Joint Blind Symbol Rate Estimation and Data Symbol Detection for Linearly Modulated Signals

Sangwoo Park, Erchin Serpedin, and Khalid Qaraqe

Original scientific paper

***Abstract*—This paper focuses on non-data aided estimation of the symbol rate and detecting the data symbols in linearly modulated signals. A blind oversampling-based signal detector under the circumstance of unknown symbol period is proposed. First, the symbol rate is estimated using the Expectation Maximization (EM) algorithm. However, within the framework of EM algorithm, it is difficult to obtain a closed form for the log-likelihood function and the density function. Therefore, these two functions are approximated in this paper by using the Particle Filter (PF) technique. In addition, a symbol rate estimator that exploits the cyclic correlation information is proposed as an initialization estimator for the EM algorithm. Second, the blind data symbol detector based on the PF algorithm is designed. Since the signal is oversampled at the receiver side, a delayed multi-sampling PF detector is proposed to manage the inter-symbol interference caused by oversampling, and to improve the demodulation performance of the data symbols. In the PF algorithm, the hybrid importance function is used to generate both data samples and channel model coefficients, and the Mixture Kalman Filter (MKF) algorithm is used to marginalize out the fading channel coefficients.**

***Index Terms*—Symbol rate estimation, Data symbol detection, Particle Filter, Cyclostationarity, Expectation Maximization**

I. INTRODUCTION

Recently, non-cooperative communication systems have attracted a lot of attention. Especially for military and civilian applications, many researchers have focused on developing efficient means for electronic interception and identification of RF signals such as automatic modulation classification (AMC) algorithms [1], [2], [8], [11], [16]. For this class of applications, notice that before or after identifying the modulation parameters, estimation of unknown channel parameters represents a major challenge for the successful deployment of AMC systems.

One of the key parameters required for the successful demodulation and decoding of unknown RF linearly modulated signals is that of symbol-rate estimation. After modulation classification, the demodulation step requires accurate symbol rate estimation [13]. Several approaches for symbol-rate estimation have been recently proposed in the literature. A symbol-rate estimator which uses the wavelet transform is suggested in [13]. However, in reference [13], this algorithm is based on the assumption that the transmitted signal has an invariant instantaneous amplitude during each symbol period. This implies that a rectangular pulse shaping filter is used at the transmitter. However, the majority of practical communication systems do not employ a rectangular transmit pulse since it requires a large bandwidth. A more general cyclic correlation (CC) based symbol-rate estimator was proposed in [6] and [12] for arbitrary linearly modulated signals. Even though the CC-based symbol-rate estimator [6] is very powerful for AMC applications since no prior information is required, the performance of the estimator should be improved for efficient demodulation of data symbols and channel tracking.

In addition to symbol-rate estimation, data symbols should be also blindly detected. In many real-world applications, narrowband mobile communication channels are generally modeled as frequency flat Rayleigh fading channels. To estimate the symbol rate, oversampling is used at the receiver side. However, oversampling causes inter-symbol interference (ISI) in the received signal. Therefore, numerous contributions have been reported in the literature for signal detection and channel estimation in the presence of ISI effects. Most of these works rely on techniques such as the maximum likelihood sequence estimation (MLSE) [17], [18]. Since these optimal solutions are based on the Viterbi algorithm and require an additional channel estimation step based on the Kalman Filter (KF) for each possible sequence, they entail huge decision delays and high computational complexity. Moreover, in the conventional MLSE, the metrics of trellis branches are evaluated based on the delayed channel parameter estimates which are then updated according to the detected data. Since the data symbol detection is based on delayed estimates of the channel, this method is not suitable for fast fading channels.

To reduce the complexity of MLSE, suboptimal detectors were proposed such as the per-survival sequence detector [14], [15]. This class of suboptimal detectors is more appropriate for fast fading channels since it avoids delayed channel estimates. However, it has a number of drawbacks. First, it still requires a huge computational complexity since it detects the data symbols based on trellis decoding. Second, it requires a separate channel estimator, which further requires preamble symbols to track the channel.

Recently, novel sequential Monte Carlo algorithms, which jointly estimate the channel and detect the data symbols, have been suggested in [7] and [9]. Without compromising the system model, they approximate the optimal solution using sequential Monte Carlo techniques. However, the assumption

S. Park, E. Serpedin, and K. Qaraqe are with the Texas A & M University, College Station, TX 77843-3128, USA.

This work was supported in part by grants from QNRF-NPRP and QTel.

of known model coefficients is required. In practice, the model coefficients should be estimated in advance. To obtain accurate estimates, a large number of training data is required. Using a blind Particle Filter (PF) detector, Huang et al. [3] suggested an improved algorithm. In [3], the proposed algorithm employs a novel resampling algorithm to prevent the error floors caused by the modeling errors. However, this detector cannot be adopted when the symbol-rate is unknown since the symbol rate estimation generally requires oversampling of the received signal. In addition, the proposed resampling algorithm presents increased computational complexity. Therefore, this algorithm is not suitable for efficient demodulation of data symbols in AMC.

This paper proposes a blind oversampling-based signal detector under the circumstance of unknown symbol period. It consists of two parts: a symbol rate estimator and a symbol detector. The symbol rate is estimated using a combination between a cyclic correlation based approach and the Expectation Maximization (EM) algorithm, a framework which requires oversampling or fractionally sampling (sampling faster than the symbol rate). To reduce the computational complexity, the EM algorithm is simplified by using a Particle Filter (PF) algorithm. Data symbols are detected by using the same PF algorithm based on the oversampled received signal. The oversampling of the received signal improves the performance of data-symbol detectors. In addition, the proposed scheme only requires general resampling steps, which are much simpler than the novel resampling steps proposed in [3]. The PF algorithm employs also a modified hybrid importance function [4] and the Mixture Kalman Filter (MKF) algorithm [9] to reduce its computational complexity. An AR(2) process is used to model the fading channel, and both the AR coefficients and channel coefficients are estimated. Finally, two resampling techniques are adopted and compared in terms of their demodulation performance of data symbols.

The rest of this paper is organized as follows. Dynamic state-space models are proposed in Section I to capture the propagation channel. A novel blind symbol-rate estimator and a data-symbol detector are introduced in Sections III and IV, respectively. In Section V, simulation results are provided to illustrate the performance of the proposed algorithm for efficient demodulation of BPSK modulated signals in Rayleigh fading. Finally, conclusions are mentioned in Section VI.

II. Dynamic Signal Model

We consider the problem of blind (or non-data aided) detection of data symbols assuming a wireless channel environment modeled in terms of a Rayleigh flat fading channel. The Rayleigh flat fading channel is modeled using Jakes' model. Because it is not feasible to directly apply Jakes' model into dynamic state-space models, alternatively, an AR process is used to approximate Jakes' model [5]. The AR(2) process is modeled as

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

where h_t denotes the fading channel coefficient, a_1 and a_2 are the AR model coefficients, and v_t is normally distributed noise with zero mean and σ_v^2 variance [19]. Based on the assumption of unit power fading process, the noise variance σ_v^2 can be expressed as

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

Herein, we only consider linearly modulated signals. We assume M-ary PSK modulated signals, and that a square-root raised cosine filter is used as a shaping filter.

In many references, e.g., [3], [9], the dynamic state-space model of one sample per symbol period was adopted. Based on the known symbol period, the received data signal is sampled at every symbol period. Such an approach not only prevents inter-symbol interference (ISI) but also reduces the computational complexity of resulting decoding algorithms. However, it cannot be adopted when the symbol rate is unknown since the symbol rate estimation generally requires oversampling of the received signal. Therefore, we suggest employing an alternative state-space dynamic model to capture the effects of oversampling.

To estimate the symbol period T, it is necessary to oversample the received signal. If we assume that the sampling period is sufficiently small relative to the symbol period, e.g., $T_s << T$, where T and T_s denote the symbol period and the sampling period, respectively, a dynamic state-space channel model can be constructed assuming multiple samples per symbol period. Herein, the dynamic state-space model is depicted by the following set of equations:

$$\begin{aligned} \text{state equations}: \quad \mathbf{a}_t &= \mathbf{a}_{t-1}, \\ \mathbf{h}_t &= \mathbf{A}_t \mathbf{h}_{t-1} + \mathbf{v}_t, \\ \text{observation equation}: \; y_t &= \mathbf{g}^{\mathrm{T}} \mathbf{h}_t s_t + e_t, \end{aligned} \tag{3}$$

where $\mathbf{h}_t = [h_t, h_{t-1}]^{\mathrm{T}}$, $\mathbf{a}_t = [a_{t,1}, a_{t,2}]^{\mathrm{T}}$, $\mathbf{v}_t = [v_t, 0]^{\mathrm{T}}$, $\mathbf{g} = [1, 0]^{\mathrm{T}}$, and

$$\mathbf{A} = \begin{bmatrix} -a_{t,1} & -a_{t,2} \\ 1 & 0 \end{bmatrix}, \tag{4}$$

$$s_t = \sum_{l=0}^{L-1} b_{\lfloor \frac{t}{\alpha} \rfloor - l} p_{t,l}, \tag{5}$$

where L denotes the number of past symbols that are correlated with the t^{th} sample, $p_{t,l}$ denotes the pulse shaping filter tap, $\lfloor \gamma \rfloor$ denotes the largest integer less than or equal to γ, and $\alpha = T/T_s$. The fading channel taps are represented by h_t, and the AR model coefficients are denoted by $a_{t,1}$ and $a_{t,2}$. The process noise v_t is assumed to be normally distributed with zero mean and σ_v^2 variance. In the observation equation, y_t denotes the received signal, b_t stands for data symbol, and e_t stands for an additive Gaussian noise (AWGN) with zero mean and σ^2 variance. Since the channel is assumed to be stationary, the AR coefficients $a_{t,1}$ and $a_{t,2}$ are considered as static parameters [3].

The Particle Filter algorithm is next developed to blindly detect the data symbols based on the proposed dynamic state-space model.

III. Blind Symbol Detection

As mentioned above, due to oversampling, inter-symbol interference is present. To exploit efficiently the information contained in the received signal and to cope with the resulting inter-symbol interference, a delayed version of the PF algorithm is adopted herein.

First, consider the joint posterior density of transmitted symbols, $b_0, \ldots, b_{\lfloor \frac{t}{\alpha} \rfloor}$, and AR coefficients, $\mathbf{a}_0, \ldots, \mathbf{a}_t$. Using Bayes' rule, the joint posterior density can be expressed as

$$\begin{aligned}
&p(b_{0:\lfloor \frac{t}{\alpha} \rfloor}, \mathbf{a}_{0:t+\Delta_1} | y_{0:t+\Delta_1}) \\
&\propto p(b_{\lfloor \frac{t}{\alpha} \rfloor} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1}) p(b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1} | y_{0:t+\Delta_1}) \\
&\propto p(b_{\lfloor \frac{t}{\alpha} \rfloor} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1}) \\
&\times p(b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{t+\Delta_3+1:t+\Delta_1} | y_{0:t+\Delta_3}) \\
&\propto p(b_{\lfloor \frac{t}{\alpha} \rfloor} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1}) \\
&\times p(\mathbf{a}_{t+\Delta_1}, y_{t+\Delta_1} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1 - 1}, y_{0:t+\Delta_1 - 1}) \cdots \\
&\times p(\mathbf{a}_{t+\Delta_3+1}, y_{t+\Delta_3+1} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_3}, y_{0:t+\Delta_3}) \\
&\times p(b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t-\lceil \alpha \rceil} | y_{0:t-\lceil \alpha \rceil}) \\
&\propto p(b_{\lfloor \frac{t}{\alpha} \rfloor} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1}) \prod_{j=0}^{\Delta_2 - 1} p(\mathbf{a}_{t+\Delta_1 - j} | \mathbf{a}_{t+\Delta_1 - j - 1}) \\
&\times \prod_{j=0}^{\Delta_2 - 1} p(y_{t-j} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t-j}, y_{t-j-1}) \\
&\times p(b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t-\Delta_2} | \mathbf{y}_{0:t-\Delta_2}), \qquad (6)
\end{aligned}$$

where Δ_1 denotes the number of delayed samples, Δ_2 stands for the number of samples per symbol period, and $\Delta_3 = \Delta_1 - \Delta_2$.

The samples are generated from the right hand side of equation (6) which is referred to as a hybrid importance function,

$$p(b_{\lfloor \frac{t}{\alpha} \rfloor} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1}) \prod_{j=0}^{\Delta_2 - 1} p(\mathbf{a}_{t+\Delta_1 - j} | \mathbf{a}_{t+\Delta_1 - j - 1}), \quad (7)$$

with $p(\mathbf{a}_{t+\Delta_1 - j} | \mathbf{a}_{t+\Delta_1 - j - 1}) = \delta(\mathbf{a}_{t+\Delta_1 - j} - \mathbf{a}_{t+\Delta_1 - j - 1})$, $\delta(\cdot)$ being the Dirac delta function, and $p(b_{\lfloor \frac{t}{\alpha} \rfloor} | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1})$ is shown in Table I. The weight of the function is updated via

$$\begin{aligned}
\hat{w}_{\lfloor \frac{t}{\alpha} \rfloor} &\propto w_{\lfloor \frac{t}{\alpha} \rfloor - 1} \\
&\times \prod_{j=0}^{\Delta_2 - 1} p(y_{t+\Delta_1 - j} | b_{0:\lfloor \frac{t}{\alpha} - 1 \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1 - j}, y_{0:t+\Delta_1 - j - 1}).
\end{aligned} \qquad (8)$$

The proposal density function (7) does not include any vector related to the channel taps. Therefore, the channel vector must be marginalized out. This is implemented using the predictive and update steps of the Kalman filter. The details are shown in Table I.

To prevent phase ambiguity, initial AR coefficients are generated via

$$\begin{aligned}
a_1 &= -2 r_d \cos(\frac{2\pi f_d T}{\sqrt{2}}), \\
a_2 &= r_d^2, \qquad (9)
\end{aligned}$$

and

$$f_d = \frac{v}{\lambda}, \qquad (10)$$

where v denotes the speed of the vehicle, λ stands for the carrier wavelength, and r_d represents the pole radius of the AR model and f_d is the maximum Doppler frequency, which are drawn from the regions $[0.9, 0.999]$ and $[0, 0.1]$, respectively. The region of $f_d T$ is decided by considering real-world communication systems. For examples, $f_d T$ must be less than 0.062 if the system assumes 2 GHz carrier frequency, symbol rates are greater than 3600 Hz, and the vehicle speeds are less than 75 miles/h [3].

Having introduced all elements required for the implementation of the PF algorithm, the resulting weighted samples, $b^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor}$ and $w^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor}$, $i = 1, \ldots, N$, approximate $p(b_{\lfloor \frac{t}{\alpha} \rfloor} | y_{0:t})$, and the minimum mean square error (MMSE) estimate is calculated via

$$\hat{b}_{\lfloor \frac{t}{\alpha} \rfloor} = \sum_{i=1}^{N} b^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor} w^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor}. \qquad (11)$$

The resampling step is conducted at the end. However, the general resampling step does not prevent AR coefficients, $\mathbf{a}_t$, from degenerating and assuming very few different values. Huang and Djurić proposed a novel resampling step [3] based on the Auxiliary Particle Filter (APF) and the smoothing kernel approach, which was originally proposed by Liu and West in [10].

Whenever the resampling step is required, instead of the general resampling step, the following procedure is performed. First, the sampled mean and covariance matrix are computed via

$$\begin{aligned}
\bar{\mathbf{a}}_{t-1} &= \sum_{i=1}^{N} w^{(i)}_{t-1} \mathbf{a}^{(i)}_{t-1}, \\
V_{t-1} &= \sum_{i=1}^{N} w^{(i)}_{t-1} (\mathbf{a}^{(i)}_{t-1} - \bar{\mathbf{a}}_{t-1})^2. \qquad (12)
\end{aligned}$$

A new mean vector is defined as $\tilde{\mathbf{a}}^{(i)}_t = \epsilon \mathbf{a}^{(i)}_{t-1} + (1-\epsilon) \bar{\mathbf{a}}_{t-1}$. An auxiliary variable is generated from the index set $\{1, \cdots, N\}$ with the probability proportional to

$$\begin{aligned}
&q(i | y_{0:t+\Delta_1}) \\
&\propto w_{\lfloor \frac{t}{\alpha} \rfloor - 1} \\
&\times \textstyle\prod_{j=0}^{\Delta_2 - 1} p(y_{t+\Delta_1 - j} | b^{(i)}_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \tilde{\mathbf{a}}^{(i)}_{t+\Delta_3+1:t+\Delta_1 - j}, \mathbf{a}^{(i)}_{0:t+\Delta_3}, y_{0:t+\Delta_1 - j - 1}).
\end{aligned} \qquad (13)$$

Now, consider the generated sample index as a new index ξ, and draw the channel model coefficients $\mathbf{a}^{(i)}_{t+\Delta_3+1:t+\Delta_1}$ from

TABLE I
POSTERIOR DENSITY FUNCTION

$$
\begin{aligned}
& p(b_{\lfloor \frac{t}{\alpha} \rfloor} = b_l | b_{0:\lfloor \frac{t}{\alpha} \rfloor - 1}, \mathbf{a}_{0:t+\Delta_1}, y_{0:t+\Delta_1}) \\
\propto & \prod_{j=0}^{\Delta_1+\Delta_2-1} p(y_{t+\Delta_1-j} | b^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor} = b_l, b^{(i)}_{1:\lfloor \frac{t}{\alpha} \rfloor - 1}, y_{1:t+\Delta_1-j-1}) \\
\propto & \prod_{j=0}^{\Delta_1+\Delta_2-1} p(y_{t+\Delta_1-j} | b_f, b^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor} = b_l, b^{(i)}_{1:\lfloor \frac{t}{\alpha} \rfloor - 1}, y_{1:t+\Delta_1-j-1}) \\
\propto & \prod_{j=0}^{\Delta_1+\Delta_2-1} \sum_{\mathbf{b}_f} \int p(y_{t+\Delta_1-j}, h_{t+\Delta_1-j} | \mathbf{b}_f, b^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor} = b_l, b^{(i)}_{1:\lfloor \frac{t}{\alpha} \rfloor - 1}, y_{1:t+\Delta_1-j-1}) dh_{t+\Delta_1-j} \\
\propto & \prod_{j=0}^{\Delta_1+\Delta_2-1} \sum_{\mathbf{b}_f} \int N(h_{t+\Delta_1-j} s^{(i)}_{t+\Delta_1-j,f,l}, \sigma^2) N(\mu^{(i)}_{t+\Delta_1-j,f,l}, \Sigma^{(i)}_{t+\Delta_1-j,f,l}) dh_{t+\Delta_1-j} \\
\propto & \prod_{j=0}^{\Delta_1+\Delta_2-1} \sum_{\mathbf{b}_f} N(\mu^{(i)}_{t+\Delta_1-j,f,l} s^{(i)}_{t+\Delta_1-j,f,l}, c^{(i)}_{t+\Delta_1-j,f,l}),
\end{aligned}
$$

where $\mathbf{b}_f$ denotes future symbols, $\mathbf{b}_f = [b_{\lfloor \frac{t}{\alpha} \rfloor + 1}, b_{\lfloor \frac{t}{\alpha} \rfloor + 2}, \cdots, b_{\lfloor \frac{t+\Delta_1}{\alpha} \rfloor - 1}, b_{\lfloor \frac{t+\Delta_1}{\alpha} \rfloor}]$, $c^{(i)}_{t+\Delta_1-j,f,l} = \mathbf{g}^T \Sigma^{(i)}_{t+\Delta_1-j,f,l} \mathbf{g} + \sigma^2 s^{(i)2}_{t+\Delta_1-j,f,l}$, and all the other parameters are obtained by means of the Kalman filter shown in the Table II.

TABLE II
KALMAN FILTER

1. Time update the channel vector

$$
\begin{aligned}
& \mu^{(i)}_{t+\Delta_1,f,l} = \mathbf{g}^{\mathrm{T}} A^{(i)}_{t+\Delta_1} \boldsymbol{\gamma}^{(i)}_{t+\Delta_1-1,f,l} \\
& \Sigma^{(i)}_{t+\Delta_1,f,l} = A^{(i)}_{t+\Delta_1} \Sigma^{(i)}_{t+\Delta_1-1,f,l} A^{(i)\mathrm{T}}_{t+\Delta_1} + \sigma^{2(i)}_{v,t+\Delta_1} \mathbf{g}\mathbf{g}^{\mathrm{T}}.
\end{aligned}
$$

2. Measurement update the channel vector

$$
\begin{aligned}
& K^{(i)}_{t+\Delta_1,f,l} = \Sigma^{(i)}_{t+\Delta_1,f,l} \mathbf{g} c^{(i)-1}_{t+\Delta_1,f,l} s^{(i)}_{t+\Delta_1,f,l} \\
& \boldsymbol{\gamma}^{(i)}_{t+\Delta_1,f,l} = A^{(i)}_{t+\Delta_1} \boldsymbol{\gamma}^{(i)}_{t+\Delta_1-1,f,l} + K^{(i)}_{t+\Delta_1,f,l} (y_{t+\Delta_1} - \mu^{(i)}_{t+\Delta_1,f,l} s^{(i)}_{t+\Delta_1,f,l}) \\
& C^{(i)}_{t+\Delta_1,f,l} = (I - K^{(i)}_{t+\Delta_1,f,l} \mathbf{g}^{\mathrm{T}} s^{(i)}_{t+\Delta_1,f,l}) \Sigma^{(i)}_{t+\Delta_1,f,l}.
\end{aligned}
$$

the density represented by

$$
\begin{aligned}
& q(\mathbf{a}_{t+\Delta_3+1:t+\Delta_1} | \mathbf{a}^{(\xi)}_{0:t+\Delta_3}) \\
= \; & p(\mathbf{a}_{t+\Delta_1} | \mathbf{a}^{(i)}_{t+\Delta_1-1}) p(\mathbf{a}^{(i)}_{t+\Delta_1-1} | \mathbf{a}^{(i)}_{t+\Delta_1-2}) \cdots \\
\times \; & p(\mathbf{a}^{(i)}_{t+\Delta_3} | \mathbf{a}^{(i)}_{t+\Delta_3-1}) p(\mathbf{a}^{(i)}_{t+\Delta_3+1} | \mathbf{a}^{(i)}_{t+\Delta_3}) \\
= \; & \delta(\mathbf{a}_{t+\Delta_1} - \mathbf{a}^{(i)}_{t+\Delta_1-1}) \delta(\mathbf{a}^{(i)}_{t+\Delta_1-1} - \mathbf{a}^{(i)}_{t+\Delta_1-2}) \cdots \\
\times \; & TN(\mathbf{a}_{t+\Delta_3+1}; \tilde{\mathbf{a}}^{(\xi)}_{t+\Delta_3+1}, h^2 V_{t+\Delta_3+1} | [a_{l1}, a_{u1}], [a_{l2}, a_{u2}]),
\end{aligned} \tag{14}
$$

where $TN(\beta; \gamma_1, \Delta | [a_{l1}, a_{u1}], [a_{l2}, a_{u2}])$ denotes a truncated multivariate normal distribution with the mean γ_1, covariance matrix Δ, and boundaries $[a_{l1}, a_{u1}]$ and $[a_{l2}, a_{u2}]$. Since the channel is assumed to be stationary, the Dirac delta function can be used as the prior function of the channel model coefficients. However, the Dirac delta function makes the algorithm depend on the initial sample values since the previous samples are transferred without any changes. Therefore, at each data symbol $b^{(i)}_t$ drawing, the first Dirac delta function is replaced by the truncated normal distribution to vary the samples. Based on the generated samples $\mathbf{a}^{(i)}_{t+\Delta_1-\Delta_2+1:t+\Delta_1}$, the other samples $b_{\lfloor \frac{t}{\alpha} \rfloor}$ are drawn from the hybrid importance function (7). The new updated weight is also evaluated via

$$
\hat{w}^{(i)}_{\lfloor \frac{t}{\alpha} \rfloor} \propto \frac{\prod_{j=0}^{\Delta_2-1} p(y_{t+\Delta_1-j} | b^{(i)}_{0:\lfloor \frac{t}{\alpha}-1 \rfloor - 1}, \mathbf{a}^{(i)}_{0:t+\Delta_1-j}, y_{0:t+\Delta_1-j+1})}{\prod_{j=0}^{\Delta_2-1} p(y_{t+\Delta_1-j} | b^{(\xi)}_{0:\lfloor \frac{t}{\alpha}-1 \rfloor - 1}, \mathbf{a}^{(\xi)}_{0:t+\Delta_1-j}, y_{0:t+\Delta_1-j+1})}. \tag{15}
$$

IV. SYMBOL PERIOD ESTIMATION

A. Symbol Period Estimation based on EM

The Expectation Maximization (EM) algorithm is adopted to estimate the symbol rate. Based on the channel model, define the vectors

$$
\begin{aligned}
\mathbf{b} &= [b_0, b_1, \cdots, b_{\lfloor \frac{t}{\alpha} \rfloor - 1}, b_{\lfloor \frac{t}{\alpha} \rfloor}], \\
\mathbf{y} &= [y_0, \cdots, y_{t+\Delta_1}], \\
\mathcal{A} &= [\mathbf{a}_0, \cdots, \mathbf{a}_{t+\Delta_1}].
\end{aligned} \tag{16}
$$

Based on the vectors in (16), the E-step in the discrete EM (D-EM) method is implemented through

$$
Q(\alpha) = \int_{\mathcal{A}} \int_{\mathbf{b}} p(\mathbf{b}, \mathcal{A} | \mathbf{y}, \alpha) \log p(\mathbf{y} | \mathbf{b}, \mathcal{A}, \alpha) d\mathbf{b} d\mathcal{A}. \tag{17}
$$

To simplify the Q-function in equation (17), we approximate both the probability density function and log-likelihood function using the Particle Filter (PF) algorithm. The joint probability density and the log-likelihood function are next re-expressed as

$$
\begin{aligned}
p(\mathbf{b}, \mathcal{A} | \mathbf{y}, \alpha) &= p(b_{0:\lfloor \frac{t}{\alpha} \rfloor}, \mathbf{a}_{0:t+\Delta_1} | y_{0:t+\Delta_1}, \alpha), & (18) \\
\log p(\mathbf{y} | \mathbf{b}, \mathcal{A}, \alpha) &= \log p(y_{0:t+\Delta_1} | b_{0:\lfloor \frac{t}{\alpha} \rfloor}, \mathbf{a}_{0:t+\Delta_1}, \alpha). & (19)
\end{aligned}
$$

Based on the Table I, we generate samples $b^{(i)}_{0:\lfloor \frac{t}{\alpha} \rfloor}$ and $\mathbf{a}^{(i)}_{0:t+\Delta_1}$ from (18). The Q-function of the D-EM is then approximated

as

$$\begin{aligned} & Q(\alpha) \\ \approx & \int_{\mathcal{A}} \int_{\mathbf{b}} \sum_{i=1}^{N} w_{\lfloor \frac{t}{\alpha} \rfloor}^{(i)} \delta(b_{0:\lfloor \frac{t}{\alpha} \rfloor} - b_{0:\lfloor \frac{t}{\alpha} \rfloor}^{(i)}) \delta(\mathbf{a}_{0:t+\Delta_1} - \mathbf{a}_{0:t+\Delta_1}^{(i)}) \\ & \times \log p(y_{0:t+\Delta_1} | b_{0:\lfloor \frac{t}{\alpha} \rfloor}, \mathbf{a}_{0:t+\Delta_1}, \alpha) d\mathbf{b} d\mathcal{A} \\ = & \sum_{i=1}^{N} w_{\lfloor \frac{t}{\alpha} \rfloor}^{(i)} \log p(y_{0:t+\Delta_1} | b_{0:\lfloor \frac{t}{\alpha} \rfloor}^{(i)}, \mathbf{a}_{0:t+\Delta_1}^{(i)}, \alpha) \\ = & \sum_{i=1}^{N} w_{\lfloor \frac{t}{\alpha} \rfloor}^{(i)} \log \prod_{j=0}^{t+\Delta_1} p(y_j | b_{0:\frac{t}{\alpha}}^{(i)}, \mathbf{a}_{0:t+\Delta_1}^{(i)}, y_{0:j-1}, \alpha) \\ = & \sum_{i=1}^{N} w_{\lfloor \frac{t}{\alpha} \rfloor}^{(i)} \sum_{j=0}^{t+\Delta_1} \log p(y_j | b_{0:\frac{t}{\alpha}}^{(i)}, \mathbf{a}_{0:t+\Delta_1}^{(i)}, y_{0:j-1}, \alpha) \\ = & \sum_{i=1}^{N} w_{\lfloor \frac{t}{\alpha} \rfloor}^{(i)} \sum_{j=0}^{t+\Delta_1} \log N(\mu_j^{(i)} s_j^{(i)}, c_j^{(i)}). \end{aligned} \tag{20}$$

Finally, the M-step of the D-EM algorithm takes the form

$$\alpha_0^{(n)} = \arg \max_{\alpha \in \mathbb{A}^{(n-1)}} Q(\alpha), \tag{21}$$

where n denotes the number of iterations of the EM algorithm and $\mathbb{A}$ represents a discrete set of possible values for α. The general procedure that we follow is to shrink the range of the discrete set $\mathbb{A}^{(n)}$ as the number of iterations increases. The procedure is diagramed as follows:

$$\begin{array}{c} \mathbb{A}^{(0)} \\ \downarrow \\ \alpha_0^{(1)}, \quad \varepsilon_{(1)} \quad , \eta_{(1)} \\ \downarrow \\ \mathbb{A}^{(1)} \\ \| \\ \{\alpha_0^{(1)} - \eta_{(1)}, \alpha_0^{(1)} - \eta_{(1)} + \varepsilon_{(1)}, \cdots, \alpha_0^{(1)}, \cdots, \alpha_0^{(1)} + \eta_{(1)} - \varepsilon_{(1)}, \alpha_0^{(1)} + \eta_{(1)}\} \\ \downarrow \\ \vdots \\ \downarrow \\ \alpha_0^{(n)} \end{array} \tag{22}$$

where ε and η are small values which satisfy the conditions $\varepsilon_{(1)} > \cdots > \varepsilon_{(n-1)}$ and $\eta_{(1)} > \cdots > \eta_{(n-1)}$, respectively. Given the $(n-1)^{\text{th}}$ discrete set $\mathbb{A}^{(n-1)}$, the estimated oversampling factor $\alpha^{(n)}$ is estimated by the D-EM. The n^{th} discrete set $\mathbb{A}^{(n)}$ consists of the number of $\lfloor 2\eta_{(n)}/\varepsilon_{(n)} \rfloor$ elements neighboring $\alpha^{(n)}$. For example, when $\alpha^{(n)} = 5.2$, $\eta_{(n)} = 0.07$, and $\varepsilon_{(n)} = 0.01$, then

$$\mathbb{A}^{(n)} = \{5.13, 5.14, \ldots, 5.26, 5.27\}. \tag{23}$$

To represent the $(n+1)^{\text{th}}$ discrete set $\mathbb{A}^{(n+1)}$, we choose values for $\eta_{(n+1)}$ and $\varepsilon_{(n+1)}$ smaller than $\eta_{(n)}$ and $\varepsilon_{(n)}$, respectively, and repeat the process until the convergence is achieved. After certain iterations, we finally obtain an accurate estimate $\hat{\alpha}$.

B. Initial Symbol Period Estimation

In the previous section, we have discussed the symbol rate estimator using the discrete EM (D-EM) algorithm. The D-EM algorithm requires an initial finite set that will be obtained by using the cyclic correlation based symbol-rate estimator. The cyclic correlation based symbol rate estimator is suitable as an initialization estimator because it only requires a sufficiently small sampling period so that $T_s < T/4$ [6]. According to [6] and [12], the initial estimate can be obtained via

$$\hat{p}_0 = \arg \max_{f_k \in I} \hat{\mathbf{C}}(f_k)^* \hat{\mathbf{C}}(f_k), \tag{24}$$

where $\hat{\mathbf{C}}(f_k)$ stands for the vector of cyclic correlations (see [6], [12] for more details). There is a reciprocal relation between the oversampling parameter α_0 and the cyclic frequency p_0. Therefore, the estimate of oversampling factor can be represented by

$$\hat{\alpha_0} = \frac{1}{\hat{p}_0}. \tag{25}$$

For more efficient estimation, based on equation (24), the symbol-rate estimator is reformulated as

$$\begin{aligned} \hat{p}_{0,j} &= \arg \max_{f_k \in I_j} \hat{\mathbf{C}}(f_k)^* \hat{\mathbf{C}}(f_k), \\ \hat{\alpha}_{0,j} &= \frac{1}{\hat{p}_{0,j}}, \end{aligned} \tag{26}$$

where $j = 1, \cdots, J$, and J stands for the number of searching sub-intervals. The searching interval I should be divided into several sub-intervals, $I_1, \cdots, I_J$, and each local maximum value should be selected from the sub-intervals. The selected local maximum values consist of the initial finite set $\mathbb{A}^{(0)}$, i.e.,

$$\mathbb{A}^{(0)} = \{\hat{\alpha}_{0,1}, \cdots, \hat{\alpha}_{0,J}\}. \tag{27}$$

V. Simulation Results

In this section, the performance of the proposed algorithm is illustrated through computer simulations. In all computer simulations, a Rayleigh flat fading channel, BPSK modulation with unit power, and a square-root raised cosine pulse shaping filter with roll-off factor ρ are assumed. In addition, all transmitted data symbols are differentially encoded to prevent phase ambiguities. The signal to noise ratio (SNR) is calculated as the averaged received SNR.

In the first computer simulation, we compared the BER performance of the multiple samples per symbol period signal data detector (MSSD) to the single sample per symbol period signal data detector (SSSD). As shown in Fig. 1, the MSSD improves the BER performance much more than the SSSD. Based on the PF with general resampling, MSSD eliminates the visible error floor which is exhibited by the SSSD. The performance gain is much larger at high SNR since the MSSD tracks the channel much better, and the overall performance is limited by the channel fading.

In the Fig. 2, the BER performances of each method, namely Mixture Kalman Filter (MKF) with known channel model coefficients, Particle Filter with general Resampling

(PF-RS), and Particle Filter with Smoothing Kernel (PF-SK), are plotted. When we oversample the received signal, the gain due to the smoothing kernel is negligible. Therefore, using PF-RS, the complexity caused by the smoothing kernel method can be reduced. Both PF-RS and PF-SK show better performance than the Dual Kalman Filter (DKF) method. To show the lower bound, the performance of the MKF with known channel model coefficients is also presented. In the Fig.

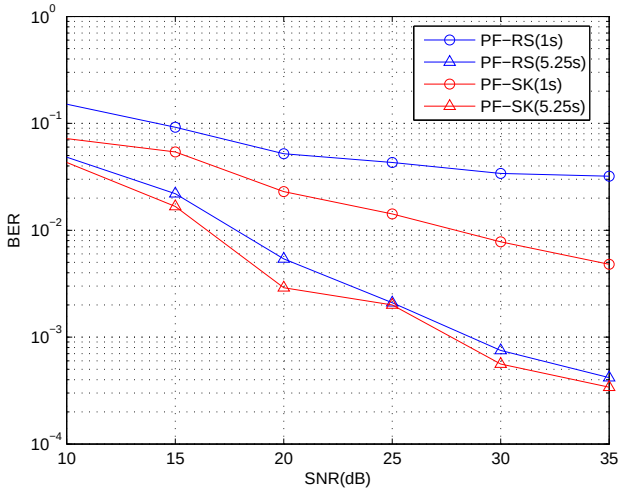

Fig. 1. BERs of PF-SK and PF-RS with one sample per symbol period and 5.25 samples per symbol period (f_dT=0.05, $\alpha = 5.25$, $\rho = 0.7$).

3, according to the number of particles that are considered, the BER performances are compared. As the number of particles increases, the BER performance is improved. Moreover, the PF-RS algorithm shows better performance than the DKF.

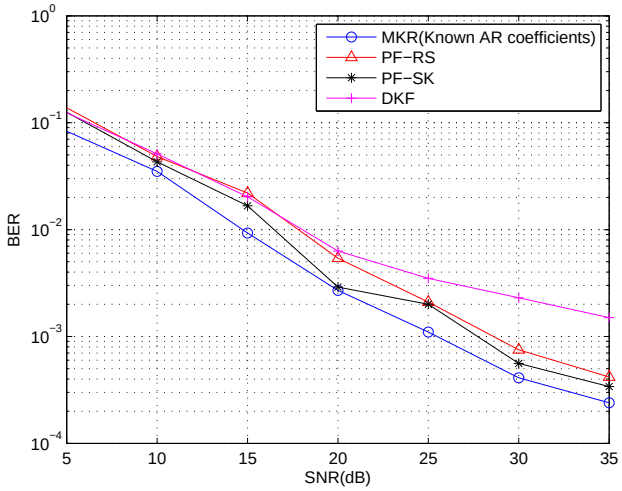

Fig. 2. BERs of PF-SK, PF-RS, and MKF with known AR coefficients with 50 particles, f_dT=0.05, $\alpha = 5.25$, and , $\rho = 0.7$.

In Tables III and IV, the performance of the symbol rate estimator based on PF-RS is compared with the classical

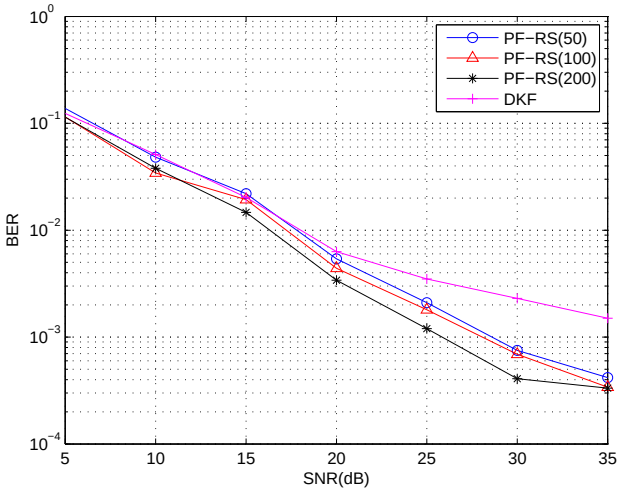

Fig. 3. BERs of the PF-RS with 50, 100, and 200 particles (f_dT=0.05, $\alpha = 5.25$, $\rho = 0.7$).

symbol rate estimator based on cyclic correlation (CC). The estimated symbol rate is considered up to the accuracy level of 10^{-2}. Namely, if the estimated symbol rate lies in the interval $[5.25, 5.26)$, then we consider the symbol rate is correctly estimated. As shown in the Tables III and IV, the proposed algorithm improve the performance a lot in both SNRs.

TABLE III
ESTIMATOR PERFORMANCE COMPARISON BETWEEN PF AND CC
($\alpha = 5.25$, 50 PARTICLES, 10DB)

	PF($\rho = 0.5$)	CC($\rho = 0.5$)
Performance(%)	99.99	24.29
	PF($\rho = 0.7$)	CC($\rho = 0.7$)
Performance(%)	100.00	75.03

TABLE IV
ESTIMATOR PERFORMANCE COMPARISON BETWEEN PF AND CC
($\alpha = 5.25$, 50 PARTICLES, 30DB)

	PF($\rho = 0.5$)	CC($\rho = 0.5$)
Performance(%)	100.00	28.03
	PF($\rho = 0.7$)	CC($\rho = 0.7$)
Performance(%)	100.00	75.92

VI. CONCLUSIONS

Recently, non-cooperative communication systems have attracted a lot of attention. Numerous researchers have focused on systems such as automatic modulation classification that is encountered in both military as well as civilian applications. The importance of blind estimation of the channel parameters and blind detection of the data symbols also comes from the increasing attention given to AMC applications.

In this paper, novel symbol rate estimators with improved performance compared to the estimator based on cyclic correlation were proposed. The EM algorithm, which is used in the symbol rate estimator, is simplified and made tractable by using the PF algorithm. A delayed oversampling based

data symbol detector is also proposed under the modeling framework of Rayleigh flat fading channels. Using the delayed oversampling data symbol detector, the performance of the data symbol detector is improved compared to the classical blind PF detector. Moreover, the general resampling technique, which is very simple, can be adopted since this detector reduces the effect of the AR coefficient estimation errors. Finally, since both the symbol rate estimator and data symbol detector rely on the same PF algorithm, the resulting algorithm presents low computational complexity.

REFERENCES

[1] O. A. Dobre, A. Abdi, Y. Bar-Ness, W. Su: *Cyclostationarity-Based Blind Classification of Analog and Digital Modulations*, Military Communications Conference, (2006) 1-7.

[2] O. A. Dobre, A. Abdi, Y. Bar-Ness, W. Su: *Blind modulation classification: a concept whose time has come*, Advances in Wired and Wireless Communication, IEEE/Sarnoff Symposium, (2005) 223-228.

[3] Y. Huang, P. M. Djurić: *A Blind Particle Filtering Detector of Signals Transmitted Over Flat Fading Channels*, IEEE Trans. Signal Processing, 52(7) (2004).

[4] Y. Huang, P. M. Djurić, *A hybrid importance function for particle filtering*, IEEE Signal Processing Letter, 11 (2004) 404-406.

[5] L. Lindbom, A. Ahlen, M. Sternad, M. Falkenstrom: *Tracking of time-varying mobile radio channels. II. A case study*, IEEE Trans. Commun., 50 (2002) 156-167.

[6] P. Ciblat, P. Loubaton, E. Serpedin, G. B. Giannakis: *Asymptotic Analysis of Blind Cyclic Correlation-Based Symbol-Rate Estimators*, IEEE Trans. Information Theory, 48(7) (2002).

[7] E. Punskaya, C. Andrieu, A. Doucet, W. J. Fitzgerald: *Particle filtering for demodulation in fading channels with non-Gaussian additive noise*, IEEE Trans. Commun., 49 (2001) 579-582.

[8] C. M. Spooner: *On the utility of sixth-order cyclic cumulants for RF signal classification*, in Proc. ASILOMAR, (2001) 890-897.

[9] R. Chen, X. Wang, J. S. Liu: *Adaptive Joint Detection and Decoding in Flat-Fading Channels via Mixture Kalman Filtering*, IEEE Trans. Information Theory, 46(6) (2000).

[10] J. Liu, M. West: "Combined parameter and state estimation in simulation based filtering", in: A. Doucet, J. F. G. De Freitas, and N. J. Gordon (Ed.), Sequential Monte Carlo Methods in Practice, Springer-Verlag, New York, 2000, pp. 197-224.

[11] C. M. Spooner, W. A. Brown, G. K. Young: *Automatic radio-frquency environment analysis*, in Proc. ASILOMAR, (2000) 1181-1186.

[12] L. Mazet, P. Loubaton: *Cyclic correlation based symbol rate estimation*, Signals, Systems, and Computers, 2(24-27) (1999) 1008-1012.

[13] Y. T. Chan, J. W. Plews, K. C. Ho: *Symbol Rate Estimation by The Wavelet Transform*, IEEE ISCAS'97, 1 (1997) 177-180.

[14] G. M. Vitetta, D. P. Taylor: *Maximum likelihood decoding of uncoded and coded PSK signal sequences transmitted over Rayleigh flat fading channels*, IEEE Trans. Commun., 43 (1995) 2750-2758.

[15] X. Yu, S. Pasupathy: *Innovations-based MLSE for Rayleigh fading channels*, IEEE Trans. Commun., 43 (1995) 1534-1544.

[16] C. M. Spooner: *Classification of cochannel communication signals using cyclic cumulants*, in Proc. ASILOMAR, (1995) 531-536.

[17] D. Makrakis, P. T. Mathiopoulos, D. P. Bouras: *Optimal decoding of coded PSK and QAM signals in correlated fast fading channels and AWGN: A combined envelop, multiple differential and coherent detection approach*, IEEE Trans. Commun., 42 (1994) 63-75.

[18] J. H. Lodge, M. L. Moher, *Maximum likelihood sequence estimation of CPM signals transmitted over Rayleigh flat-fading channels*, IEEE Trans. Commun., 38 (1990) 787-794.

[19] H. Y. Wu, A. Duel-Hallen, *Multiuser detection and channel estimation for flat Rayleigh fading CDMA*, [Online]. Available: citeseer.nj.nec.com/406 501.html.

Sangwoo Park is currently working towards obtaining his PhD degree in Electrical Engineering from Texas A&M University, College Station, TX. His research interests lie in the fields of signal processing and wireless communications.

Erchin Serpedin received received the specialization degree in signal processing and transmission of information from Ecole Superiéure D'Electricité, Paris, France, in 1992, the M.Sc. degree from Georgia Institute of Technology, Atlanta, in 1992, and the Ph.D. degree in Electrical Engineering from the University of Virginia, Charlottesville, VA, in January 1999. In July 1999, he joined Texas A&M University in College Station, where he currently holds the position of associate professor. His research interests lie in the areas of signal processing and telecommunications. He received the NSF Career Award in 2001, the Outstanding Faculty Award in 2004, the TEES Award in 2005, and several best paper awards. He served as an associate editor for the IEEE Communications Letters, IEEE Transactions on Signal Processing, IEEE Signal Processing Letters, IEEE Transactions on Wireless Communications and EURASIP Journal on Advances in Signal Processing. Dr. Serpedin served also as a technical co-chair of the Communications Theory Symposium at Globecom 2006 Conference, and for the VTC Fall 2006: Wireless Access Track Symposium. He is currently serving as associate editor for the journals Signal Processing-Elsevier, IEEE Transactions on Communications, IEEE Transactions on Information Theory, EURASIP Journal on Advances in Signal Processing, and EURASIP Journal on Bioinformatics and Systems Biology. Dr. Serpedin is also the author of the research monograph: Synchronization in Wireless Sensor Networks, published in August 2009 by Cambridge University Press.

Khalid Qaraqe (M'97-S'00) was born in Bethlehem. Dr Qaraqe received the B.S. degree in EE from the University of Technology, Baghdad in 1986, with honors. He received the M.S. degree in EE from the University of Jordan, Jordan, in 1989, and he earned his Ph.D. degree in EE from Texas A&M University, College Station, TX, in 1997. From 1989 to 2004 Dr Qaraqe has held a variety positions in many companies and he has over 12 years of experience in the telecommunication industry. Dr Qaraqe has worked for Qualcomm, Enad Design Systems, Cadence Design Systems/Tality Corporation, STC, SBC and Ericsson. He has worked on numerous GSM, CDMA, WCDMA projects and has experience in product development, design, deployments, testing and integration. Dr Qaraqe joined the department of Electrical Engineering of Texas A&M University at Qatar, in July 2004, where he is now a senior associate professor. Dr Qaraqe research interests include communication theory and its application to design and performance, analysis of cellular systems and indoor communication systems. Particular interests are in the development of WCDMA and broadband wireless communications and diversity techniques.