Symbol Detection with Time-varying Unknown Phase by Expectation Propagation

Tao Wei, Yufei Huang
Department of Electrical Engineering
The University of Texas at San Antonio
San Antonio, TX 78249 USA
twei@lonestar.utsa.edu, yhuang@utsa.edu

Yuan Qi
MIT Media Lab
Cambridge, MA, 02139 USA
yuanqi@media.mit.edu

***Abstract*— In digital communications, symbol detection in phase noise is an important topic that has been discussed in many papers under different conditions. In this paper, we consider symbol detection with time-varying unknown phase. We propose a solution based on expectation propagation (EP). EP is an extension to belief propagation and developed in machine learning. We point out that the developed EP solution can be considered as iterated extended Kalman smoother (EKS). However, a crucial step of recycling the likelihoods in EP makes possible the further improvement over EKS. We show in the simulation that EP can produce very good performance with relatively low complexity. Since it produce soft information, the EP solution can be readily applied to iterative detection of coded systems.**

I. Introduction

Reliable signal detection is a fundamental task in digital communications. When performing detection at the receiver, phase shift due to channel delays must be compensated. The problem becomes, however, very challenging when phase changes dynamically in time, a very realistic scenario in communications. The difficulties of the problem rise from the nonlinear, time-varying nature of the phase model and an exponentially increasing detection space. To address the difficulties, solutions based on the extended Kalman filter (EKF)[1], the unscented Kalman filter (UKF)[2], and other suboptimal adaptive estimation techniques have been reported in the literature. However, a decision-directed scheme is often adopted for symbol detection, whose performance is limited by error propagation.

Recently, more advanced techniques were proposed for the solution of the problem. Among them, particle filtering [3] and belief propagation (BP) on factor graph [4], [5] are very appealing. However, despite the near-optimum performance, the complexity of particle filtering can be very high. Although its complexity can be affordable, BP is, however, developed for discrete random variables or the continuous variables whose distributions are in the exponential family, such as Gaussian distribution.

In this paper, we develop a novel solution based on expectation propagation (EP) [6]. EP is an extension to belief propagation (BP) and is specially developed for distributions outside of the exponential family. EP has relatively less complexity and may produce comparable performance to particle filtering. Moreover, EP is potentially more powerful and applicable to wider problem than BP. We provide in this paper detailed derivations of the EP solution and point out that the EP solution can be considered as an extension to extended Kalman smoother (EKS) [7]. However, a crucial step of recycling the likelihood makes possible the improvement over the EKS. We also show in the simulation that EP can produce very good performance with relatively low complexity. Since it produce soft information, the EP solution can be applied to iterative detection of coded systems.

The remaining of the paper is organized as follows. In section II, system model is formulated and the concerned problem is presented. The background of EP is described in section III. In section IV, solution based on EP is derived in detail. Simulations are included in V and conclusion is drawn in section VI.

II. Problem Formulation

Consider detection of K symbols from a block of K observations. We assume the symbols, when transmitted through communication channels, undergo time-varying phase shift. At time k, the observation y_k can be modeled by a state space representation as

$$\theta_k = \theta_{k-1} + w_k \tag{1}$$

$$y_k = s_k e^{j\theta_k} + n_k \tag{2}$$

where θ_k is the time-varying phase, whose dynamics is modeled by a random walk process (1), $w_k \sim \mathcal{N}(0, \sigma_w^2)$ is the driving noise of the phase process, s_k is the data symbol generated from an alphabet set $\mathcal{A}$ of size M, and $n_k \sim \mathcal{N}(0, \sigma_n^2)$ is complex white Gaussian noise. Our objective is to detect s_k from a block of observations with, however, time-varying unknown phase. In what follows, we use subscript $_{1:K}$ to denote a collection of variables from $k = 1$ to K and subscript $_{1:K,-k}$ to denote a collection of variables from $k = 1$ to K except the kth.

From a Bayesian perspective, the *a posteriori* probabilities (APP) $p(s_k|y_{1:K})$ $\forall k$ must be calculated for detection. Due to the unknown phase, the APPs are obtained by marginalizing the joint posterior distribution as follows,

$$p(s_k|y_{1:K}) = \sum_{\mathbf{s}_{1:K,-k}} \int_{\theta_{1:K}} p(s_{1:K}, \theta_{1:K}|y_{1:K}) d\theta_{1:K} \tag{3}$$

There are two major difficulties in (3). First, since y_k is nonlinear in θ_k, the high dimensional integration cannot be solved analytically. Second, the size of summation increases exponentially with k, which makes the problem NP hard. Suboptimum approaches must be adopted as a result. We show in the following how EP can be used for the solution.

III. Background on Expectation Propagation

Consider K independent observations generated from a statistical parametric model $y_k = f(\theta, n_k)$, where $f(\cdot)$ is a parametric function, θ the unknown parameter, and n_k random noise. Given the prior distribution $p(\theta)$, our objection is to obtain the posterior distribution $p(\theta|y_{1:K})$. Except for limited cases where, for instance, $f(\cdot)$ is linear and n_k is Gaussian, the posterior distribution cannot, however, be derived analytically. Instead, EP can be applied to approximate the desired posterior distribution.

EP consists of two major parts: initial density estimation and iterative refinement. In the first part, an initial estimate on $p(\theta|y_{1:k})$ is constructed by sequentially incorporating the observations from $k = 1$ to K using ADF. Note that the order is imposed for convenience of our composition and would be natural for dynamic systems, but it is not necessary in EP nevertheless. To see the procedure in detail, we assume that, at step $k-1$, we have obtained an approximation of $p(\theta|y_{1:k-1})$, say, $q(\theta|y_{1:k-1})$. We, however, restrict $q(\theta|y_{1:k-1})$ to be from the exponential family, a key requirement of EP when approximating posterior distribution. This is because that, with a distribution from the exponential family, only a fixed number of expectations (the sufficient statistics) needs to be propagated. Now, to incorporate the new likelihood $p(y_k|\theta)$ at step k and to obtain the new approximation $q(\theta|y_{1:k})$ from $p(\theta|y_{1:k})$, we start from the following relationship

$$\begin{aligned} p(\theta|y_{1:k}) &= \frac{p(y_k|\theta)p(\theta|y_{1:k-1})}{Z} \\ &\approx \frac{p(y_k|\theta)q(\theta|y_{1:k-1})}{Z} \\ &= \hat{q}(\theta|y_{1:k}) \end{aligned} \tag{4}$$

where Z is the normalizing constant. Since $\hat{q}(\theta|y_{1:k})$ may not be in the exponential family, we need to project $\hat{q}(\theta|y_{1:k})$ to the exponential family distribution to obtain the required approximation $q(\theta|y_{1:k})$. Based on the criterion that the Kullback-Leibler (KL) distance between the original and the projected is minimized, it is shown in [6] that the projection is equivalent to moment matching. For example, if $q(\theta|y_{1:k})$ is chosen to be Gaussian, then the moment matching matches the mean and variance of $q(\theta|y_{1:k})$ to those of $\hat{q}(\theta|y_{1:k})$. When $f(\cdot)$ is not linear, the moments of $\hat{q}(\theta|y_{1:k})$ cannot be obtained analytically. Techniques including quadratic approximation and unscented transformation can be used instead to approximate these moments. Note, with $q(\theta|y_{1:k})$ and $q(\theta|y_{1:k-1})$, the corresponding approximated likelihood function can be also obtained as

$$q(y_k|\theta) \propto \frac{q(\theta|y_{1:k})}{q(\theta|y_{1:k-1})} \tag{5}$$

When the above steps are finished at $k = K$, we obtain an initial estimate on $p(\theta|y_{1:K})$.

In the second part of EP, refinement on the approximation $q(\theta|y_{1:K})$ is performed in an iterative fashion. In each iteration, refinement is also performed sequentially from 1 to K by recycling the K likelihoods. In specific, at the kth step, two sub-steps are included

1) *Removal of the approximated likelihood*: The approximated likelihood is removed according to

$$q(\theta|y_{1:K;-k}) \propto \frac{q(\theta|y_{1:K})}{q(y_k|\theta)} \tag{6}$$

where $y_{1:K;-k}$ represents the collection of the observations except y_k.

2) *True likelihood recycling and moment matching*: The true likelihood $p(y_k|\theta)$ is then combined with $q(\theta|y_{1:K;-k})$ by the same fashion as in (4) and the refined approximation $q(\theta|y_{1:K})$ is obtained through moment matching.

After one sweep from $k = 1$ to K, this refinement iterates again until the convergence of $q(\theta|y_{1:K})$ and EP outputs the converged $q(\theta|y_{1:K})$ as the final approximation to $p(\theta|y_{1:K})$.

IV. Symbol Detection in time-varying Phase Noise with Expectation Propagation

For our problem, the objective is to calculate the marginal *a posteriori* probabilities (APPs) $p(s_k|y_{1:K})\ \forall k$. EP approximates the desired APPs indirectly by approximating the joint density $p(s_k, \theta_k|y_{1:K})\ \forall k$ first.

As discussed in section III, EP includes two parts. In the first part, initial estimates of $p(s_k, \theta_k|y_{1:K})\ \forall k$ are formulated by sequentially incorporating the likelihoods. For the concerned dynamic system, since $p(s_k, \theta_k|y_{1:K})$ is a smoothing density, an estimate is conveniently calculated through a forward (filtering) and a backward (smoothing) process. However, the moment an estimate on a smoothing density of, say, s_k is derived, the second part of EP should be performed right away to refine the estimate. This is because in dynamic systems we deal with the different unknowns at different k. As a result, the refinement part is intertwined in the smoothing process.

Let us first discuss the filtering process, suppose at $k-1$ that we have obtained $q(s_{k-1}|y_{1:k-1})$ and $q(\theta_{k-1}|y_{1:k-1})$ as the approximation of $p(s_{k-1}|y_{1:k-1})$ and $p(\theta_{k-1}|y_{1:k-1})$, respectively. In particular, $q(s_{k-1}|y_{1:k-1})$ is a discrete distribution defined on the symbol space $\mathcal{A}$ and $q(\theta_{k-1}|y_{1:k-1})$ is a Gaussian distribution, i.e.,

$$q(s_{k-1}|y_{1:k-1}) = Discrete(r_{k-1,1}, \cdots, r_{k-1,M})$$
$$q(\theta_{k-1}|y_{1:k-1}) = \mathcal{N}(\bar{\theta}_{k-1|k-1}, \sigma^2_{\theta_{k-1|k-1}}).$$

Now we want to obtain the estimates of $p(s_k|y_{1:k})$ and $p(\theta_k|y_{1:k})$ from approximating $p(s_k, \theta_k|y_{1:k})$. For a uniform prior of $p(s_k) = 1/M$, $p(s_k, \theta_k|y_{1:k})$ can be expressed as

$$p(s_k, \theta_k|y_{1:k}) \propto p(y_k|s_k, \theta_k)p(\theta_k|y_{1:k-1}). \tag{7}$$

To compute (7), we first calculate the predict density $p(\theta_k|y_{1:k-1})$ as

$$\begin{aligned} p(\theta_k|y_{1:k-1}) &= \int p(\theta_k|\theta_{k-1})p(\theta_{k-1}|y_{1:k-1})d\theta_{k-1} \\ &\approx \int p(\theta_k|\theta_{k-1})q(\theta_{k-1}|y_{1:k-1})d\theta_{k-1} \\ &= \mathcal{N}(\bar{\theta}_{k-1|k-1}, {\sigma^2}_{\theta_{k|k-1}}) \end{aligned} \tag{8}$$

where

$${\sigma^2}_{\theta_{k|k-1}} = \sigma^2_{\theta_{k-1|k-1}} + \sigma^2_{w_k}. \tag{9}$$

Next, we calculate the likelihood $p(y_k|s_k, \theta_k)$. We observe, however, that unless y_k is linear in θ, the moments of $p(s_k, \theta_k|y_{1:k})$ cannot be calculated analytically. Thus, linearization on the model over θ must be performed. In this paper, we adopt the quadratic approximation and the likelihood after linearization can be expressed as a Gaussian distribution

$$p(y_k|s_k, \theta_k) \approx \mathcal{N}(y_k^{new}|\bar{y}_k, \sigma^2_{y_k}) \tag{10}$$

where

$$\begin{aligned} y_k^{new} &= y_k - s_k g(\bar{\theta}_{k-1|k-1}) && (11) \\ \bar{y}_k &= s_k G_k \bar{\theta}_{k-1|k-1} && (12) \\ \sigma^2_{y_k} &= |s_k|^2 |G_k|^2 (\sigma^2_{\theta_{k-1|k-1}} + \sigma^2_{\omega_k}) + \sigma^2_n && (13) \\ G_k &= \frac{\partial g(\theta_k)}{\partial \theta_k} |_{\theta_k = \bar{\theta}_{k-1|k-1}} = j e^{j\bar{\theta}_{k-1|k-1}} && (14) \\ g(\bar{\theta}_{k-1|k-1}) &= e^{j\bar{\theta}_{k-1|k-1}}. && (15) \end{aligned}$$

As a result, an estimate $\hat{q}(s_k, \theta_k|y_{1:k})$ on the joint posterior density $p(s_k, \theta_k|y_{1:k})$ can be derived analytically from (7) based on (8) and (10), i.e.,

$$\begin{aligned} &\hat{q}(s_k, \theta_k|y_{1:k}) \\ &\quad \propto \mathcal{N}(y_k^{new}|\bar{y}_k, \sigma^2_{y_k})\mathcal{N}(\theta_k|\bar{\theta}_{k-1|k-1}, \sigma^2_{\theta_{k-1|k-1}} + \sigma^2_{w_k}). \end{aligned} \tag{16}$$

Marginalizing (16) over θ_k, we obtain an estimate on $p(s_k|y_{1:k})$ as

$$\begin{aligned} r_{k,m} &= \hat{q}(s_k = a_m|y_{1:k}) \\ &= \int_{\theta_k} \hat{q}(s_k = a_m, \theta_k|y_{1:k})d\theta_k \\ &= \frac{\mathcal{N}(y_k|a_m G_k \bar{\theta}_{k-1|k-1}, \sigma^2_{y_k})}{Z} \quad \forall a_m \end{aligned} \tag{17}$$

where

$$Z = \sum_{s_k \in \mathcal{A}} \mathcal{N}(y_k|\bar{y}_k, \sigma^2_{y_k}). \tag{18}$$

Notice that $\hat{q}(s_k|y_{1:k})$ is already in the exponential family and no projection is thus needed. We then obtain our desired estimate on the APP at step k as

$$q(s_k|y_{1:k}) = \hat{q}(s_k|y_{1:k}) = Discrete(r_{k,1}, \cdots, r_{k,M}) \tag{19}$$

Next, we calculate an estimate on $p(\theta_k|y_{1:k})$, which is needed for computations at the $k+1$th step. Marginalizing (16) alternatively over s_k, we obtain an estimate as

$$\begin{aligned} \hat{q}(\theta_k|y_{1:k}) &= \sum_{s_k \in \mathcal{A}} \hat{q}(s_k, \theta_k|y_{1:k}) \\ &= \sum_{s_k \in \mathcal{A}} \hat{q}(\theta_k|s_k, y_{1:k}) q(s_k|y_{1:k}) \\ &= \sum_{m=1}^{M} \mathcal{N}(\theta_k|\ \bar{\theta}^{(m)}_{k|k}, \bar{\sigma}^2_{\theta_{k|k}}) r_{k,m} \end{aligned} \tag{20}$$

where

$$\bar{\theta}^{(m)}_{k|k} = \bar{\theta}_{k-1|k-1} + K_k(y_k - a_m g(\bar{\theta}_{k-1|k-1})) \tag{21}$$

$$\bar{\sigma}^2_{\theta_{k|k}} = \sigma^2{}_{\theta_{k|k-1}} - K_k a_m G_k \sigma^2{}_{\theta_{k|k-1}} \tag{22}$$

$$K_k = \sigma^2{}_{\theta_{k|k-1}} G^*_k a^*_m \sigma^{-2}_{y_k} \tag{23}$$

Note that the current estimate (20) is a mixture Gaussian distribution, which is not in the exponential family, and therefore projection must be performed. As a result of moment matching, we obtain an approximation of $\hat{q}(\theta_k|y_{1:k})$ by a single Gaussian distribution

$$q(\theta_k|y_{1:k}) = \mathcal{N}(\theta_k|\ \bar{\theta}_{k|k}, \sigma^2_{\theta_{k|k}}), \tag{24}$$

where

$$\bar{\theta}_{k|k} = \sum_{m=1}^{M} r_{k,m} \bar{\theta}^{(m)}_{k|k} \tag{25}$$

$$\sigma^2_{\theta_{k|k}} = \bar{\sigma}^2_{\theta_{k|k}} - \bar{\theta}^2_{k|k} + \sum_{m=1}^{M} r_{k,m} (\bar{\theta}^{(m)}_{k|k})^2. \tag{26}$$

Now, we obtained the desired filtering posterior distributions at k. As the last part of the filtering process, we compute the estimated likelihood functions as

$$\begin{aligned} q(y_k|s_k, \theta_k) &= q(y_k|\theta_k) q(y_k|s_k) \\ &= Z \frac{q(\theta_k, |y_{1:k})}{q(\theta_k|y_{1:k-1})} \frac{q(s_k, |y_{1:k})}{q(s_k|y_{1:k-1})} \end{aligned} \tag{27}$$

and obtain

$$q(y_k|\theta_k) \propto \mathcal{N}(\theta_k|\hat{\mu}_k, \hat{\lambda}_k) \tag{28}$$

$$q(y_k|s_k = a_m) \propto \bar{r}_{k,m} \tag{29}$$

where, to avoid numerical problems, $\hat{\lambda}_k$ and $\hat{\mu}_k$ are indirectly computed by a natural parameterization of the exponential family

$$\mu_k = \hat{\lambda}^{-1}_k \hat{\mu}_k = (\sigma^2_{\theta_{k|k}})^{-1} \bar{\theta}_{k|k} - (\sigma^2_{\theta_{k|k-1}})^{-1} \bar{\theta}_{k-1|k-1} \tag{30}$$

$$\lambda_k = \hat{\lambda}^{-1}_k = (\sigma^2_{\theta_{k|k}})^{-1} - (\sigma^2_{\theta_{k|k-1}})^{-1} \tag{31}$$

$$\bar{r}_{k,m} = \frac{r_{k,m}}{r_{k-1,m}} \quad \forall m \tag{32}$$

Next, in the smoothing process and at k, without explicit derivation, we first obtain the estimates on the smoothing posterior distributions from Kalman smoothing [8]

$$q(\theta_k|y_{1:K}) = \mathcal{N}(\theta_k|\ \bar{\theta}_{k|K}, \sigma^2_{\theta_{k|K}}) \tag{33}$$

$$q(s_k|y_{1:K}) = Discrete(r^s_{k,1}, \cdots, r^s_{k,M}) \tag{34}$$

where

$$J_k = \sigma^2_{\theta_{k|k}} (\sigma^2_{\theta_{k|k-1}})^{-1} \tag{35}$$

$$\bar{\theta}_{k|K} = \bar{\theta}_{k|k} + J_k(\bar{\theta}_{k+1|K} - \bar{\theta}_{k|k}) \tag{36}$$

$$\sigma^2_{\theta_{k|K}} = \sigma^2_{\theta_{k|k}} + J^2_k(\sigma^2_{\theta_{k+1|K}} - \sigma^2_{\theta_{k|k}}). \tag{37}$$

Then, the refinement part of EP on $q(\theta_k|y_{1:K})$ and $q(s_k|y_{1:K})$ is triggered. As discussed in section III, two steps are involved. In the first step, we remove the estimated likelihoods in (28) and (29) from the smoothing densities to obtained

$$q(\theta_k|y_{1:K,-k}) = \mathcal{N}(\theta_k|\ \bar{\theta}_{-k|K}, \sigma^2_{\theta_{-k|K}}) \tag{38}$$

$$q(s_k|y_{1:K,-k}) = Discrete(r^s_{-k,1}, \cdots, r^s_{-k,M}) \tag{39}$$

where

$$\sigma^2_{\theta_{k|K}} = ((\sigma^2_{\theta_{k|k}})^{-1} - \lambda_k)^{-1} \tag{40}$$

$$\bar{\theta}_{-k|K} = \sigma^2_{\theta_{k|K}}((\sigma^2_{\theta_{k|k}})^{-1} \bar{\theta}_{k|k} - \mu_k) \tag{41}$$

$$r^s_{-k,m} = r_{k,m}/\bar{r}_{k,m}. \tag{42}$$

Then, in the second step of the refinement, the true likelihood is incorporated and projection is performed to obtain an new estimation on the smoothing densities. The procedure is the same as in the filtering processes and we omit the detail. When the smoothing process is finished, more iterations of filtering and smoothing-refinement are carried out until the estimated smoothing posterior distributions converge. We, however, want to point out that in the filtering process of later iterations, only the estimated likelihoods from the previous smoothing process are incorporate and therefore linearization is no longer needed and the moments of the new estimates can be conveniently calculated by

$$\sigma^2_{\theta_{k|k}} = ((\sigma^2{}_{\theta_{k|k-1}})^{-1} + \lambda_k)^{-1} \tag{43}$$

$$\bar{\theta}_{k|k} = \sigma^2_{\theta_{k|k}}(\mu_k + (\sigma^2{}_{\theta_{k|k-1}})^{-1} \bar{\theta}_{k-1|k-1}) \tag{44}$$

$$r_{k,m} = \bar{r}_{k,m} r^s_{-k,m} \tag{45}$$

The algorithm of EP for our problem can be summarized in the following:

---

1. Initial estimate
   - For k = 1 : K
     - Get the predictive density of θ_k from (9)
     - Moment matching via (25) to (26)
     - Obtain the estimated likelihood via (30) to (32)

2. Loop by increasing i until the maximum number of iteration is reached or convergence happens:

   - For k = 1 : K (Skip on the first iteration)

TABLE I
NUMBER OF ITERATIONS AT $\sigma = 0.01$ *and* 0.02 WITH BLOCK DATA LENGTH K=5000.

SNR (dB)	0.5	1	1.5	2	2.5	3	3.5	4
# of iteration ($\sigma = 0.01$)	9	7	7	6	6	6	6	5
# of iteration ($\sigma = 0.02$)	9	8	8	7	7	7	7	6

  - Get the predictive density of θ_k from (9)
  - Incorporate the true likelihoods into posterior distribution via (43) to (45)
- For k = K : -1 : 1
  - Kalman smoothing via (35) to (37) when $k < K$
  - Delete old likelihood via (40) to (42)
  - Moment matching via (25) to (26)
  - Obtain new likelihood for the next refinement via (30) to (32)

In the very end, when $p(s_k|y_{1:K})$s are obtained, the maximum *a posteriori* detection can be performed as

$$\hat{s}_k = \arg\max_{s_k \in \mathcal{A}} p(s_k|y_{1:K}).$$

It is interesting to see that the initial estimate process is very similar to a extended Kalman smoother (EKS). However, EP differs from EKS with a refinement process that recycles the likelihood. The refinement process enables continuous improvement on the global approximation whenever a local improvement is produced and such improvement will not stop until the globally best solutions are reached. As a result, EP can produce much better approximations than EKS.

V. SIMULATION RESULTS

We performed simulations to demonstrate the performance of the proposed EP detection. In our simulation, the data symbols were transmitted by blocks with the block length $K = 5000$. The data were differentially encoded to combat phase ambiguity. We also assumed the unit energy for symbols and thus the signal-to-noise ratio (SNR) is calculated by $-10\log\sigma_n^2$. Bit-error-rates (BERs) of different detectors were calculated at different SNRs as an evaluation of performance. In particular, to obtain a desired BER, we require at least 50 errors be collected.

First, in Figure 1, we plotted the BER vs. SNR for EP, differential detection, and detector with known phase when symbols are BPSK modulated. The BER of the detector with known phase is served as the lower bound. In this experiment, we set $\sigma_w = 0.05$. We see clearly that the performance EP is better than differential detection and almost overlaps with that of the detector with known phase. In Figure 2, we tested QPSK modulation. We can see the performance of EP is close to the performance of known phase, but is much better than differential detection especially at high SNR.

Next, we explore the convergence of EP by examining the number of iterations at different phase noise variance. We presented the results in Table I for $\sigma_w = 0.01$ and 0.02. We see that, overall, EP converges in less than 10 iterations. In addition, we observe that the number of iterations decreases with the increase of SNR and/or phase variance. Since high SNR region is of more interest for practical transmission, the reduced complexity of EP at high SNRs is a potentially attractive feature.

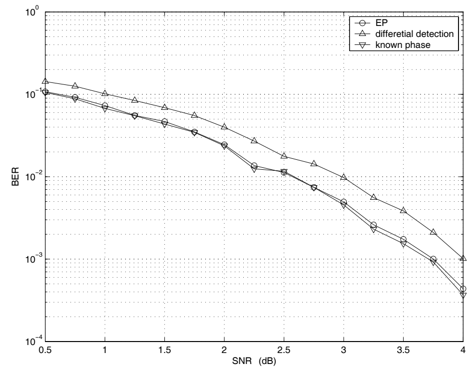

Fig. 1. Plots of BER vs. SNR for BPSK.

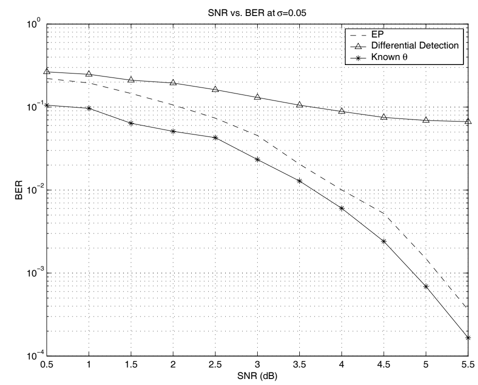

Fig. 2. Plots of BER vs. SNR for QPSK.

VI. CONCLUSIONS

In this paper, we presented a novel solution based on expectation propagation to symbol detection in the presence of dynamic phase change. We demonstrated through simulation very good performance of EP and its relative small complexity.

REFERENCES

[1] T. J. Lim and L. K. Rasmussen, "Adaptive symbol and parameter estimation in asynchronous multiuser cdma detectors," *IEEE Trans. Commun.*, vol. 45, pp. 213–220, Jan. 1997.

[2] J. Caffery and G. L. Stüber, "Nonlinear multiuser parameter estimation and tracking in CDMA systems," *IEEE Trans on Commun.*, vol. 48, no. 12, pp. 2053–2063, Dec 2000.

[3] Y. Yin and Y. Huang, "Joint symbol detection and timing estimation with stochastic m-algorithm," in *Proceedings of the IEEE International Conference on Acoustics,Speech, and Signal Processing*, May 2004.

[4] J. Dauwels and H.-A. Loeliger, "Phase estimation by message passing," in *Proc. 2004 IEEE Int. Conf. on Communications*, Zurich, Switzerland, June 2004.

[5] V. K. H.-A. Loeliger, J. Dauwels and S. Korl, "Signal processing with factor graphs: examples," in *Proc. First Int. Symp. on Control, Communications and Signal Processing*, March 2004.

[6] T. P. Minka, "A family of algorithms for approximate bayesian inference," Ph.D. dissertation, MIT, Jan 2001.

[7] Y. Qi, "Extending expectation propagation for graphical models," Ph.D. dissertation, MIT, Oct 2004.

[8] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms*. Prentice Hall, 2000.