Joint Maximum Likelihood and Expectation Maximization methods for Unsupervised Iterative Soft Bit Error Rate Estimation

Samir Saoudi*, Jia Dong* and Tarik Ait-Idir†
*Institut Mines-Telecom / Telecom Bretagne / LabSticc / UeB, Brest, France
Email: samir.saoudi@telecom-bretagne.eu
†INPT-Rabat, Morocco,
Email: aitdir@ieee.org

***Abstract*—This paper addresses the problem of unsupervised soft bit error rate (BER) estimation for any communications system, where no prior knowledge either about transmitted information bits, or the transceiver scheme is available. We show that the problem of BER estimation is equivalent to estimating the conditional probability density functions (pdf)s of soft receiver outputs. Assuming that the receiver has no analytical model of soft observations, we propose a non parametric Kernel-based pdf estimation technique, with Maximum Likelihood based smoothing parameter computation. We then introduce an iterative Stochastic Expectation Maximization algorithm for the estimation of both *a priori* and *a posteriori* probabilities of transmitted information bits, and the classification of soft observations according to transmitted bit values. These inputs serve in the iterative Kernel-based estimation procedure of conditional pdfs. We analyze the performance of the proposed unsupervised BER estimator in the framework of a multiuser code division multiple access (CDMA) system with single user detection, and show that attractive performance are achieved compared with conventional Monte Carlo-aided techniques.**

I. Introduction

Monte Carlo (MC) techniques are generally used to evaluate the bit error rate (BER) or block error rate (BLER) of digital communication systems. In [1], a tutorial exposition of different MC-aided techniques has been provided, with particular reference to four methods: i) Modified Monte Carlo simulation (i.e., importance sampling); ii) extreme value theory; iii) tail extrapolation; and iv) quasi-analytical method. Except for the MC method, all these techniques assume perfect knowledge of the type and/or the form of noise statistics. In [2], high order statistics of the bit log-likelihood-ratio are used to evaluate the error rate performance of turbo-like codes. In [3], we have suggested a BER estimation technique where the transmitter sends a fixed information bit value. This technique is called *soft* BER estimation since the BER is estimated using soft outputs without requiring hard decisions about information bits. This provides reliable BER estimates and reduces the number of required samples compared with hard decision-based BER estimation techniques. Unfortunately, all these methods assume that the estimator perfectly knows transmitted data.

In many practical communication systems, the BER is required to be on-line estimated in order to perform system-level functions such as scheduling, resource allocation, power control, or link adaptation where the transmission scheme is adapted to the channel conditions (See for instance [4]). Under this framework, the unsupervised BER estimation problem becomes very challenging because of the following main reasons:

- In MC-based techniques, the *unknown* transmitted information bit values are required for computing the BER estimate, while in practical communications systems the BER estimation should be performed in an *unsupervised* fashion because the estimator has no information about transmitted data.
- Most of practical communication channels are quasi-static block fading and randomly changes from block to block. Therefore, in order to provide the transmitter with a reliable BER information feedback for a given block, i.e., channel state information (CSI), only the soft observations corresponding to the actual block have to be used for estimating the instantaneous BER. In the case of MC-based techniques, this results in unreliable BER estimates because the number of observations is not sufficient.
- The knowledge of the transmitter scheme, channel and interference model, and receiver technique greatly impacts the reliability of the BER estimate. In other words, it is generally quite hard to derive analytical expressions of the bit error probability (BEP) when the system suffers from interference or in the case of non-linear receivers.

This paper was mainly motivated by the above considerations. We assume that the estimator has no knowledge either about transmitted data, or the transmitter/receiver scheme and the communication model. We focus on the problem of BER estimation in a completely unsupervised fashion. We first provide a problem formulation where we show that BER estimation is equivalent to estimating the pdfs of conditional soft observations corresponding to transmitted information

978-1-4673-1881-5/12/$31.00 ©2012 IEEE

bits. Then, we introduce a non parametric Kernel-based pdf estimation technique where no analytical model of soft observations is assumed. As pdf estimation requires the knowledge of both *a priori* probabilities of information bits and the classification of soft observations according to the $+1$ or -1 values of transmitted bits, we introduce a Stochastic Expectation Maximization (EM) algorithm for iteratively computing these parameters. We analyze the performance of the proposed unsupervised technique in the case of a multiuser code division multiple access (CDMA) system with conventional single-user detection. Performance comparison shows that the proposed estimator clearly outperforms supervised MC-aided estimation. Interestingly, reliable BER estimates are achieved even in the high signal to noise ratio (SNR) and using only a few number of soft observations.

The remainder of the paper is organized as follows. In Section II, we provide some preliminaries about the proposed BER estimator and detail soft output pdf estimation; In Section III, we introduce the Kernel-based BER estimator; Section IV details the Stochastic EM iterative algorithm; In Section V, we carry out performance evaluation; The paper is concluded in Section VI.

II. Preliminaries and Output PDF Estimation

A. System Model

We consider a general communication system where a sequence of N independent and identically distributed (i.i.d) information bits $(b_i)_{1\leq i\leq N} \in \{+1, -1\}$ is transmitted using any transmission scheme.

Let X denote the random variable (RV) corresponding to receiver soft outputs and $(X_i)_{1\leq i\leq N}$ be the sequence of the realizations of X where each X_i corresponds to the decision statistic that serves for the computation of hard decision $\hat{b}_i$ about transmitted information bit b_i as $\hat{b}_i = sgn\,(X_i)$. Note that X_i may contain any type of interference such as co-channel interference (CCI), intersymbol interference (ISI), and multiple antenna interference (MAI), etc... . We assume no knowledge either about the transmission scheme, or about the channel model and the receiver technique. Let π_+ and π_- denote the probability that the transmitted bit b_i is equal to $+1$ and -1, respectively, i.e.,

$$\begin{cases} \pi_+ & \triangleq P\left[b_i = +1\right], \\ \pi_- & \triangleq P\left[b_i = -1\right], \end{cases} \tag{1}$$

where $\pi_+ + \pi_- = 1$. Note that π_+ and π_- are not known at the receiver. The soft outputs $(X_i)_{1\leq i\leq N}$ are random variables having the same pdf $f_X(x)$. The BEP is then given by

$$p_e \;=\; \pi_+ \int_{-\infty}^{0} f_X^{b_+}(x)\; dx + \pi_- \int_{0}^{+\infty} f_X^{b_-}(x)\; dx, \tag{2}$$

where $f_X^{b_+}(\cdot)$ (respectively, $f_X^{b_-}(\cdot)$) is the conditional pdf of X such that $b_i = +1$ (respectively, $b_i = -1$). $f_X(x)$ is a mixture of the two conditional pdfs $f_X^{b_+}(x)$ and $f_X^{b_-}(x)$ and can therefore be written as

$$f_X(x) \;=\; \pi_+ f_X^{b_+}(x) + \pi_- f_X^{b_-}(x). \tag{3}$$

B. Brief Overview of the Proposed BER Estimator

As it can be seen from the generic expression of the BEP p_e in (2), the BER can be estimated using the two conditional pdfs $f_X^{b_+}(\cdot)$, $f_X^{b_-}(\cdot)$, and the *a priori* probabilities π_+ and π_-. Note that all these parameters are unknown, and are required to be estimated before computing the BER. The estimation of conditional pdfs $f_X^{b_+}(\cdot)$ and $f_X^{b_-}(\cdot)$ can be performed based on observations $X_1, \cdots, X_N$, where only those corresponding to transmitted information bit value $+1$ (respectively, -1) are used to estimate $f_X^{b_+}(\cdot)$ (respectively, $f_X^{b_-}(\cdot)$). Unfortunately, the problem of classifying soft outputs $X_1, \cdots, X_N$ according to transmitted information bit values is itself a detection problem. The unsupervised BER estimator we propose in this paper allows us to classify soft outputs $X_1, \cdots, X_N$ and estimate both the conditional pdfs $f_X^{b_+}(\cdot)$ and $f_X^{b_-}(\cdot)$ and *a priori* probabilities π_+ and π_- in an iterative fashion using *a posteriori* probability (APP) values $P[b_i = +1 \mid X_i]$ and $P[b_i = -1 \mid X_i]$ $\forall\; i = 1, \cdots, N$.

Let T denote the number of iterations (index $t = 1, \cdots, T$). At each iteration t, all APP values are computed using *a priori* probabilities and pdf estimates obtained at the previous iteration $t-1$. The resulting APPs allow us to update the estimates of *a priori* probabilities and classify soft outputs $X_1, \cdots, X_N$ into two classes according to transmitted information bit values. The two conditional pdfs are then re-estimated. This process is repeated T times, and the BER is estimated with the aid of parameter estimates obtained at the last iteration T. Note that, at the first iteration, since no prior information is available, the classification of soft outputs $X_1, \cdots, X_N$ can be performed with respect to the sign of each observation X_i, while the initial values of *a priori* probabilities π_+ and π_- can be derived using this initial classification. Fig. 1 shows the diagram of the proposed iterative BER estimation algorithm.

C. Output PDF Estimation

Both conditional pdfs $f_X^{b_+}(x)$ and $f_X^{b_-}(x)$ depend on the communication channel model and receiver scheme. Therefore, it is extremely difficult to find out the exact parametric model of these distributions. In this subsection, we introduce a Kernel-based method [6] to estimate both pdfs $f_X^{b_+}(x)$ and $f_X^{b_-}(x)$. The proposed technique is non parametric, and only requires soft outputs for each class.

Let us assume that we can classify the set $\mathcal{C} = \{X_1, \cdots, X_N\}$ into two classes (i.e., partitions) $\mathcal{C}_+$ and $\mathcal{C}_-$, where $\mathcal{C}_+$ (respectively, $\mathcal{C}_-$) contains the observed received soft output X_i such that the corresponding transmitted bit is $b_i = +1$ (respectively, $b_i = -1$). In Section IV, we will introduce a Stochastic EM-based algorithm that allows us to classify the sets $\mathcal{C}_+$ and $\mathcal{C}_-$. Let N_+ (respectively, N_-)

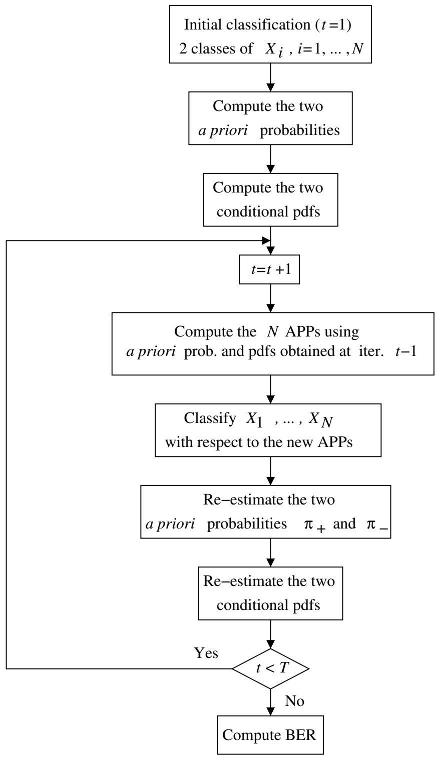

Figure 1. The diagram of the proposed iterative BER estimation algorithm.

denote the cardinality of $\mathcal{C}_+$ (respectively, $\mathcal{C}_-$). Using the Kernel technique, it follows that each conditional distribution can be estimated using the elements of sets $\mathcal{C}_+$ and $\mathcal{C}_-$ as,

$$\hat{f}^{b_+}_{X,N_+}(x) = \frac{1}{N_+ h_{N_+}} \sum_{X_i \in \mathcal{C}_+} K\left(\frac{x - X_i}{h_{N_+}}\right), \tag{4}$$

$$\hat{f}^{b_-}_{X,N_-}(x) = \frac{1}{N_- h_{N_-}} \sum_{X_i \in \mathcal{C}_-} K\left(\frac{x - X_i}{h_{N_-}}\right), \tag{5}$$

where h_{N_+} (respectively, h_{N_-}) is a smoothing parameter which depends on the number of observed samples, i.e., N_+ (respectively, N_-). Function $K(\cdot)$ is any pdf (called the Kernel) with zero mean and unit variance. In the following, we provide equations only in the case of class C_+ as those corresponding to C_- can easily be obtained by replacing "$+$" by "$-$".

The choice of the optimal smoothing parameter is very critical since it impacts the accuracy of the pdf estimate. It has been shown in [3] that the optimal smoothing parameter $h^*_{N_+}$ in the integrated mean squared error (IMSE) criterion is difficult to compute. Here, in this paper, we will show how the optimal smoothing parameter h_{N_+} can be derived in the Maximum Likelihood criterion. To use the Maximum Likelihood method, we firstly specify the joint density function of all observations. For an independent and identically distributed sample of the set $\mathcal{C}_+$, the log-likelihood function can be given as

$$\ln\left(\prod_{X_i \in \mathcal{C}_+} \hat{f}^{b_+}_{X,N_+}(X_i)\right) = \sum_{X_i \in \mathcal{C}_+} \ln\left(\frac{1}{N_+ h_{N_+}} \sum_{j \neq i} K\left(\frac{X_i - X_j}{h_{N_+}}\right)\right) \tag{6}$$

The optimal smoothing parameter $h^*_{N_+}$ is computed by canceling the derivative of log-likelihood function. We suggest to use Newton's method by finding successively better approximations to the root. We begin with an initial condition $h^{(0)}_{N_+} = \left(\frac{4}{3N_+}\right)^{\frac{1}{5}} \sigma_+$, where σ^2_+ is the variance of $\mathcal{C}_+$. Then at each iteration t, a new value of $h^{(t)}_{N_+}$ is given by using the previous value $h^{(t-1)}_{N_+}$ such as

$$h^{(t)}_{N_+} = h^{(t-1)}_{N_+} - \frac{g(h^{(t-1)}_{N_+})}{g'(h^{(t-1)}_{N_+})}, \tag{7}$$

where $g(h_{N_+}) = \left[\ln\left(\prod_{X_i \in \mathcal{C}_+} \hat{f}^{b_+}_{X,N_+}(X_i)\right)\right]'$.

Let us note that expression (7) can directly be given by only using the soft output observations. Indeed, let us denote, for any integer $n \geq 0$, $A_{n,i} = \sum_{j \neq i} (X_i - X_j)^n K(\frac{X_i - X_j}{h})$, we can easily show that:

$$g(h) = -h^2 N + \sum_{i=1}^{N} \frac{A_{2,i}}{A_{0,i}} \tag{8}$$

and,

$$g'(h) = -2hN + \frac{1}{h^3} \sum_{i=1}^{N} \left(\frac{A_{4,i}}{A_{0,i}} - \frac{A^2_{2,i}}{A^2_{0,i}}\right) \tag{9}$$

The proposed initial value $h^{(0)}_{N_+}$ is optimal in the IMSE criterion for Gaussian outputs distribution and Gaussian Kernel.

III. BER Estimation

In this section, we derive the expression of the BER estimate assuming Gaussian Kernel-based pdf estimator. Let $\theta \triangleq \left(\pi_+, N_+, h_{N_+}, \pi_-, N_-, h_{N_-}\right)$. The value of θ is unknown and is estimated by iteratively classifying soft outputs $X_1, \cdots, X_N$ into two classes (see sectionIV). At the last iteration T, a reliable estimate of θ is reached and the BER can be computed using the obtained estimate. Let us recall the expression of the BEP (2). Replacing the two conditional pdfs by their Gaussian Kernel-based estimates (4) and (5), and given the value of θ, the BER estimate is simply computed as

$$\hat{p}_{e,N} = \frac{\pi_+}{N_+} \sum_{X_i \in \mathcal{C}_+} Q\left(\frac{X_i}{h_{N_+}}\right) + \frac{\pi_-}{N_-} \sum_{X_i \in \mathcal{C}_-} Q\left(-\frac{X_i}{h_{N_-}}\right), \tag{10}$$

where $Q(\cdot)$ denotes the complementary unit cumulative Gaussian distribution.

IV. Stochastic EM-Based Parameter Estimation

In this section, we introduce a Stochastic EM-based algorithm to iteratively classify soft outputs $X_1, \cdots, X_N$ into two classes $\mathcal{C}_+$ and $\mathcal{C}_-$, and estimate θ. The EM algorithm iteratively computes, with the aid of the estimation and maximization steps, maximum likelihood (ML) estimates of different missing information.

A. Estimation Step

In the estimation step of iteration t, we estimate the APPs

$$\begin{cases} \rho_{i_+}^{(t)} & \triangleq P[b_i = +1 \mid X_i, \theta^{(t-1)}], \\ \rho_{i_-}^{(t)} & \triangleq P[b_i = -1 \mid X_i, \theta^{(t-1)}], \end{cases} \tag{11}$$

of unobserved information bits b_i, for $i = 1, \cdots, N$, conditioned on observations $X_1, \cdots, X_N$ and the estimate $\theta^{(t-1)}$ of θ obtained at the maximization step of previous iteration $t-1$. Using simple mathematical manipulations, we show that the likelihood probabilities of bit b_i at iteration t can be computed as,

$$\begin{cases} \rho_{i_+}^{(t)} & = \frac{\pi_+^{(t-1)} \hat{f}^{b_+}_{X,N_+^{(t-1)}}(X_i)}{\pi_+^{(t-1)} \hat{f}^{b_+}_{X,N_+^{(t-1)}}(X_i) + \pi_-^{(t-1)} \hat{f}^{b_-}_{X,N_-^{(t-1)}}(X_i)}, \\ \rho_{i_-}^{(t)} & = \frac{\pi_-^{(t-1)} \hat{f}^{b_-}_{X,N_-^{(t-1)}}(X_i)}{\pi_+^{(t-1)} \hat{f}^{b_+}_{X,N_+^{(t-1)}}(X_i) + \pi_-^{(t-1)} \hat{f}^{b_-}_{X,N_-^{(t-1)}}(X_i)}. \end{cases} \tag{12}$$

B. Maximization Step

At iteration t, the maximization step allows us to compute the estimate $\theta^{(t)}$ based on conditional probabilities obtained at the estimation step of the same iteration t. The estimate $\theta^{(t)}$ is obtained by maximizing the conditional expectation of the log-likelihood of the joint event at iteration t which leads to the new estimates of *a priori* probabilities,

$$\begin{cases} \pi_+^{(t)} & = \frac{1}{N} \sum_{i=1}^N \rho_{i_+}^{(t)}, \\ \pi_-^{(t)} & = \frac{1}{N} \sum_{i=1}^N \rho_{i_-}^{(t)}. \end{cases} \tag{13}$$

C. Classification Step

The remaining parameters $N_+^{(t)}$, $h_{N_+}^{(t)}$, $N_-^{(t)}$, $h_{N_-}^{(t)}$ and the two conditional pdf estimates $\hat{f}^{b_+}_{X,N_+^{(t)}}(.)$ and $\hat{f}^{b_-}_{X,N_-^{(t)}}(.)$ depend on the outcome of the classification procedure of subsets $\mathcal{C}_+$ and $\mathcal{C}_-$ at iteration t. Therefore, this procedure should be carefully performed since it greatly impacts the reliability of both pdf estimates and *a posteriori* probabilities in subsequent iterations, and consequently the accuracy of the BER estimate. Given estimates $\theta^{(t-1)}$, $\hat{f}^{b_+}_{X,N_+^{(t-1)}}(.)$, and $\hat{f}^{b_-}_{X,N_-^{(t-1)}}(.)$ available at iteration t, we can classify soft outputs $X_1, \cdots, X_N$ according to joint probabilities $P\left[X_i, b_i = +1\right]$, and $P\left[X_i, b_i = -1\right]$.

The Stochastic EM technique uses a random Bayesian rule. At iteration t, it combines likelihood probabilities (12) and realizations $U_1^{(t)}, \cdots, U_N^{(t)}$ of a uniform random variable U defined over the interval $[0,1]$. The classification of soft outputs is performed as follows,

$$\begin{cases} \mathcal{C}_+^{(t)} & = \left\{ X_i : \ \rho_{i_+}^{(t)} \geq U_i^{(t)} \right\}, \\ \mathcal{C}_-^{(t)} & = \bar{\mathcal{C}}_+^{(t)}. \end{cases} \tag{14}$$

The optimal smoothing parameters $h_{N_+}^{(t)}$ and $h_{N_-}^{(t)}$ are therefore computed with the aid of Newton's method as suggested in subsection II-C.

V. Performance Evaluation

A. Considered Framework

To evaluate the performance of the proposed unsupervised BER estimator, we consider the framework of a synchronous CDMA system with two users using binary phase-shift keying (BPSK) and operating over an additive white Gaussian noise (AWGN) channel. We restrict ourselves to the conventional single user CDMA detector. Performance assessment in the case of advanced signaling/receivers is not reported in this paper due to space limitation and is left for future contributions.

In all simulations, we consider $T = 6$ iterations for the Stochastic EM-based parameter estimation while at each iteration $t = 1, \cdots, T$, four iterations are used for computing the optimal smoothing parameters $h_{N_+}^*$ and $h_{N_-}^*$ as mentioned in Subsection II-C.

B. Numerical Results

1) Performance for Uniform Sources: First, we consider the case of equiprobable information bits, i.e., ($\pi_+ = \pi_- = 1/2$). The number of soft outputs that serve for estimating the BER is $N = 10^4$ observations. In Fig. 2 we present both the theoretical and the estimated conditional pdfs (at the last iteration $T = 6$) for $\mathrm{SNR} = 10$dB. We observe that the proposed BER estimator provides accurate estimates of conditional pdfs. In Fig. 3, we provide the BER estimation performance. We notice that the proposed unsupervised BER estimator offers the same performance as the MC-aided method where this last one has perfect knowledge about the transmitted information bits, i.e., perfect classification of soft outputs.

2) Performance for Non Uniform Sources: We now turn to the case when the information bits are not equiprobable. We consider the scenario where $\pi_+ = 0.75$ and $\pi_- = 0.25$. The number of soft outputs is kept to $N = 10^4$. In Fig. 4, we report both the theoretical and estimated weighted conditional pdfs for $\mathrm{SNR} = 10$dB. A quick inspection of the performance graph shows that even if the *a priori* probabilities are not equal, the proposed unsupervised BER estimator achieves reliable estimates of conditional pdfs for

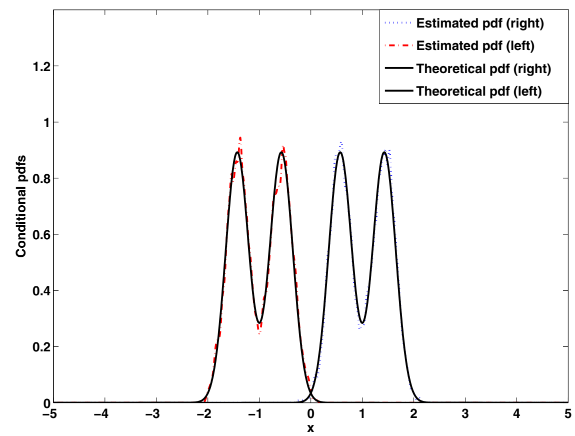

Figure 2. Estimated conditional pdfs for $\pi_+ = \pi_- = 1/2$, and $\mathrm{SNR} = 10\mathrm{dB}$.

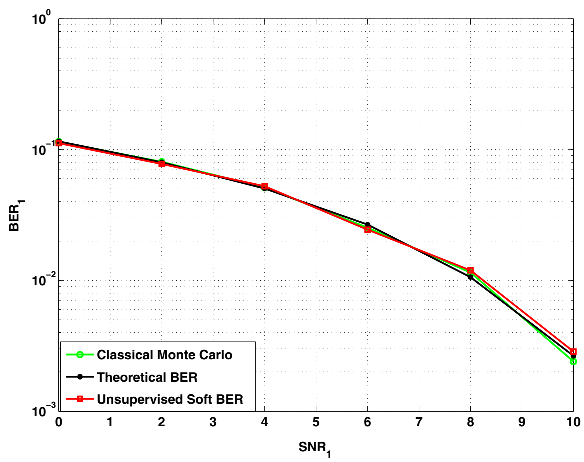

Figure 3. BER performance comparison when $\pi_+ = \pi_- = 1/2$, and $N = 10^4$ soft observations.

high SNR independently of the distribution of information bits.

VI. Conclusions

In this paper, we considered the problem of unsupervised BER estimation for any communication system using any signal processing technique. We proposed a BER estimation algorithm where only soft observations that serve for computing hard decisions about information bits are used to estimate the BER, and no prior knowledge about transmitted information bits is required. First of all, we provided a formulation of the problem where we showed that BER estimation is equivalent to the estimation of conditional pdfs of soft observations. We then proposed a BER computation technique using Gaussian Kernel-based pdf estimation. Then, we introduced an iterative Stochastic EM technique

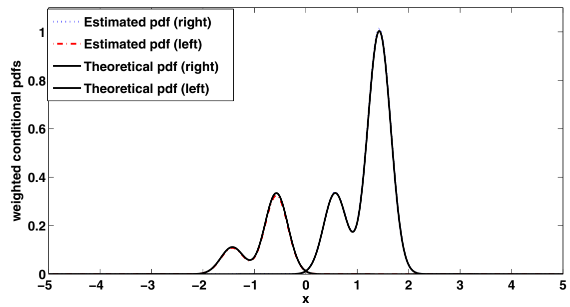

Figure 4. Estimated conditional pdfs for $\pi_+ = 0.75$, $\pi_- = 0.25$, and $\mathrm{SNR} = 10\mathrm{dB}$.

to compute the parameters that serve for the estimation of conditional pdfs based only on soft observations. The proposed method involves the EM steps to estimate the *a priori* probabilities of transmitted information bits, and a Stochastic classification step to classify soft observations according to information bits. Finally, we evaluated the performance of the proposed unsupervised BER estimation technique in the framework of CDMA systems to corroborate the theoretical analysis. Interestingly, we showed that the proposed estimator provides reliable estimates using only few soft observations.

References

[1] M.C. Jeruchim, “Techniques for estimating the bit error rate in the simulation of digital communication systems,” *IEEE J. Selec. Areas, Comm.,* vol. 2, no. 1, pp. 153-170, Jan. 1984.

[2] A. Abedi, M.E. Thompson and A.K. Khandani, “Application of cumulant method in performance evaluation of turbo-like codes,” *IEEE Intern. Conf. Commun., ICC’07,* 24-28 Jun. 2007 pp. 986 – 989.

[3] S. Saoudi, M. Troudi and F. Ghorbel, “An iterative soft bit error rate estimation of any digital communication systems using a non parametric probability density function,” *Eurasip Journal on wirless Communications and Networking,* Hindawi Publishing Corporation, Volume 2009, Article ID512192, 9 pages, doi:10.1155/2009/512192.

[4] M-S. Alouini, A. J. Goldsmith, “Capacity of rayleigh fading channels under different adaptive transmission and diversity-combining techniques,” *IEEE Trans. Veh. Technol,* vol. 48, no. 4, Jul. 1999.

[5] M. Tüchler, R. Koetter, and A. Singer, “Turbo equalization: principles and new results,” *IEEE Trans. Commun.,* vol. 50, no. 5, pp. 754–767, May 2002.

[6] E. Parzen, “On estimation of a probability density function and mode,” *Ann. Math. Statist.,* vol. 33, Issue 3, pp. 1065–1076, 1962.

[7] S. Saoudi, T. Derham, T. Ait-Idir and P. Coupé, “A fast soft bit error rate estimation method ” *EURASIP journal on wireless communications and networking*, Hindawi Publishing Corporation, volume 2010. doi:10.1155/2010/372370.

[8] S. Saoudi, F. Ghorbel and A. Hillion, “Some statistical properties of the kernel-diffeomorphism estimator,” *Applied Stochastic Models and Data Analysis Journal,* John Wiley, Vol. 13, Issue 1, pp. 39-58, March 1997.

[9] A. P. Dempster, N. M. Laird, D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society.* Series B, vol. 39, No. 1 (1977), pp. 1–38

[10] S. Saoudi, T. Ait-Idir and Y. Mochida, “A novel non-parametric iterative soft bit error rate estimation technique for digital communications systems,” *ICC2011: IEEE International Conference on Communications*, IEEE, Kyoto, Japan, 2011, pp. 1-6, 05-09 june 2011.