gibbs sampling move with its invariant distribution being the desired posterior distribution [27]. The method, however, requires considerable memory to store the trajectories that might not be feasible for the communication application. We, therefore, adopted an alternative scheme that combines the auxiliary particle filter [28] and a kernel smoothing technique, which were proposed in [29]. In its implementation, modifications were made to adapt the use of the hybrid importance function. Specifically, when resampling is needed, the proposed procedure is inserted to replace the original particle filtering step. If assuming (a_{1l}, a_{1u}) and (a_{2l}, a_{2u}) are the constrained regions for a_1 and a_2 derived from Section III-B, the detailed algorithm is summarized in Chart 2. It is suggested in [29] that $\alpha = \sqrt{1-h^2}$ and $h^2 = 1-((3\epsilon-1)/2\epsilon)^2$ and that ϵ is a discount factor typically from the set [0.95,0.99]. We want to point out that this smoothing kernel approach is an approximation method that cannot guarantee samples from the right distribution.

Chart 2. PFD with resampling and smoothing kernel

At time t:

For $j = 1, \ldots, N$, compute $\mathbf{m}_t^{(j)} = \alpha \mathbf{a}_{t-1}^{(j)} + (1-\alpha)\bar{\mathbf{a}}_{t-1}$, where $\bar{\mathbf{a}}_{t-1} = \sum_{j=1}^{N} w_{t-1}^{(j)} \mathbf{a}_{t-1}^{(j)}$.

For $j = 1, \cdots, N$

- Sample an auxiliary variable from the set $\{1, \cdots, N\}$ with probability proportional to

$$\lambda_t^{(j)} \propto w_{t-1}^{(j)} p\left(y_t | \mathbf{m}_t^{(j)}, \mathbf{a}_{0:t-1}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, \mathbf{y}_{0:t-1}\right) \tag{18}$$

  and call the sample index k.
- Sample $\mathbf{a}^{(j)} \sim \mathcal{TN}(\mathbf{m}_t^{(k)}, h^2 \mathbf{V}_t | (a_{1l}, a_{1u}), (a_{2l}, a_{2u}))$, where h is a smoothing parameter, $\mathbf{V}_t$ is the weighted sample covariance matrix, and $\mathcal{TN}(.,.|.,.)$ represents a truncated Gaussian distribution with a_1 and a_2 constrained in the region (a_{1l}, a_{1u}) and (a_{2l}, a_{2u}), respectively.

- Sample $s_t^{(j)}$ from the hybrid importance function.
- Evaluate the corresponding weight by

$$w_t^{(j)} \propto \frac{p\left(y_t|\mathbf{a}_t^{(k)}, \mathbf{a}_{0:t}^{(k)}, \mathbf{s}_{0:t-1}^{(k)}, y_{0:t-1}\right)}{p\left(y_t|\mathbf{m}_t^{(j)}, \mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}^{(j)}, y_{0:t-1}\right)}. \tag{19}$$

IV. Discussion

The PFD presented in Section III is general in structure and is flexible to accommodate changes of system requirements and conditions, evoking only small additional computations. In the following, we focus our discussion on coping with a semiblind implementation and time varying AR coefficients.

A. Semiblind Implementation

A semiblind implementation may be desired to reduce the initial transition period in a blind implementation and further combat the phase and coefficient ambiguities. In the semiblind implementation, pilot data are transmitted before or in between the implementation of the blind PFDs. These pilot data are used to obtain extra information about the AR coefficients, which is then fed to the subsequent blind detection to improve its performance. The implementation of the PFD on the pilot data sequence still follows the described procedure. However, since the symbols s_t are known to the receiver during the pilot transmission, no sampling procedure on s_t is needed. Therefore, the weight w_t is solely associated with the samples $\mathbf{a}_t^{(j)}$ and is calculated by

$$\bar{w}_t^{(j)} = w_{t-1}^{(j)} p\left(y_t|\mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t-1}, y_{0:t-1}\right) \tag{20}$$

where $\mathbf{s}_{0:t-1}$ is the sequence of the pilot signals up to $t-1$.

One distinct feature of the semiblind PFD is its ability to allow the exchange of soft information rather than hard decisions of the AR coefficients within the semiblind implementation. The soft information is expressed in terms of the prior distribution and approximated by the particles. Namely, at the beginning of the pilot data sequence, the particles of $\mathbf{a}$ obtained from the blind implementation at the previous time instant are fed into the particle filtering detectors as initial particles. Then, at the end of the sequence, the particles generated from the pilot sequence are used as initial samples from the prior distribution by the subsequent blind PFD. Notice that in a conventional deterministic implementation, a hard decision is made on $\mathbf{a}$ at the end of the pilot sequence and then used as a fixed true value of $\mathbf{a}$ for subsequent detection. Therefore, no update on $\mathbf{a}$ is acquired during the detection, although the new observations contain extra information of $\mathbf{a}$. Compared with the deterministic implementation, the PFD has clear advantages since, through exchange of soft information between the blind and pilot implementation, it can use all the available information to make a decision. This advantage is more evident for time varying AR coefficients.

B. Time-Varying AR Coefficients

The Rayleigh fading assumes that the mobile has a constant speed. However, in reality, the mobile speed changes continuously, which implies that the coefficients of the adopted AR model also change with time. This type of AR model is called a time-varying AR model. The time-varying AR model imposes great mathematical and computational difficulties for conventional solutions. One possible conventional solution is to use semiblind implementation. However, as indicated in Section IV-A, in these implementations, the use of information about $\mathbf{a}$ is very inefficient. Thus, for cases with fast varying coefficients, the performance of the system deteriorates. To improve on it, one must send pilot sequences more frequently, which in turn reduces the bandwidth. The PFD demonstrates its great versatility in that it does not need pilot signals. When a parametric model of coefficient change is available, the implementation of the detector follows the proposed procedure, except that now, at each time t, $\mathbf{a}_t$ is sampled from the prior distributions associated with the assumed model. If the model of parameter variation is not known, one can use the method based on forgetting factors, as suggested in [30].

V. Simulation Results

In this section, the performance of the proposed PFD is studied through experiments. In all the experiments, the transmitted signal was DBPSK modulated. The bit error rate (BER) was computed by transmitting a symbol stream continuously until 300 errors were collected. The initial samples of a_1 and a_2 were drawn from the shaded region in Fig. 4.

In the first experiment, we first studied the effect of using a smoothing kernel in the resampling. In the simulation of a fading channel, the coefficients of the AR(2) model were $a_1 = -1.9305$ and $a_2 = 0.9793$. They reflect a physical scenario of $\Omega_d = 0.05$. This AR process was normalized to have a unit power, and thus, the signal-to-noise ratio (SNR) was obtained by $10\log(1/\sigma^2)$. In Fig. 4, we provide the (BERs) under various SNRs for the PFDs with residual resampling (PFD-RS) [29] and those with smoothing kernel (PFD-SK). For the PFD-RS, we displayed the BERs of two different implementations: the one using 300 trajectories (PFD-RS-300) and another one using 500 trajectories (PFD-RS-500). In the same figure, we also plotted the performance of the MKF with known AR coefficients, which serves as a lower bound. We observe that the increase of sample size from 300 to 500 does not improve the performance of the PFD-RS too much. However, a drastic improvement on the BER is seen for the PFD-SK-200 over the PFD-RS-300 and particularly at high SNRs. There, the BER improvement is almost tenfold. Since only 200 trajectories were maintained by the PFD-SK-200, it is apparent that the use of smoothing kernel in resampling is advantageous. Further, we studied the effect of the sample size N and show in Fig. 4 the performance of using 300 (PFD-SK-300) and 400 (PFD-SK-400) trajectories. We see clear improvement over PFD-SK-200, whereas the PFD-SK-300 and PFD-SK-400 perform similarly. Therefore, 300 trajectories were maintained

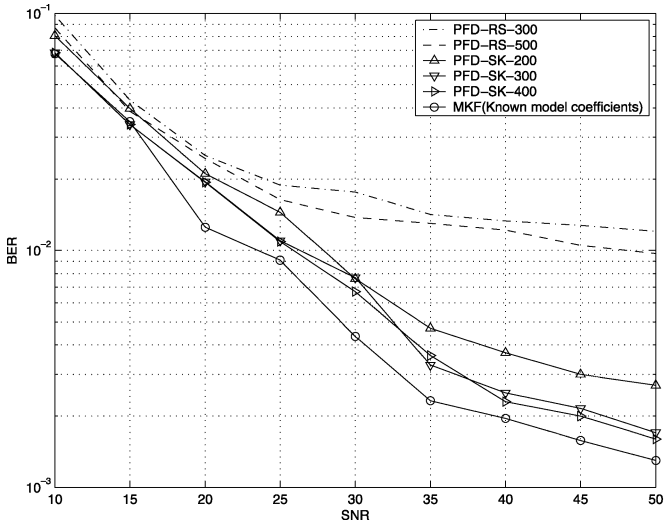


Fig. 4. BERs of the PFD-RSs, the PFD-SK, and the MKF with known model coefficients. $\Omega = 0.05$.

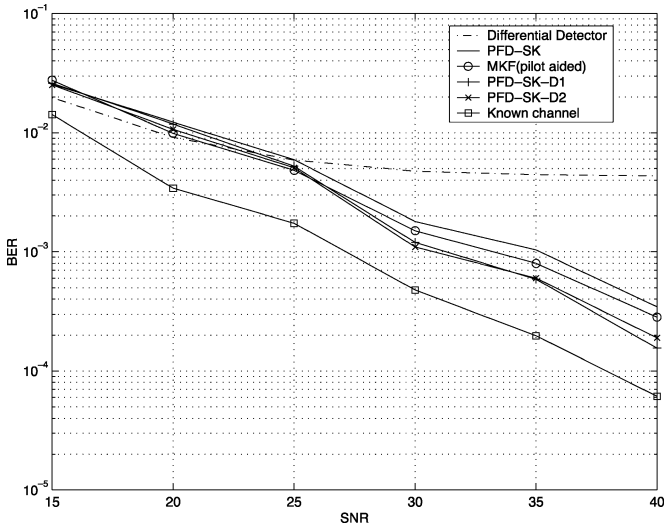


Fig. 5. BERs versus SNR of the PFD-SKs, the pilot-aided MKF, and the detector with known channel. The fading channel was generated by the Jakes' method and $\Omega_d = 0.03$.

for all the PFDs in the following simulations. In addition, the BERs of the PFD-SK-300 and PFD-SK-400 also follow closely the performance of the MKF with known coefficients.

Next, we studied the performance of the PFD-SK under various conditions. To be more realistic, the fading channel was simulated by the Jakes' method with eight oscillators [18], and it was normalized to have a unit power. In addition to the PFD-SK, we examined the performance of the differential detector (DD) [23], the PFD-SK's using 1-step (PFD-SK-D1) and 2-step (PFD-SK-D2) delayed weight, the pilot-aided MKF, and the detector with known channel. In the pilot–aided MKF, 1000 pilot symbols were first used to estimate the AR coefficients (the modified covariance method [31], [32] was employed for the estimation of the AR parameters), and then, MKF was implemented with the estimated coefficients set as the true coefficients. In Fig. 5 and 6, we plot the BERs versus SNR for all the examined methods at $\Omega_d = 0.03$ and $\Omega_d = 0.05$, respectively. From the two plots, first, we observe that unlike the DD, there is no visible error floor for the PFD-SKs. Second, the performance of the PFD-SK approaches that of the pilot-aided MKF and by adding small extra memory, the delayed weight PFD-SKs outperform the pilot-aided MKF. Since no pilot symbols are needed, the PFD-SK's are clearly more bandwidth efficient than the pilot-aided MKF. Third, although there is still some margin for improvement of the BERs of the PFD-SKs, the gap can be reduced by introducing the delayed sample method [15], [33], smoothing [34]–[36], or diversity (for example, fractional sampling) [37].

In Figs. 7 and 8, we demonstrate the performance of the PFD-SKs with the change of Ω_d. The SNR was fixed at 30 dB for Figs. 7 and 40 dB for Fig. 8. Although the DD, PFD-SKs, and pilot-aided MKF have better performance at 40 dB than at 30 dB, their performance deteriorates with the increase of Ω_d. Nonetheless, compared with the DD, the PFD-SK and PFD-SK with delayed weights have around a three and five times gain in BER, respectively. Compared with the pilot-aided MKF,

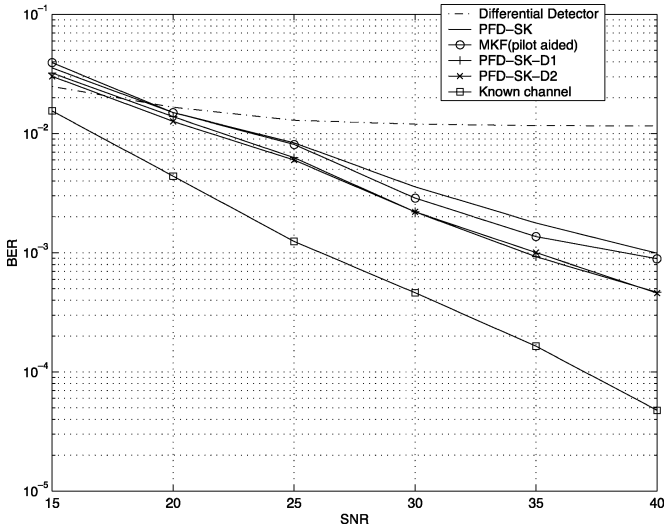


Fig. 6. BERs versus SNR of the PFD-SKs, the pilot-aided MKF, and the detector with known channel. The fading channel was generated by the Jakes' method and $\Omega_d = 0.05$.

similar conclusions can be reached about the PFD-SKs, as in the last experiment.

VI. Conclusions

A particle filtering detector was proposed for blind joint estimation of the parametric channel model coefficients and signal detection. A novel hybrid importance function and the employed MKF method led to efficient implementation of the detector. A connection between the physical interpretation of the AR(2) channel model and the underlying fading channel was used to avoid ambiguity in detection. The structure of the PFD was shown to be versatile to incorporate different implementation requirements. The simulation results demonstrated very good performance of the proposed PFDs.

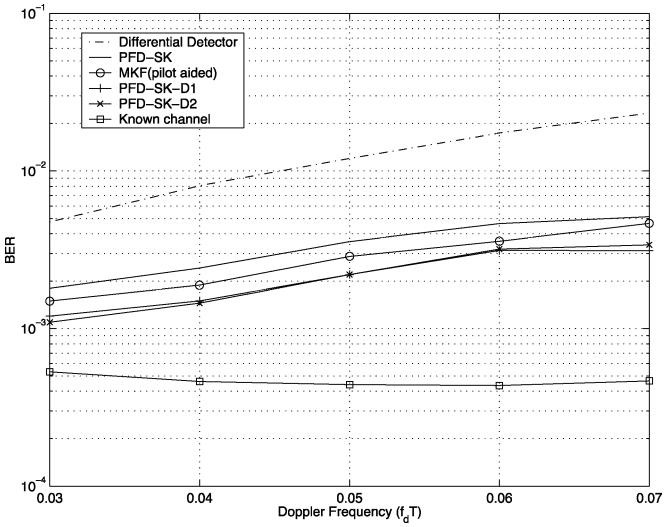


Fig. 7. BERs versus Ω_d for the PFD-SKs, the pilot-aided MKF, and the detector with known channel. The SNR was 30 dB.

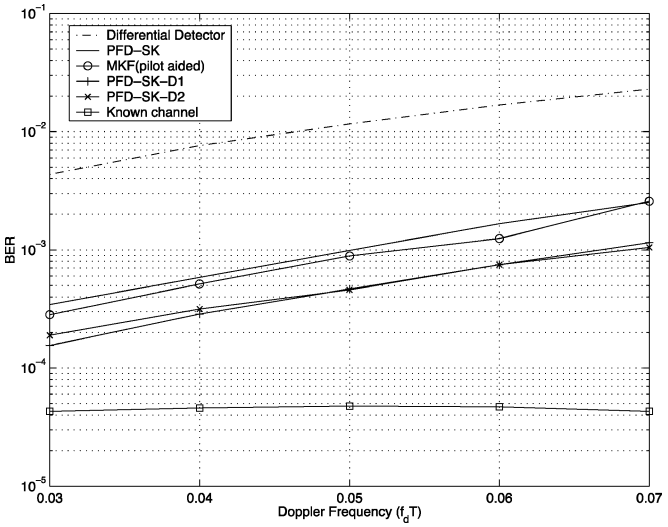


Fig. 8. BERs versus Ω_d for the PFD-SKs, the pilot-aided MKF, and the detector with known channel. The SNR was 40 dB.

APPENDIX A
DERIVATION OF (13)

Note that

$$\begin{aligned} & p\left(y_t|s_t,\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right) \\ &= \int p\left(y_t,h_t|s_t,\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right)dh_t \\ &= \int p\left(y_t|h_t,s_t,\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right) \\ &\qquad \times p\left(h_t|\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right)dh_t. \end{aligned} \tag{21}$$

From the observation equation, we know that

$$p\left(y_t|h_t,s_t,\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right)=\mathcal{N}(h_t s_t,\sigma^2). \tag{22}$$

Furthermore, $p(h_t|\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1})$ is the predictive density of h_t, which can be obtained from the predictive step of the Kalman filter [31], [38], i.e.,

$$p\left(h_t|\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right)=\mathcal{N}\left(\mu_t^{(j)},\Sigma_t^{(j)}\right) \tag{23}$$

where $\mu_t^{(j)}=\mathbf{g}^{\mathsf{T}}\mathbf{D}_t^{(j)}\boldsymbol{\gamma}_{t-1}^{(j)}$, and $\boldsymbol{\Sigma}_t^{(j)}=\mathbf{D}_t^{(j)}\mathbf{C}_{t-1}^{(j)}(\mathbf{D}_t^{(j)})^{\mathsf{T}}+\sigma_{v,t}^{2(j)}\,\mathbf{g}\mathbf{g}^{\mathsf{T}}$ with $\sigma_{v,t}^{2(j)}=P(1+a_2^{(j)})((1-a_2^{(j)})^2-a_1^{(j)}2)/(1-a_2^{(j)})$

$$\mathbf{D}_t^{(j)}=\begin{bmatrix}-a_{1,t}^{(j)} & -a_{2,t}^{(j)}\\ 1 & 0\end{bmatrix}$$

and $\boldsymbol{\gamma}_{t-1}^{(j)}$ and $\mathbf{C}_{t-1}^{(j)}$ are computed from the update steps of the Kalman filter that are expressed at t as

$$\boldsymbol{\gamma}_t^{(j)}=\mathbf{D}_t^{(j)}\boldsymbol{\gamma}_{t-1}^{(j)}+\mathbf{K}_t^{(j)}\left(y_t-\mu_t^{(j)}s_t^{(j)}\right) \tag{24}$$

and

$$\mathbf{C}_t^{(j)}=\left(\mathbf{I}-\mathbf{K}_t^{(j)}\mathbf{g}^{\mathsf{T}}s_t^{(j)}\right)\boldsymbol{\Sigma}_t^{(j)} \tag{25}$$

where $\mathbf{K}_t^{(j)}=\boldsymbol{\Sigma}_t^{(j)}\mathbf{g}c_t^{(j)-1}s_t^{(j)}$. Now, the integration in (21) is readily derived as

$$p\left(y_t|s_t,\mathbf{a}_{0:t}^{(j)},\mathbf{s}_{0:t-1}^{(j)},y_{0:t-1}\right)=\mathcal{N}\left(m_t^{(j)},c_t^{(j)}\right) \tag{26}$$

where $m_t^{(j)}=\mu_t^{(j)}s_t$, and $c_t^{(j)}=\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_t^{(j)}\mathbf{g}+\sigma^2$.

APPENDIX B
PROOF OF THE AMBIGUITY IN A BPSK-MODULATED SYSTEM

For a BPSK-modulated system, define two particles $\{\mathbf{a}_{0:t},\mathbf{s}_{0:t}\}$ and $\{\hat{\mathbf{a}}_{0:t},\hat{\mathbf{s}}_{0:t}\}$, where $a_{1,t}=-\hat{a}_{1,t}$, $a_{2,t}=\hat{a}_{2,t}$ for all t, $s_t=\hat{s}_t$ when t is even and $s_t=-\hat{s}_t$ when t is odd. Here, for convenience, we call the two particles a complement pair. The existence of the ambiguity in detection is established through the following two propositions.

Proposition 1: At any time t, $p(y_t|\mathbf{a}_{0:t},\mathbf{s}_{0:t},y_{0:t-1})=p(y_t|\hat{\mathbf{a}}_{0:t},\hat{\mathbf{s}}_{0:t},y_{0:t-1})$, where the two likelihood functions are calculated in the way, as discussed in Appendix A. Furthermore, we also have $\boldsymbol{\gamma}_{t,1}=\hat{\boldsymbol{\gamma}}_{t,1}$, $\boldsymbol{\gamma}_{t,2}=\hat{\boldsymbol{\gamma}}_{t,2}$, $\mathbf{C}_{t,ii}=\hat{\mathbf{C}}_{t,ii}$, and $\mathbf{C}_{t,ij}=-\hat{\mathbf{C}}_{t,ij}$ for $i\neq j$, where $\boldsymbol{\gamma}_{t,i}$ and $\hat{\boldsymbol{\gamma}}_{t,i}$ are the ith element of $\boldsymbol{\gamma}_t$ and $\hat{\boldsymbol{\gamma}}_t$, and $\mathbf{C}_{t,ij}$ and $\hat{\mathbf{C}}_{t,ij}$ represent the ijth element of $\mathbf{C}_t$ and $\hat{\mathbf{C}}_t$, respectively.

Proof: Since the two likelihoods are Gaussian, it is equivalent to show that $m_t=\hat{m}_t$ and that $c_t=\hat{c}_t$. This is proved by induction.

First, we prove that the proposition is true at $t=0$. At time $t=0$, we have no prior knowledge on the state vector $\mathbf{h}_0$, and the initial condition is chosen as $\boldsymbol{\gamma}_{-1}=\hat{\boldsymbol{\gamma}}_{-1}=0$ and $\mathbf{C}_{-1}=\hat{\mathbf{C}}_{-1}=\sigma_s^2\mathbf{I}$ with σ_s^2 taking a sufficiently large value. The initial conditions indicate $\boldsymbol{\gamma}_{-1,1}=\hat{\boldsymbol{\gamma}}_{t-1,1}$, $\boldsymbol{\gamma}_{-1,2}=-\hat{\boldsymbol{\gamma}}_{-1,2}$, $\mathbf{C}_{-1,ii}=\hat{\mathbf{C}}_{-1,ii}$, and $\mathbf{C}_{-1,ij}=-\hat{\mathbf{C}}_{-1,ij}$ for $i\neq j$ $\mu_0=\hat{\mu}_0=0$. Therefore, we can prove the validity of proposition at $t=0$ in the same vein as what is developed in the following for time t.

At time t, suppose that the proposition holds for time up to $t-1$, and therefore, at $t-1$, we have $\boldsymbol{\gamma}_{t-1,1} = \hat{\boldsymbol{\gamma}}_{t-1,1}$, $\boldsymbol{\gamma}_{t-1,2} = -\hat{\boldsymbol{\gamma}}_{t-1,2}$, $\mathbf{C}_{t-1,ii} = \hat{\mathbf{C}}_{t,ii}$, and $\mathbf{C}_{t,ij} = -\hat{\mathbf{C}}_{t,ij}$ for $i \neq j$. We also assume that $t-1$ is an odd number, which indicates that $s_{t-1} = \hat{s}_{t-1}$. At time t, by assumption, $s_t = -\hat{s}_t$, and $m_t = \mathbf{g}^{\mathsf{T}}\mathbf{D}_t\boldsymbol{\gamma}_{t-1}s_t = -(\boldsymbol{\gamma}_{t-1,1}a_{1,t} + \boldsymbol{\gamma}_{t-1,2}a_{2,t})s_t = \hat{m}_t$. Next, we have from Appendix A that $\boldsymbol{\Sigma}_t$ follows (27), shown at the bottom of the page. Since $\sigma_{v,t}^2 = \hat{\sigma}_{v,t}^2$, we obtain $\Sigma_{t,ii} = \hat{\Sigma}_{t,ii}$ $\forall i = 1, 2$, and $\Sigma_{t,ij} = -\hat{\Sigma}_{t,ij}$ $\forall i \neq j$. It then follows that $c_t = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_t\mathbf{g} + \sigma^2 = \boldsymbol{\Sigma}_{t,11} + \sigma^2 = \hat{c}_t$. Therefore, we obtain $p(y_t|\mathbf{a}_{0:t}, \mathbf{s}_{0:t}, y_{0:t-1}) = p(y_t|\hat{\mathbf{a}}_{0:t}, \hat{\mathbf{s}}_{0:t}, y_{0:t-1})$. Next, we derive $\mathbf{C}_t$ and $\boldsymbol{\gamma}_t$. Since $\mathbf{K}_t = \boldsymbol{\Sigma}_t\mathbf{g}c_ts_t = [\boldsymbol{\Sigma}_{t,11}c_ts_t\ \boldsymbol{\Sigma}_{t,21}c_ts_t]^{\mathsf{T}}$, we find that $\mathbf{K}_{t,1} = -\hat{\mathbf{K}}_{t,1}$ and that $\mathbf{K}_{t,2} = \hat{\mathbf{K}}_{t,2}$. Now, by definition, we first have

$$\begin{aligned}\mathbf{C}_t &= (\mathbf{I} - \mathbf{K}_t\mathbf{g}^{\mathsf{T}}s_t)\boldsymbol{\Sigma}_t \\ &= \begin{bmatrix} (1-\mathbf{K}_{t,1}s_t)\boldsymbol{\Sigma}_{t,11} & (1-\mathbf{K}_{t,1}s_t)\boldsymbol{\Sigma}_{t,12} \\ -\mathbf{K}_{t,2}s_t\boldsymbol{\Sigma}_{t,11} + \boldsymbol{\Sigma}_{t,21} & -\mathbf{K}_{t,2}s_t\boldsymbol{\Sigma}_{t,12} + \boldsymbol{\Sigma}_{t,22} \end{bmatrix}.\end{aligned}$$

Since $\mathbf{K}_{t,1}s_t = \hat{\mathbf{K}}_{t,1}\hat{s}_t$ and $\mathbf{K}_{t,2}s_t = -\hat{\mathbf{K}}_{t,2}\hat{s}_t$, we conclude that $\mathbf{C}_{t,ii} = \hat{\mathbf{C}}_{t,ii}$ $\forall i$, and $\mathbf{C}_{t,ij} = -\hat{\mathbf{C}}_{t,ij}$ for $\forall i \neq j$. Second

$$\begin{aligned}\boldsymbol{\gamma}_t &= \mathbf{D}_t\boldsymbol{\gamma}_{t-1} + \mathbf{K}_t(y_t - m_t) \\ &= \begin{bmatrix} m_ts_t + \mathbf{K}_{t,1}(y_t - m_t) \\ \boldsymbol{\gamma}_{t-1,1} + \mathbf{K}_{t,2}(y_t - m_t) \end{bmatrix}.\end{aligned}$$

If $m_t = \hat{m}_t$, $s_t = -\hat{s}_t$, $\mathbf{K}_{t,1} = -\hat{\mathbf{K}}_{t,1}$, and $\mathbf{K}_{t,2} = \hat{\mathbf{K}}_{t,2}$, it is straightforward to see that $\boldsymbol{\gamma}_{t,1} = \hat{\boldsymbol{\gamma}}_{t,1}$ and that $\boldsymbol{\gamma}_{t,2} = \hat{\boldsymbol{\gamma}}_{t,2}$. ◇

Proposition 2: For a BPSK-modulated system, the particle-filtering detector has the same probability of generating the particles $\{\mathbf{a}_{0:t}, \mathbf{s}_{0:t}\}$ and $\{\hat{\mathbf{a}}_{0:t}, \hat{\mathbf{s}}_{0:t}\}$.

Proof: To prove the proposition, it is equivalent to show

$$p(\mathbf{a}_{0:t}, \mathbf{s}_{0:t}|y_{0:t}) = p(\hat{\mathbf{a}}_{0:t}, \hat{\mathbf{s}}_{0:t}|y_{0:t}) \tag{28}$$

which, again, can be proved by induction.

First, at $t = 0$, we have

$$p(\mathbf{a}_0, s_0|y_0) \propto p(y_t|\mathbf{a}_0, s_0)p(s_t)p(\mathbf{a}_0). \tag{29}$$

From the proof of Proposition 1, we know that $p(y_t|\mathbf{a}_0, s_0) = p(y_t|\hat{\mathbf{a}}_0, \hat{s}_0)$, $p(s_t) = p(\hat{s}_t)$, and $p(\mathbf{a}_0) = p(\hat{\mathbf{a}}_0)$. Thus, the proposition is true at $t = 0$.

Next, suppose that the proposition holds for time up to $t-1$. Thus, at time $t-1$, we have $p(\mathbf{a}_{0:t-1}, \mathbf{s}_{0:t-1}|y_{0:t-1}) = p(\hat{\mathbf{a}}_{0:t-1}, \hat{\mathbf{s}}_{0:t-1}|y_{0:t-1})$, and we want to show that $p(\mathbf{a}_{0:t}, \mathbf{s}_{0:t}|y_{0:t}) = p(\hat{\mathbf{a}}_{0:t}, \hat{\mathbf{s}}_{0:t}|y_{0:t-1})$. First, we factor $p(\mathbf{a}_{0:t}, \mathbf{s}_{0:t}|y_{0:t})$ as

$$\begin{aligned}p(\mathbf{a}_{0:t}, \mathbf{s}_{0:t}|y_{0:t}) &= p(y_t|\mathbf{a}_{0:t}, \mathbf{s}_{0:t}, y_{0:t-1}) \\ &\quad \times p(s_t)p(\mathbf{a}_t|\mathbf{a}_{t-1})\frac{p(\mathbf{a}_{0:t-1}, \mathbf{s}_{0:t-1}|y_{0:t-1})}{p(y_t|y_{0:t-1})}\end{aligned} \tag{30}$$

where the same factorization applies to $p(\hat{\mathbf{a}}_{0:t-1}, \hat{\mathbf{s}}_{0:t-1}|y_{0:t-1})$. Since $p(s_t) = p(\hat{s}_t)$ and $p(\mathbf{a}_t|\mathbf{a}_{t-1}) = p(\hat{\mathbf{a}}_t|\hat{\mathbf{a}}_{t-1}) = 1$, if comparing the factorization of the two posterior distributions, we notice that (28) holds as long as the likelihoods $p(y_t|\mathbf{a}_{0:t}, \mathbf{s}_{0:t}, y_{0:t-1}) = p(y_t|\hat{\mathbf{a}}_{0:t}, \hat{\mathbf{s}}_{0:t}, y_{0:t-1})$, which has been proved true in Proposition 1. As a result, the proposition is true at time t. ◇

Using Proposition 2, we deduce that by time t, if the PFD generated a set of N particles $\{\mathbf{a}_{0:t}^{(j)}, \mathbf{s}_{0:t}^{(j)}\}_{j=1}^N$, it would then be equally likely to produce another set of N particles $\{\hat{\mathbf{a}}_{0:t}^{(j)}, \hat{\mathbf{s}}_{0:t}^{(j)}\}_{j=1}^N$ with each corresponding particle (with the same index) in the two sets being the complement pair. As a result, with the same probability, the PFD could produce either the estimates $\bar{a}_1$, $\bar{a}_2$, and $s_{\mathrm{t,MMSE}}$ obtained from the first set of particles or the estimates $-\bar{a}_1$ and $\bar{a}_2$, $\hat{s}_{\mathrm{t,MMSE}}$ computed by the second set of particles. Furthermore, if $\{s_{\mathrm{t,MMSE}}\}_{t=0}^T$ produces a BER of 0, then $\{\hat{s}_{\mathrm{t,MMSE}}\}_{t=0}^T$ will produce a BER of 0.5. In conclusion, the ambiguity in detection and estimation exists in PDF and the ambiguity could dramatically deteriorate the performance of the detector.

REFERENCES

[1] T. Kailath, "Correlation detection of signals perterbed by a random channel," *IRE Trans. Inform. Theory*, vol. 6, pp. 361–366, June 1960.

[2] J. H. Lodge and M. J. Moher, "Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels," *IEEE Trans. Communications*, vol. 38, pp. 787–794, June 1990.

[3] X. Yu and S. Pasupathy, "Innovations-based MLSE for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 1534–1544, Feb./Mar./Apr. 1995.

[4] G. M. Vitetta and D. P. Taylor, "Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat-fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 2750–2758, Nov. 1995.

[5] G. D. Forney Jr., "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.

[6] R. Haeb and H. Meyr, "A systematic approach to carrier recovery and detection of digitally phase modulated signals on fading channels," *IEEE Trans. Commun.*, vol. 37, pp. 748–754, July 1989.

[7] J. H. Painter and S. C. Gupta, "Recursive ideal observer detection of known M-ary signals in multiplicative and additive Gaussian noise," *IEEE Trans. Commun.*, vol. COMM-21, pp. 948–953, Aug. 1973.

[8] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2079–2094, Sept. 2000.

[9] E. Punskaya, C. Andrieu, A. Doucet, and W. J. Fitzgerald, "Particle filtering for demodulation in fading channels with non-Gaussian additive noise," *IEEE Trans. Commun.*, vol. 49, pp. 579–582, Apr. 2001.

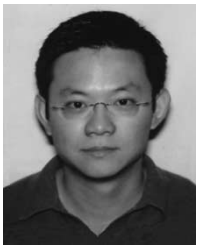
[10] J. Kotecha and P. M. Djurić, "Sequential Monte Carlo sampling detector for Rayleigh fast-fading channels," in *Proc. ICASSP*, Istanbul, Turkey, 2000.

[11] C. Andrieu, A. Doucet, and E. Punskaya, "Sequential Monte Carlo methods for optimal filtering," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. de Freitas, and N. Gordon, Eds. New York: Springer-Verlag, 2001.

[12] J. Kotecha and P. M. Djurić, "Hybrid Monte Carlo—recursive identification algorithms for blind detection over a Rayleigh fading channel," in *Proc. EUSIPCO*, Tampere, Finland, 2000.

$$\begin{aligned}\boldsymbol{\Sigma}_t &= \mathbf{D}_t\mathbf{C}_{t-1}\mathbf{D}_t^{\mathsf{T}} + \sigma_{v,t}^2 \\ &= \begin{bmatrix} a_{1,t}^2\mathbf{C}_{t,11} + a_{1,t}a_{2,t}\mathbf{C}_{t,21} + a_{1,t}a_{2,t}\mathbf{C}_{t,12} + a_{2,t}^2\mathbf{C}_{t,22} + \sigma_{v,t}^2 & -a_{1,t}\mathbf{C}_{t,11} - a_{2,t}\mathbf{C}_{t,21} \\ -a_{1,t}\mathbf{C}_{t,11} - a_{2,t}\mathbf{C}_{t,12} & \mathbf{C}_{t,11} \end{bmatrix}.\end{aligned} \tag{27}$$

[13] Y. Huang and P. M. Djurić, "A hybrid importance function for particle filtering," *IEEE Signal Processing Lett.*, vol. 11, pp. 404–406, Mar. 2004.
[14] R. Chen and J. Liu, "Mixture Kalman filters," *J. R. Statist. Soc. B*, vol. 62, pp. 493–508, 2000.
[15] X. Wang and R. Chen, "Adaptive Bayesian multiuser detection for synchronous CDMA with Gaussian and impulsive noise," *IEEE Trans. Signal Processing*, vol. 47, pp. 2013–2027, July 2000.
[16] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.
[17] N. J. Gordon, A. Doucet, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. Signal Processing*, vol. 49, pp. 613–624, Mar. 2001.
[18] G. L. Stüber, *Principles of Mobile Communication*. Dordrecht, The Netherlands: Kluwer, 1996.
[19] Multiuser detection and channel estimation for flat Rayleigh fading CDMA, H. Y. Wu and A. Duel-Hallen. [Online]. Available: citeseer.nj.nec.com/406 501.html
[20] L. Lindbom, A. Ahlen, M. Sternad, and M. Falkenstrom, "Tracking of time-varying mobile radio channels. II. A case study," *IEEE Trans. Commun.*, vol. 50, pp. 156–167, Jan. 2002.
[21] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Channel estimation and equalization in fading," in *Conf. Rec. Thirty-Third Asilomar Conf. Signals, Syst., Comput.*, vol. 2, 1999, pp. 1159–1163.
[22] L. M. Davis, I. B. Collings, and P. Hoeher, "Joint map equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, pp. 2106–2114, Dec. 2001.
[23] J. G. Proakis, *Digital Communication*, Third ed. New York: McGraw-Hill, 1995.
[24] A. Doucet, J. de Freitas, and N. Gordon, Eds., *Sequential Monte-Carlo Methods in Practice*. New York: Springer-Verlag, 2001.
[25] J. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Statist. Assoc.*, vol. 93, pp. 1032–1044, 1998.
[26] E. R. Beadle and P. M. Djurić, "Uniform random parameter generation of stable minimum phase real ARMA(p,q) processes," *IEEE Signal Processing Lett.*, vol. 4, pp. 259–261, Sept. 1997.
[27] W. R. Gilks and C. Berzuini, "Following a moving target—Monte Carlo inference for dynamic Bayesian models," *J. R. Statist. Soc. B*, no. 63, 2001.
[28] M. Pitt and N. Shephard, "Filtering via simulation: auxiliary particle filter," *J. Amer. Statist. Assoc.*, vol. 94, pp. 590–599, 1999.
[29] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G. De Freitas, and N. J. Gordon, Eds. New York: Springer-Verlag, 2000.
[30] P. M. Djurić, J. Kotecha, F. Esteve, and E. Perret, "Sequential parameter estimation of time-varying non-Gaussian autoregressive processes," *EURASIP J. Applied Signal Process.*, vol. 8, pp. 865–875, 2002.
[31] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.
[32] S. Kay, *Modern Spectrum Estimation: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[33] X. Wang, R. Chen, and D. Guo, "Delayed-pilot sampling for mixture Kalman filter with application in fading channels," *IEEE Trans. Signal Processing*, vol. 50, pp. 241–254, Feb..
[34] T. C. Clapp and S. J. Godsill, "Fixed-lag smoothing using sequential importance sampling," *Bayesian Statist.*, vol. 6, 1998.
[35] A. Doucet, S. J. Godsill, and M. West, "Monte Carlo filtering and smoothing with application to time-varying spectral estimation," in *Proc. ICASSP*, vol. 2, 2000, pp. 11 701–11704.
[36] J. Zhang and P. M. Djurić, "Reduced complexity Monte Carlo smoothing with application to communications," in *Proc. Conf. Inform. Sci. Syst.*. Princeton, NJ, Mar. 2002.
[37] H. Zamiri-Tafarian and S. Pasupathy, "Oversampled blind MLSDE receiver," in *Proc. IEEE Int. Conf. Commun.*, vol. 1, 2000, pp. 382–386.
[38] S. Haykin, *Adaptive Filter Theory*, Third ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.



Yufei Huang (M'03) was born in Shanghai, China, in 1973. He received the B.S. degree in applied electronics from Northwestern Polytechnic University, Xi'an, China, in 1995 and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York (SUNY) at Stony Brook in 1997 and 2001, respectively.

He was a Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, SUNY Stony Brook. He is now with the Department of Electrical Engineering, University of Texas at San Antonio. His current research interests are in the theory of Monte Carlo methods and their applications to array processing and wireless communications.

Peter M. Djurić (SM'99) received the B.S. and M.S. degrees in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1981 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Rhode Island, Kingston, in 1990.

From 1981 to 1986, he was Research Associate with the Institute of Nuclear Sciences, Vinca, Belgrade, Yugoslavia. Since 1990, he has been with the State University of New York at Stony Brook, where he is Professor with the Department of Electrical and Computer Engineering. He works in the area of statistical signal processing, and his primary interests are in the theory of modeling, detection, estimation, and time series analysis and its application to a wide variety of disciplines, including telecommunications, bio-medicine, and power engineering.

Prof. Djuric has served on numerous technical committees for the IEEE and SPIE and has been invited to lecture at universities in the United States and overseas. He is the Area Editor of Special Issues of the IEEE SIGNAL PROCESSING MAGAZINE and Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is also Chair of the IEEE Signal Processing Society Committee on Signal Processing—Theory and Methods and is on the Editorial Board of *Digital Signal Processing*, the *EURASIP Journal on Applied Signal Processing*, and the *EURASIP Journal on Wireless Communications and Networking*. He is a Member of the American Statistical Association and the International Society for Bayesian Analysis.