14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

BLIND PHASE NOISE ESTIMATION AND DATA DETECTION BASED ON SMC TECHNIQUE AND UNSCENTED FILTERING

Erdal Panayırcı[†], *Hakan A. Çırpan*[‡], *Marc Moeneclaey*[§], *and Nele Noels*[§]

[†]Department of Electronics Engineering, Kadir Has University, Cibali 34230, Istanbul, Turkey
[‡]Department of Electrical Engineering, Istanbul University, Avcilar 34320, Istanbul, Turkey
[§]TELIN/DIGCOM Department, Ghent University, B9000 Gent, Belgium

ABSTRACT

In this paper, a computationally efficient algorithm is presented for tracing phase noise with linear drift and blind data detection jointly, based on a sequential Monte Carlo(SMC) method. Tracing of phase noise is achieved by Kalman filter and the nonlinearity of the observation process is taken care of by unscented filter rather that using extended Kalman technique. On the other hand,SMC method treats the transmitted symbols as " missing data" and draw samples sequentially of them based on the observed signal samples up to time t. This way, the Bayesian estimates of the phase noise and the incoming data are obtained through these samples, sequentially drawn, together with their importance weights. The proposed receiver structure is seen to be ideally suited for high-speed parallel implementation using VLSI technology.

1. INTRODUCTION

The problem of carrier phase synchronization is of great importance in coherent digital communication systems. A considerable amount of research has been carried out for data detection in the presence of the time-varying phase noise as well as the fixed phase offset [1]. Estimating the phase offset and detecting the data jointly by maximum likelihood (ML) technique does not seem to be analytically tractable. Even if the likelihood function can be evaluated offline, however, it is invariably a nonlinear function of the parameter to be estimated, which makes the maximization step (which must be performed in real-time) computationally infeasible. A number of suboptimal algorithms have thus been proposed, most of which employ a two-stage receiver structure with a phase noise estimation stage followed by the data detection [2]. Phase synchronization is typically implemented by a decision directed(or data aided) or non-decision directed (or non-data aided). Decision directed schemes depend on availability of reliably detected symbol for obtaining the phase estimate, and therefore, they usually require transmission of pilot or training data. However, in applications where bandwidth is the most precious resource, training data can significantly reduce the overall system capacity. Thus blind or non-data aided techniques become an attractive alternative [3, 4].

Unlike data-aided techniques, non-data-aided methods do not require knowledge of the transmitted data, and instead, they exploit statistics of digital transmitted signal. ML estimation techniques can also be used in non-decision-directed methods if the symbols transmitted are treated as random variables with known statistics so that the likelihood function can be averaged over the data sequence received. Unfortunately, except for few simple cases, this averaging process is mathematically impracticable and it can be obtained only by some approximations which are valid only either at high or low SNR values [5].

On the other hand, in order to provide an implementable solution, recently there have been a substantial amount of work on iterative formulation of the parameter estimation problem based on the Expectation-Maximization (EM) technique [6]. It is known that the EM algorithm derives iteratively and converges to the true ML estimation of these unknown parameters. The main drawbacks of this approach are that the algorithm is sensitive to the initial starting values chosen for the parameters, it does not necessarily converge to the global extremum and the convergence can be slow. Furthermore, in situation where the posterior distribution must be constantly updated with arrival of the new data with missing parts, EM algorithm can be highly inefficient, because the whole iteration process must be redone with additional data. The sequential Monte Carlo(SMC) methodology [7] that has emerged in the field of statistics and engineering has shown great promise to solve such problems. This technique can approximate the optimal solution directly without compromising the system model. Additionally, the decision made at time t does not depend on any decisions made previously, and thus, no error is propagated in their implementation. More importantly, the SMC yields a fully blind algorithm and allows for both Gaussian and non-Gaussian ambient noise as well as high-speed parallel implementations. Furthermore, the tracking the time-varying phase noise and the data detection are naturally integrated. The algorithm is self-adaptive and no training/pilot symbols or decision feedback are needed [8].

The main objective of this paper is to investigate the use of the SMC method to the problem of jointly detecting the data and tracking the phase noise with linear drift in the presence of white Gaussian noise. The algorithm is based on

This research has been conducted within the NEWCOM Network of Excellence in Wireless Communications funded through the EC 6th Framework Programme. This work was also supported in part by the Turkish Scientific and Technical Research Institute(TUBITAK) under Grant 104E166 and the Research Fund of the University of Istanbul under Project: UDP-732/05052006.

a Bayesian formulation. The tracking of phase noise is implemented by a Kalman filter algorithm. Rather than applying the extended Kalman algorithm, an *unscented filter* technique is employed to handle nonlinearity of the observation process. On the other hand, the basic idea SMC method is to treat the transmitted symbols as "missing data" and to sequentially draw samples of them based on the current observation and computing appropriate importance sampling weights. The technique does not require iterations and updating with new data can be done cheaply.

2. SYSTEM DESCRIPTION

We consider a digital communication system in the presence of random phase noise and the additive Gaussian noise. The input binary information bit d_t are encoded using some channel code, resulting in a code bit stream b_t. The code bits are passed to a symbol mapper, yielding complex data symbols s_t, which take values from a finite alphabet set $A = \{a_1, a_2, \cdots, a_{|A|}\}$, where $|A|$ represents the cardinality of the set A. Each data symbol is then transmitted through a channel whose input-output relationship is given by

$$y_t = s_t e^{i\theta_t} + n_t, \qquad t = 0, 1, \cdots \tag{1}$$

where y_t, s_t, θ_t, are the received signal, the transmitted symbols and the phase noise, respectively, and n_t the additive complex Gaussian noise with mean zero and the variance $\sigma_n^2 = E[|n_t|^2]$. The phase noise process θ_t at tth sampling instant is defined as a Wiener process determined as

$$\begin{aligned} \theta_t &= \theta_{t-1} + u_t, \qquad t = 1, 2, \cdots \\ \theta_0 &\sim \text{uniform}(-\pi, +\pi) \end{aligned} \tag{2}$$

where $\{u_t\}$ is a sequence of independent and identically distributed (i.i.d.) zero-mean random variables with variance equal to σ_u^2. Note that as Wiener phase noise is the accumulation of white noise, its variance increase linearly with t. It is assumed that u_t and n_t are independent.

Our main objective is to solve the problem of online detection of the symbols s_t and estimation of the phase noise θ_t, completely blindly, based on the received signals up to time t, $\{y_i\}_{i=0}^t$. Defining the vectors, $\boldsymbol{S}_t = [s_0, s_1, \cdots, s_t]^T$, $\boldsymbol{Y_t} = [y_0, y_1, \cdots y_t]^T$, $\boldsymbol{\theta}_t = [\theta_0, \theta_1, \cdots, \theta_t]^T$, the the problem may be formulated by making Bayesian inference with respect to the posterior distribution

$$\begin{aligned} p(\boldsymbol{\theta}_t, \boldsymbol{S}_t|\boldsymbol{Y}_t) &\propto p(\theta_0)p(\boldsymbol{S}_t)p(y_0|\theta_0, s_0)\prod_{j=1}^{t} p(\theta_j|\theta_{j-1})p(y_j|\theta_j, s_j) \\ &\propto p(\theta_0)p(\boldsymbol{S}_t)\exp\left(-\frac{1}{\sigma^2}|y_0 - s_0 e^{j\theta_0}|^2\right) \\ &\times \prod_{j=1}^{t}\exp\left(-\frac{1}{\sigma_u^2}(\theta_j - \theta_{j-1})^2 - \frac{1}{\sigma^2}|y_j - s_j e^{j\theta_j}|^2\right) . \end{aligned}$$

Although this joint distribution can be written out explicitly up to a normalizing constant, the computation of the corresponding marginal joint distributions $p(s_t, \theta_t|\boldsymbol{Y}_t)$, necessary for online joint symbol detection and phase noise estimation involve very high dimensional integration. Therefore, the task is mathematically infeasible in practice. The Gibbs samples [9] is a Monte Carlo method for overcoming this difficulty. However it is not an adaptive procedure and has difficulty dealing with sequentially observed data. With new data coming the whole computation must be repeated to incorporate new information. we now present an adaptive blind algorithm for the joint symbol detection and the phase noise estimation which is based on a Bayesian formulation of the problem called Sequential Monte Carlo(SMC) method first developed by [9].

Consider the case of uncoded system, where the symbols are assumed to independent and identically distributed, i.e.,

$$P(s_t = a_i|\boldsymbol{S}_{t-1}) = P(s_t = a_i), a_i \in A \quad . \tag{3}$$

For simplicity the symbols are chosen from a QPSK constellation.When no prior information about the symbols is available, the symbols are assumed to take each possible value in A with equal probability, i.e., $P(s_t = a_i) = 1/|A|$. Since we are interested in jointly estimating the symbol s_t and the phase noise θ_t, at time t based on the observation $\boldsymbol{Y}_t$, the Bayes solution requires the posterior distribution

$$p(s_t, \theta_t|\boldsymbol{Y}_t) = \int p(\theta_t|\boldsymbol{Y}_t, \boldsymbol{S}_t)p(\boldsymbol{S}_t|\boldsymbol{Y}_t)d\boldsymbol{S}_{t-1}. \tag{4}$$

Note that with a given $\boldsymbol{S}_t$, the nonlinear (Kalman filter) model (1), (2) can be solved using the extended Kalman filter (EKF) technique [10] by linearizing the observation equation (1) [8]. The EKF is the most widely used estimation algorithm for nonlinear systems. However, the long past experience has shown that it is difficult to implement, difficult to tune, and only reliable for systems that are almost linear on the time scale of the updates. Many of these difficulties are mainly due to the linearization process inherent in the EKF technique. To overcome this limitation, the *unscented filtering (UF)* technique was developed as a method of propagate mean and covariance information through nonlinear transformation. It was shown that it is more accurate, easier to implement and uses the same order of calculation as linearization. Therefore, we apply in this paper the UF technique for solving the Kalman filtering part of our problem. However, before the details of the Kalman algorithm is given, the UF technique is explained briefly in the following section.

3. THE UNSCENTED FILTERING

The Unscented Filtering is a technique for calculating the mean and covariance of a random variable which undergoes a nonlinear transformation. The details of technique can be found in [14] and is summarized as follows: Suppose that $\boldsymbol{x}$ is an $n \times 1$ dimensional random vector with mean $\boldsymbol{\mu_x}$ and covariance $\mathbf{P_{xx}}$. A second random variable , $\boldsymbol{y}$ is related to $\boldsymbol{x}$ through the nonlinear function

$$\boldsymbol{y} = \boldsymbol{\psi}(\boldsymbol{x}).$$

The mean $\boldsymbol{\mu_y}$ and covariance $\mathbf{P_{yy}}$ of $\boldsymbol{y}$ can be calculated as follows: The n-dimensional random vector $\boldsymbol{x}$ with mean $\boldsymbol{\mu_x}$ and covariance $\mathbf{P_{xx}}$ is approximated by $2n + 1$ points $\boldsymbol{\chi}_i$, called the *sigma-points*, and the weights W_i given by

$$\boldsymbol{\chi}_0 = \boldsymbol{\mu_x}, \qquad W_0 = \kappa/(n + \kappa)$$

$$\begin{aligned}\boldsymbol{\chi}_i &= \boldsymbol{\mu_x} + (\sqrt{(n+\kappa)\mathbf{P}_{\boldsymbol{xx}}})_i, \qquad W_i = 1/2(n+\kappa)\\ \boldsymbol{\chi}_{i+n} &= \boldsymbol{\mu_x} - (\sqrt{(n+\kappa)\mathbf{P}_{\boldsymbol{xx}}})_i, \qquad W_{i+n} = 1/2(n+\kappa)\end{aligned}$$

where $\kappa \in \Re$ and $(\sqrt{(n+\kappa)\mathbf{P}_{\boldsymbol{xx}}})_i$ represents the ith row or column of the matrix square root of $(n+\kappa)\mathbf{P}_{\boldsymbol{xx}}$ and W_i is the weight which is associated with the ith point. The transformation procedure is as follows:

1. Obtain the set of transformed sigma-points,

$$\mathcal{Y}_i = \psi(\boldsymbol{\chi}_i), \qquad i = 0, 1, \cdots, 2n$$

2. Compute the mean given by the weighted average of the transformed points

$$\boldsymbol{\mu_y} = \sum_{i=0}^{2n} W_i \mathcal{Y}_\mathbf{i}$$

3. Compute the covariance by the weighted outer product of the transform points,

$$\mathbf{P}_{\boldsymbol{yy}} = \sum_{i=0}^{2n} W_i(\mathcal{Y}_\mathbf{i} - \boldsymbol{\mu_y})(\mathcal{Y}_\mathbf{i} - \boldsymbol{\mu_y})^\dagger$$

The detailed properties of the algorithms can be found in [14]. Note that κ in the algorithm provides an extra degree of freedom to adjust the higher order moments of the approximation and can be used to to reduce the overall prediction errors. It was shown in [14] that when $\boldsymbol{x}$ is Gaussian, a useful choice of the κ is $\kappa = n - 3$.

4. KALMAN FILTERING BASED ON UNSCENTED TRANSFORMATION

The phase noise process (1) is a Gaussian process. Hence,

$$p(\theta_t|\boldsymbol{S}_t, \boldsymbol{Y}_t) \sim N(\mu_{\theta_t}(\boldsymbol{S}_t), \sigma^2_{\theta_t}(\boldsymbol{S}_t)), \tag{5}$$

where the mean $\mu_{\theta_t}(\boldsymbol{S}_t)$ and the variance $\sigma^2_{\theta_t}(\boldsymbol{S}_t)$ can be obtained as follows. Denoting $\widehat{\theta}_{t|t-1}$ as the estimator of θ_t based on the observations $\boldsymbol{Y}_{t-1} = (y_0, y_1 \cdots, y_{t-1})$ and

$$\mu_{\theta_t}(\boldsymbol{S}_t) \stackrel{\triangle}{=} \widehat{\theta}_{t|t} \quad \text{and} \quad \sigma^2_{\theta_t}(\boldsymbol{S}_t) \stackrel{\triangle}{=} M_{t|t}, \tag{6}$$

$\widehat{\theta}_{t|t}$ and $M_{t|t}$ can be calculated recursively by using the Kalman Technique [10, page 449-452] along with the unscented transformation, given $\boldsymbol{S}_t$ as:

$$\begin{aligned}\widehat{\theta}_{t|t} &= \widehat{\theta}_{t|t-1} + K_t\nu &(7)\\ M_{t|t} &= M_{t|t-1} - K_t M^{\nu\nu}_{t|t-1} K^*_t &(8)\\ \nu &= y_t - \widehat{y}_{t|t-1} &(9)\\ K_t &= M^{\theta\nu}_{t|t-1}(M^{\nu\nu}_{t|t-1})^{-1}, &(10)\end{aligned}$$

where

$$\begin{aligned}M^{\nu\nu}_{t|t-1} &= E\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})^*|\boldsymbol{Y}_{t-1}\right]\\ M^{\theta\nu}_{t|t-1} &= E\left[(\theta_t - \widehat{\theta}_{t|t-1})(y_t - \widehat{y}_{t|t-1})^*|\boldsymbol{Y}_{t-1}.\right]\end{aligned}$$

In order to implement the Kalman filter algorithm given above, one needs to compute

- Prediction of the new state of the phase noise $\widehat{\theta}_{t|t-1}$ and its variance $M_{t|t-1}$
- Prediction of the expected observation $\widehat{y}_{t|t-1}$ and the innovation variance $M^{\nu\nu}_{t|t-1}$
- Prediction of the cross correlation $M^{\theta\nu}_{t|t-1}$.

Note that, since the state equation (2) is linear we can easily obtain $\widehat{\theta}_{t|t-1}$ and $M_{t|t-1}$ as follows:

$$\begin{aligned}\widehat{\theta}_{t|t-1} &= \widehat{\theta}_{t-1|t-1} &(11)\\ M_{t|t-1} &= M_{t|t-1} + \sigma^2_u. &(12)\end{aligned}$$

Furthermore, it can be easily shown from (1) and (2) that

$$\begin{aligned}\widehat{y}_{t|t-1} &= E\{\exp(j\theta_t)|\boldsymbol{Y}_{t-1}\} &(13)\\ M^{\nu\nu}_{t|t-1} &= 1 + \sigma^2_w - |\widehat{y}_{t|t-1}|^2 &(14)\\ M^{\theta\nu}_{t|t-1} &= E\{\theta_t \exp(-j\theta_t)|\boldsymbol{Y}_{t-1}\} - \widehat{y}^*_{t|t-1}\widehat{\theta}_{t|t-1}. &(15)\end{aligned}$$

The expectations above can be computed by the unscented filtering technique as follows: Since the θ_t's are one-dimensional and Gaussian, three sigma-points would be sufficient to implement the algorithm. The three sigma points and the corresponding weights are chosen according to the general formulation explained in Section 3 as

$$\begin{aligned}\Theta^{(0)}_{t|t-1} &= \widehat{\theta}_{t|t-1}, &W_0 = \kappa/(1+\kappa)\\ \Theta^{(1)}_{t|t-1} &= \widehat{\theta}_{t|t-1} + \sqrt{(1+\kappa)M_{t|t-1}}, &W_1 = 1/2(1+\kappa)\\ \Theta^{(2)}_{t|t-1} &= \widehat{\theta}_{t|t-1} - \sqrt{(1+\kappa)M_{t|t-1}}, &W_2 = 1/2(1+\kappa).\end{aligned}$$

Note that since θ_t is Gaussian, and $n = 3$, the value of $\kappa = 0$ as pointed out in Section 3. Therefore $W_0 = 0$ and the only following two sigma-points are sufficient to implement the algorithm.

$$\begin{aligned}\Theta^{(1)}_{t|t-1} &= \widehat{\theta}_{t|t-1} + \sqrt{M_{t|t-1}}, &W_1 = 1/2 \quad (16)\\ \Theta^{(2)}_{t|t-1} &= \widehat{\theta}_{t|t-1} - \sqrt{M_{t|t-1}}, &W_2 = 1/2 \quad (17)\end{aligned}$$

We now summarize the algorithm to compute the expectations (13) and (15).

1. Select the two sigma points $\Theta^{(0)}_{t|t-1}, \Theta^{(1)}_{t|t-1}$ according to (16) and (17).
2. for $i = 1, 2$, compute the transform sets

$$\begin{aligned}\widehat{\mathcal{Y}}^{(i)} &= \exp(j\Theta^{(i)}_{t|t-1})\\ \widehat{\mathcal{Z}}^{(i)} &= \Theta^{(i)}_{t|t-1}\exp(-j\Theta^{(i)}_{t|t-1})\end{aligned}$$

3. Compute

$$\begin{aligned}\widehat{y}_{t|t-1} &= \sum_{i=1}^{2} W_i\widehat{\mathcal{Y}}^{(i)}\\ M^{\theta\nu}_{t|t-1} &= \sum_{i=1}^{2} W_i\widehat{\mathcal{Z}}^{(i)} - \widehat{y}^*_{t|t-1}\widehat{\theta}_{t|t-1}\\ M^{\nu\nu}_{t|t-1} &= 1 + \sigma^2_w - |\widehat{y}_{t|t-1}|^2.\end{aligned}$$

5. SMC TECHNIQUE FOR BLIND DETECTION AND ESTIMATION

We can now make timely estimates of θ_t and detection of s_t based on the currently available observation $\boldsymbol{Y}_t$, up to time t, blindly, as follows. With the Bayes theorem, we realize that the optimal solution to this problem is

$$\begin{aligned}\widehat{\theta}_t &= E\{\theta_t|\boldsymbol{Y}_t\} = \int \theta_t p(\theta_t|\boldsymbol{Y}_t)d\theta_t \qquad (18)\\ &= \int_{\boldsymbol{S}_t} \underbrace{\left[\int_{\theta_t} \theta_t p(\theta_t|\boldsymbol{S}_t,\boldsymbol{Y}_t)d\theta_t\right]}_{\mu_{\theta_t}(\boldsymbol{S}_t)} p(\boldsymbol{S}_t|\boldsymbol{Y}_t), d\boldsymbol{S}_t.\end{aligned}$$

It then follows that

$$\widehat{\theta}_t = E\{\theta_t|\boldsymbol{Y}_t\} = \int_{\boldsymbol{S}_t} \mu_{\theta_t}(\boldsymbol{S}_t)p(\boldsymbol{S}_t|\boldsymbol{Y}_t)d\boldsymbol{S}_t \ . \qquad (19)$$

Similarly, the data can be detected by the hard decisions on the symbol s_t by

$$\widehat{s_t} = \arg\max_{a_i\in A} P(s_t = a_i|\boldsymbol{Y}_t) \qquad (20)$$

where

$$P(s_t = a_i|\boldsymbol{Y}_t) = E\{1(s_t = a_i)|\boldsymbol{Y}_t\} \ . \qquad (21)$$

$1\{.\}$ in (21) is an indicator function defined as

$$1(s_t = a_i)\begin{cases} 1 & \text{if } s_t = a_i \\ 0 & \text{otherwise.}\end{cases}$$

In most cases, an exact evaluations of the expectations (19) and (21) are analytically intractable. SMC technique can provide us an alternative way for the required computation. Specifically, following the notation adopted in [11], if we can draw m independent random samples $\{\boldsymbol{S}_t^{(j)}\}_{j=1}^m$ from the distribution $p(\boldsymbol{S}_t|\boldsymbol{Y}_t)$, then we can approximate the quantities of interest $E\{\theta|\boldsymbol{Y}_t\}$ and $E\{1(s_t = a_i)|\boldsymbol{Y}_t\}$ in (11) and (13), respectively, by

$$E\{\theta|\boldsymbol{Y}_t\} \cong \frac{1}{m}\sum_{j=1}^m \mu_{\theta_t}(\boldsymbol{S}_t^{(j)}) \qquad (22)$$

$$E\{1(s_t = a_i)|\boldsymbol{Y}_t\} \cong \frac{1}{m}\sum_{j=1}^m 1(s_t^{(j)} = a_i) \qquad (23)$$

But, usually drawing samples from $p(\boldsymbol{S}_t|\boldsymbol{Y}_t)$ directly is usually difficult. Instead, sample generation from some *trial distribution* may be easier. In this case, the idea of *importance sampling* can be used. Suppose a set of random samples $\{\boldsymbol{S}_t^{(j)}\}_{j=1}^m$ is generated from a trial distribution $q(\boldsymbol{S}_t|\boldsymbol{Y}_t)$, which

- is strictly positive, $q(.|.) > 0$, and
- has the domain as $p(.|.)$.

By associating the weight

$$w_t^{(j)} = \frac{p(\boldsymbol{S}_t^{(j)}|\boldsymbol{Y}_t)}{q(\boldsymbol{S}_t^{(j)}|\boldsymbol{Y}_t)} \qquad (24)$$

to the samples, the quantities of interest, $E\{1(s_t = a_i)|\boldsymbol{Y}_t\}$ and $E\{\theta_t|\boldsymbol{S}_t\}$ can be approximated as follows.

$$\begin{aligned}E\{\theta|y_t\} &\cong \frac{1}{W_t}\sum_{j=1}^m \mu_t(\boldsymbol{S}_t^{(j)})w_t^{(j)} \qquad (25)\\ E\{1(s_t = a_i)|\boldsymbol{Y}_t\} &\cong \frac{1}{W_t}\sum_{j=1}^m 1(s_t^{(j)} = a_i)w_t^{(j)}, \ \ i = 1,2,..,|A|\end{aligned}$$

with $W_t = \sum w_t^{(j)}$. The pair $(\boldsymbol{S}_t^{(j)}, w_t^{(j)}), j = 1,2,\cdots,m$ is called a properly weighted sample with respect to distribution $p(\boldsymbol{S}_t|\boldsymbol{Y}_t)$. Note that the samples $\boldsymbol{S}_t^{(j)}$ can be drawn from the distribution $q(\boldsymbol{S}_t|\boldsymbol{Y}_t)$ sequentially as follows. We can choose $q(.)$ to satisfy

$$q(\boldsymbol{S}_{t-1}|\boldsymbol{Y}_t) = q(\boldsymbol{S}_{t-1}|\boldsymbol{Y}_{t-1}).$$

Then, it can be easily shown that

$$q(\boldsymbol{S}_t|\boldsymbol{Y}_t) = q(s_t|\boldsymbol{Y}_t,\boldsymbol{S}_{t-1})q(\boldsymbol{S}_{t-1}|\boldsymbol{Y}_{t-1}),$$

and one can draw samples $s_t^{(j)}$ from a trial distribution $q(s_t|\boldsymbol{Y}_t,\boldsymbol{S}_{t-1}^{(j)})$ and let $\boldsymbol{S}_t^{(j)} = (s_t^{(j)},\boldsymbol{S}_{t-1}^{(j)})$ for $t = 0,1,\cdots$. Specifically, it was shown in [11] that a suitable choice for the trial distribution is of the form:

$$q(s_t|\boldsymbol{Y}_t,\boldsymbol{S}_{t-1}^{(j)}) = p(s_t|\boldsymbol{Y}_t,\boldsymbol{S}_{t-1}^{(j)}) \ . \qquad (26)$$

For this trial distribution, it is shown in [11] that the importance weight is updated according to

$$w_t^{(j)} = w_{t-1}^{(j)}p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)}), \quad t = 0,1,\cdots \qquad (27)$$

The predictive distribution in (27) is given by

$$\begin{aligned}p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)}) &= \sum_{a_i\in A} p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)},s_t = a_i)P(s_t = a_i|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)})\\ &= \sum_{a_i\in A} p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)},s_t = a_i)P(s_t = a_i) \qquad (28)\end{aligned}$$

where (28) holds because s_t is independent of $\boldsymbol{S}_{t-1}$ and $\boldsymbol{Y}_{t-1}$. Furthermore, it can be shown from the state and observation equations in (1) and (2),respectively, that

$$p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)},s_t = a_i) \sim N(\mu_{y_t}^{(j)}(i),\sigma_{y_t}^{2(j)}(i)) \qquad (29)$$

with mean and variance given by

$$\begin{aligned}\mu_{y_t}^{(j)}(i) &= E\{y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)},s_t = a_i\}\\ &= a_i(H_t\mu_{\theta_{t-1}}^{(j)} + Q_t) \qquad (30)\\ \sigma_{y_t}^{2(j)}(i) &= \mathrm{Var}\{y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)},s_t = a_i\}\\ &= \sigma_{\theta_{t-1}}^{2(j)} + \sigma_n^2 + \sigma_p^2 \qquad (31)\end{aligned}$$

where the quantities $\mu_{\theta_t^{(j)}}$ and $\sigma_{\theta_t^{2(j)}}$ in (30)and (31) can be computed recursively for the Extended Kalman equations given in (7-8). The trial distribution in (30) can be computed as follows:

$$\begin{aligned}p(s_t = a_i|\boldsymbol{Y}_t,\boldsymbol{S}_{t-1}^{(j)}) &= p(y_t|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)},s_t = a_i)\\ &\quad \times P(s_t = a_i|\boldsymbol{Y}_{t-1},\boldsymbol{S}_{t-1}^{(j)})\\ &\triangleq \xi_{t,i}^{(j)} \qquad (32)\end{aligned}$$

where it follows from (2) that

$$\xi_{t,i}^{(j)} = \frac{1}{\pi\sigma_{y_t}^{2(j)}(i)} \exp\left(-\frac{||y_t - \mu_{y_t}^{(j)}(i)||^2}{\sigma_{y_t}^{2(j)}(i)}\right) P(s_t = a_i). \quad (33)$$

We now summarize the SMC blind data detection and phase noise estimation algorithm as follows:

Step 1- *Initialization:*

- Initialize the Kalman filter: Choose the initial mean and the variance of the estimated θ_t as the mean and the variance of a uniform distribution defined on $-\pi, +\pi$) as

$$\begin{aligned} \mu_{\theta_0}^{(j)} &= \widehat{\theta}_{0|0}^{(j)} = 0 \\ \sigma_{\theta_0}^{2(j)} &= M_{0|0}^{(j)} = \pi^2/12, \quad j = 1, 2, \cdots, m. \end{aligned} \quad (34)$$

- Initialize the importance weights: All importance weights are initialized as $w_0^{(j)} = 1, j = 1, 2, \cdots, m$. Since the data symbols are are assumed to be independent, initial symbols are not needed be generated.

Step 2- *Compute* $\xi_{t,i}^{(j)}$*:*
For each $a_i \in A$ compute the $\mu_{y_t}^{(j)}(i), \sigma_{y_t}^{2(j)}(i)$ and $\xi_{t,i}^{(j)}$ from (30), (31) and (33),respectively.

Step 3- *Draw samples* $s_t^j, j = 1, 2, \cdots, m$
Draw $s_t^{(j)}$ from the set A with probabilities

$$P(s_t^{(j)} = a_i) \propto \xi_{t,i}^{(j)}, \quad a_i \in A. \quad (35)$$

Append $s_t^{(j)}$ to $\boldsymbol{S}_{t-1}^{(j)}$ to obtain $\boldsymbol{S}_t^{(j)}$.

Step 4- *Compute the importance weights:*

$$w_t^{(j)} = w_{t-1}^{(j)} \sum_{a_i \in A} \xi_{t,i}^{(j)}.$$

Step 5-*Detect the symbol* s_t*:*
Detect the symbol s_t from (23).

Step 6-*Update the a posteriori mean and variance of the phase noise:*
If the samples drawn up to time t is $\boldsymbol{S}_t$ in Step 2, set

$$\begin{aligned} \mu_{\theta_t}(\boldsymbol{S}_t^{(j)}) &\triangleq \mu_{\theta_t}^{(j)} = \widehat{\theta}_{t|t}^{(j)} \\ \sigma_{\theta_t}^{2(j)}(\boldsymbol{S}_t^{(j)}) &\triangleq \sigma_{\theta_t}^{2(j)} = M_{t|t}^{(j)} \quad j = 1, 2, \cdots, m. \end{aligned}$$

and update according to the Kalman equations (7), (8).

Step 7- *Do the restamping as described in Section 6.*

6. RESAMPLING METHOD

A major problem in the practical implementation of the SMC method described so far is that after a few iteration most of the importance weights have negligible values that is $w_t^{(j)} \approx 0$. A relatively small weight implies that the sample is drawn far from the main body of the posterior distribution and has a small contribution in the final estimation. Such sample is said to be ineffective. The SMC algorithm becomes ineffective if there are too many ineffective samples. The common solution to this problem is *resampling.* Restampling is a an algorithmic step that stochastically eliminates those samples with small weights. Basically, the resampling method takes the samples, to be generated sequentially $\Xi_t = \{\boldsymbol{S}_t^{(j)}, \mu_{\theta_t}^{(j)}, \sigma_{\theta_t}^{2(j)}\}_{j=1}^m$ with corresponding weights $\{w_t^{(j)}\}_{j=1}^m$ as an input and generates a new set of samples $\widetilde{\Xi}_t = \{\widetilde{\boldsymbol{S}}_t^{(j)}, \widetilde{\mu}_{\theta_t}^{(j)}, \widetilde{\sigma}_{\theta_t}^{2(j)}\}_{j=1}^m$ with equal weights, i.e $\{w_t^{(j)} = 1/m\}_{j=1}^m$, assuming they are normalized to $\sum_{j=1}^m w_t^{(j)} = 1$. A simple procedure to achieve this goal is, for each $j = 1, 2, \cdots, m$, to choose $(\widetilde{\boldsymbol{S}}_t^{(j)}, \widetilde{\mu}_{\theta_t}^{(j)}, \widetilde{\sigma}_{\theta_t}^{2(j)}) = (\boldsymbol{S}_t^{(j)}, \mu_{\theta_t}^{(i)}, \sigma_{\theta_t}^{2(i)})$ with probability $w_t^{(i)}$.

In this paper, a resampling technique suggested by [12] is employed. This technique forms a new set of weighted samples $\widetilde{\Xi}_t = \{\widetilde{\boldsymbol{S}}_t^{(j)}, \widetilde{\mu}_{\theta_t}^{(j)}, \widetilde{\sigma}_{\theta_t}^{2(j)}\}_{j=1}^m$ according to the following algorithm. (assume that $\sum_{j=1}^m w_t^j = m$)

1. For $j = 1, 2, \cdots, m$, retain $\ell_j = w_t^j$ copies of the samples $(\boldsymbol{S}_t^{(j)}, \mu_{\theta_t}^{(i)}, \sigma_{\theta_t}^{2(i)})$. Denote $L_r = m - \sum_{j=1}^m \ell_j$.
2. Obtain L_r i.i.d. draws from the original sample set $\{(\boldsymbol{S}_t^{(j)}, \mu_{\theta_t}^{(i)}, \sigma_{\theta_t}^{2(i)})\}_{j=1}^m$, with probabilities proportional to $(w_t^j - \ell_j), j = 1, 2, \cdots, m$.
3. Assign equal weights, that is, set $w_t^j = 1$, for each new sample.

It is shown in [12] that the samples drawn by the above procedure are properly weighted with respect to $p(\boldsymbol{S}_t|\boldsymbol{Y}_t)$, provided that m is sufficiently large. Note that resampling at every time step is not needed in general. In one way the resampling can be done every k_0 recursions where k_0 is a prefixed resampling interval. On the other hand, the resampling can be carried out whenever the effective sample size, approximated as

$$\widehat{N}_{eff} = \frac{1}{\sum_{j=1}^m (w_t^j)^2} \leq m \quad (36)$$

goes below a certain threshold, typically a fraction of m. Intuitively, $\widehat{N}_{eff}$ reflects the equivalent size of i.i.d samples from the true posterior densities of interest for the set of m weighted ones. It is suggested in [13] that resampling should be performed when $\widehat{N}_{eff} < m/10$. Alternatively, one can conduct the first approach to conduct resampling at every fixed-length time interval say every five steps.

7. SIMULATION RESULTS

In this section, we provide some computer simulation examples to demonstrate the performance of the proposed SMC and UF approach for joint blind phase noise estimation and data detection. The phase process is modelled by AR process driven by a white Gaussian noise with $\sigma_u^2 = 0.0314$. It is assumed BPSK modulation is employed. In order to demonstrate the performance of the adaptive SMC approach, we first present the performance (in terms of the phase error $\phi(k) = \theta_t - \widehat{\theta}_t$) during one simulation run for different initial phase errors $\phi(k) = 0, \pi/4, \pi/2, 3\pi/4, \pi$. The phase error for several values of $\phi(0)$ at $SNR = 10dB$ is shown in Fig. 1.

The performance of the proposed algorithm is further exploited by the evaluation of average BER over observed block for different SNRs and different intial phase errors. The un-

coded average BER performance of this adaptive approach is plotted in Fig. 2.

Our simulations indicate that

- as the initial phase error $\phi(0)$ approaches π, the probability that the phase error converges to the dual equilibrium point becomes very high
- as the initial phase error $\phi(0)$ approaches π, the BER increases, for $\phi(0) = \pi/2$, the BER is almost equal to 1 (due to ambiguity).

8. CONCLUSIONS

We have developed a new adaptive Bayesian approach for blind phase noise estimation and data detection based on sequential Monte Carlo methodology. The optimal solutions to joint symbol detection and phase noise tracking problem is computationally prohibitive to implement by conventional methods. Thus the proposed sequential approach offers an novel and powerful approach to tackling this problem at a reasonable computational cost.

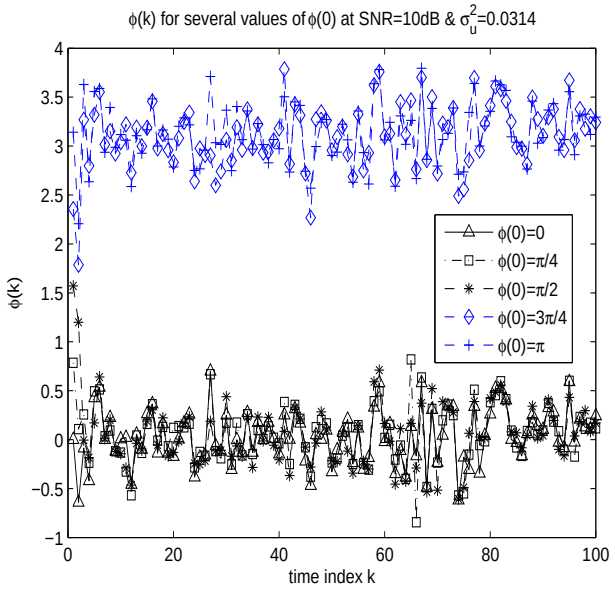

Figure 1: Tracking performance for different initializations at SNR=10dB

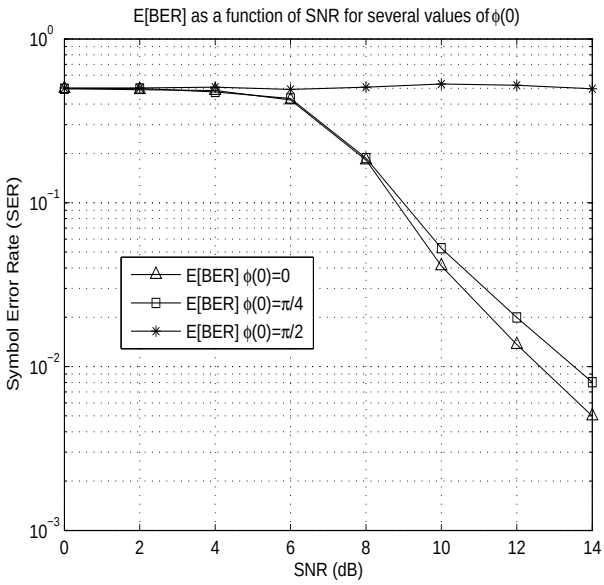

Figure 2: BER performance

REFERENCES

[1] H. Meyr, M. Moeneclaey and S.A. Fechtel, *Digital Communications Receivers*, John Wiley, New York, 1998.

[2] C. N. Georghiades and J.C. Han, "Optimum decoding of trellis coded modulation in the presence of phase noise", *Proc. 1990 Int. Symp. on Information Theory and Its Applications (ISITA'90)* Hawaii, Nov. 1990.

[3] A. T. Huq, E. Panayirci and C. N. Georghiades, " ML NDA carrier phase recovery for OFDM systems with M-PSK signaling" , *Proc. IEEE Int. Conf. on Communications (ICC'99)*, June 6-10, 1999,Vancouver, Canada.

[4] E. Panayırcı and C. N. Georghiades," Non-data-aided ML carrier frequency and phase synchronization in OFDM systems", *European Transactions on Telecommunications (ETT)*, 2000.

[5] U. Mengali, A.L. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum Press, New York 1997.

[6] A. P. Dempster, N. M. Laird and D. B. Rubin,"Maximum likelihood from incomplete data via the EM algorithm",*Ann. Roy. Stat. Soc.*, pp. 1-38, December 1977.

[7] M. Pitt and N. Shephard," Filtering via simulation: auxiliry particle filter",*J. Amer. Statist. Soc. B* , no.63, 2001.

[8] E. Panayirci, H. A. Cirpan and M. Moeneclaey,A Sequential Monte Carlo Method for Blind Phase Noise Estimation and Data Detection, 13th European Signal Processsing Conference, EUSIPCO 2005, Sep. 4-8 2005, Antalya, Turkiye.

[9] S. Geman and D. Geman," Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images",*IEEE Trans. on Pattern Analysis nad Machine Intelligence* vol. 6 ,pp. 721-741, 1984.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Upper Saddle River, NJ: Prentice Hall, 1993.

[11] Z. Yang and X. Wang, " A sequential Monte Carlo blind receiver for OFDM systems in frequency-selective fading channels",*IEEE Trans. Signal Proc.*,vol. 50, no. 2, pp.271-280, Feb. 2003.

[12] R. Chen, X. Wang and J. S. Liu, " Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filering", *IEEE Trans.Information Theo.*,vol. 46, no. 6, pp.2070-2094, Sept. 2000.

[13] J. S. Liu and R. Chen,"Blind deconvolution via sequential imputations", *J. Amer. Statist. Assoc.*, vol. 90, pp. 567-576, 1995.

[14] S. J. Julier and J.K. Uhlmann,"Unscented filtering and nonlinear estimation", *Proc. of the IEEE*, vol. 92, no.3, pp,401-422, 2004.