A MAXIMUM LIKELIHOOD CDMA RECEIVER USING THE EM ALGORITHM AND THE DISCRETE WAVELET TRANSFORM

Ilan Sharfer *Alfred O. Hero*

Dept. of Electrical Engineering and Computer Science
The University of Michigan, Ann Arbor, MI 48109

ABSTRACT

A Maximum Likelihood (ML) method for joint estimation of amplitude, phase, time delay, and data demodulation in a single-user direct sequence spread spectrum communication system is developed. The likelihood function is analytically intractable, so a recursive estimation algorithm is considered. The Expectation Maximization (EM) algorithm has been used in similar problems, however, in this case it is not computationally efficient. Recently, a variant of the EM algorithm, called Space Alternating Generalized EM (SAGE), has been derived. In this work we apply the SAGE algorithm to the sequence estimation problem in a way which results in simple sequential updates of all the estimated parameters. An important feature of the proposed algorithm is the use of a discrete wavelet decomposition of the received signal as a sufficient statistic. The consequence is that all the information is still available to the receiver, while the complicated estimation problem is considerably simplified. Computer simulations of a single user system were performed. It is shown that the algorithm has fast convergence, and essentially achieves optimal performance.

1. INTRODUCTION

The emergence of new technologies of multi-user wireless communication systems, requires advanced signal processing methods for improved efficiency and reliability. In order to optimally decode the desired information the receiver should benefit from knowledge of the nuisance parameters of the received signal which typically consist of the amplitude, phase, and time delay. Usually, these parameters are estimated by a combination of several techniques, each specialized to a particular parameter. For example, carrier phase and time delay estimation are mostly done with a Phase Locked Loop (PLL) and a Delay Locked Loop (DLL) respectively [4]. A considerable research activity has been directed at improving the performance of these basic synchronization techniques, e.g. by using decision feedback in a Data Aided Loop (DAL) configuration [5]. Clearly, an optimal receiver is one which jointly estimates the nuisance parameters as well as the data symbols.

In this work we consider the problem of Maximum Likelihood (ML) estimation of all the parameters given the received signal. This problem is analytically intractable even for the simple AWGN channel, hence the need for a recursive estimation algorithm. The Space Alternating Generalized EM (SAGE) algorithm which has recently been developed in [3] is a variant of the EM algorithm [2], both are recursive algorithms which generate a sequence of parameter estimates whose likelihood increases monotonically. The SAGE algorithm is more flexible than the EM algorithm, because it is possible to update subsets of the parameter, resulting in simpler updates and faster convergence. We have chosen to focus on the single user problem, although the same approach can be generalized to the multi-user case.

One of the basic concepts in estimation theory is that a signal can be represented by a sufficient statistic. The signal expansion on an orthonormal wavelet basis is one such possible representation. This choice has been motivated by the excellent time-frequency localization properties of the wavelet bases. The SAGE algorithm has therefore been formulated in terms of the decomposition of the received signal in the discrete wavelet transform domain. The result is a causal, fully digital receiver which makes a "single pass" on the information, without requiring any buffering or delayed processing. The principle of using a sufficient statistic in a similar problem has been discussed in [6].

This paper is organized as follows. In section 2 we define the system model, and briefly review the SAGE algorithm and the concept of a hidden data space. In section 3 we develop the single user algorithm, and outline a strategy of choosing the hidden data spaces. In section 4 we describe a Fourier based method for numerically solving the maximization step of the algorithm. The recursive implementation of the algorithm is given in section 5. We conclude with simulation results and a performance comparison with other techniques.

2. SYSTEM MODEL AND SAGE REVIEW

We consider the following single-user complex baseband CDMA model:

$$y(t) = a \sum_{n=0}^{N-1} b_n p(t - nT_p - d) + u(t) \qquad -\infty < t < \infty\,.$$

The unknown parameters are the complex gain $a \in \mathbb{C} - \{0\}$, the data symbols $b_n \in \{e^{i\beta} | \beta \in [-\pi, \pi)\}, n = 0, \ldots, N-1$, and the time delay $d \in [-T_d/2, T_d/2]$. The data symbol model corresponds to phase modulation formats such as BPSK or QPSK. The signaling waveform $p(t)$ is a known PN code. For simplicity we assume that the period of the PN code is equal to T_p. Notice that we do not assume that $p(t)$ has a compact support, so the algorithm can be used for ISI channels as well. The noise $u(t)$ is a complex white Gaussian process with power spectral density $N_0/2$. Following is a brief review of the SAGE algorithm [3].

Let θ be a parameter taking values in a parameter space Θ, which is a subset of the p dimensional Euclidean space $\mathbb{R}^p$ ($\mathbb{C}^p$). The goal is to find the (penalized) maximum likelihood estimate of θ given the observation $Y = y$. We define an index set S to be a non-empty subset of the set of integers $I_p = \{1, \cdots, p\}$, and we denote its complement by $\bar{S}$. Corresponding to these index sets we define θ_S and $\theta_{\bar{S}}$ as the elements of θ indexed by S and $\bar{S}$ respectively. In the SAGE algorithm the maximization of the penalized log likelihood function is replaced by a maximization of a sequence of other objectives $\{\phi^{S^{(i)}}(\theta_S; \theta^{(i-1)})\}_{i=1,2,\ldots}$. For this purpose a random variable X^S, called *hidden data space*, is defined such that it satisfies the following admissibility condition

$$p(y, x; \theta) = p(y|x; \theta_{\bar{S}})p(x; \theta), \quad x \in X^S, y \in Y, \tag{1}$$

i.e. the conditional pdf of Y given X^S does not depend on θ_S. This condition includes the EM complete data space requirement as a special case. The penalized SAGE objective is given by

$$\begin{aligned}\phi^S(\theta_S; \theta') &= Q^S(\theta_S; \theta') - P(\theta_S, \theta'_{\bar{S}}) \qquad (2)\\ Q^S(\theta_S; \theta') &\equiv E\{\log p(X^S; \theta_S, \theta'_{\bar{S}}|Y = y; \theta')\}\end{aligned}$$

where $P(\theta)$ is an optional penalty function. The SAGE algorithm generates a sequence of estimates $\{\theta^{(i)} : i = 0, 1, \ldots\}$ starting from an initial parameter guess $\theta^{(0)}$, by defining a sequence of hidden data spaces $X^{S^{(i)}}$, computing $\phi^{S^{(i)}}(\theta_S; \theta^{(i-1)})$ using (2), and maximizing it over θ_S.

3. SINGLE-USER SAGE ALGORITHM

Let $\{\psi_{jk}; j, k \in \mathbb{Z}\}$ be a real orthonormal wavelet basis derived from some multiresolution analysis of $L^2(\mathbb{R})$. This orthonormal basis is obtained by scaling and translations of a fixed function $\psi(t)$, i.e.

$$\psi_{jk}(t) = 2^{-j/2}\psi(2^{-j}t - k) \qquad j, k \in \mathbb{Z}.$$

The observation $y(t)$ is decomposed in this basis, yielding a set of random variables:

$$Y_{jk} = \langle y, \psi_{jk}\rangle \quad j, k \in \mathbb{Z}, \tag{3}$$

where $\langle f, g\rangle$ denotes the inner product $\int f(t)g^*(t)dt$. The set $Y = \{Y_{jk}; j, k \in \mathbb{Z}\}$ can be used to reconstruct the time signal $y(t)$ by the synthesis formula of the discrete wavelet transform, therefore Y is a sufficient statistic for the parameters. The random variables $\{Y_{jk}\}$ can also be written as:

$$Y_{jk} = a\sum_{n=0}^{N-1} b_n w_{jk}(n) + U_{jk} \quad j, k \in \mathbb{Z}, \tag{4}$$

where

$$w_{jk}(n) = \int p(t - nT_p - d)2^{-j/2}\psi(2^{-j}t - k)\, dt\,, \tag{5}$$

and U_{jk}, the projections of $u(t)$ on ψ_{jk}, are independent zero mean complex Gaussian random variables with variance $N_0/2$. Notice that the parameters $w_{jk}(n)$ are implicitly dependent on d, and that the observation model in (4) is multi-linear, as opposed to the original non-linear model of $y(t)$. The parameter which needs to be estimated is defined as $\theta = [a, b_n : n \in I_N, w_{jk}(n) : n \in I_N, j, k \in \mathbb{Z}]$ where $I_N = \{0, \ldots, N-1\}$.

We now outline the strategy of choosing the hidden data spaces X^S. Let $S \subset \mathbb{Z}^2$ be a non empty set of indices, and let $\bar{S}$ denote its complement. Define a set of random variables X^S as follows:

$$X^S = \{a\sum_{n=0}^{N-1} b_n w_{jk}(n) + U_{jk} \,:\, (j, k) \in S\}. \tag{6}$$

It is easy to show that if the parameter θ is chosen such that $w_{jk}(n) = 0$ when $(j, k) \in \bar{S}, n \in I_N$, then X^S is an admissible hidden data space. In this application we consider a sequence of hidden data spaces $\{X^{S^{(i)}}\}$, where the index sets $S^{(i)}, i \in \mathbb{N}$ satisfy the following conditions:

$$(i).\ S^{(i)} \subset S^{(i+1)}, i \in \mathbb{N} \qquad (ii). \bigcup_{i\in\mathbb{N}} S^{(i)} = \mathbb{Z}^2.$$

Thus, the algorithm incorporates new information in each iteration, which is used to update the previous estimates. In this way the algorithm makes only a "single pass" over the set of observations Y.

4. NUMERICAL SOLUTION OF THE MAXIMIZATION STEP

We now describe a method by which the maximization steps of the SAGE algorithm can be done numerically by solving a polynomial equation. This method is particularly useful in the case of time delay estimation, where one usually resorts to a high resolution line search. The same type of solution can be used for the data symbol estimate. The method is based on Fourier series expansion of the correlation parameters $w_{jk}(n)$ (see (5)), whose dependence on d is explicitly denoted by $w_{jk}(n; d)$. The Fourier series of $w_{jk}(n; d)$ on an interval $[-T_0/2, T0/2]$ can be written in the following z notation as

$$w_{jk}(n; d) = \sum_{m\in\mathbb{Z}} c_{jk,m}(n)z^m, \tag{7}$$

with $z = e^{i\omega_0 d}$ and $c_{jk,m}(n)$ given by

$$c_{jk,m}(n) = \frac{1}{\sqrt{T_0}}\int_{-T_0/2}^{T_0/2} w_{jk}(n; t)e^{-im\omega_0 t}dt\,.$$

It follows that the non penalized part of the objective can be written as a conjugate symmetric power series in z. Additionally, it may be necessary to include a penalty function, e.g. to account for the apriori density of z. This can be accomplished by choosing a complex function $\Pi(z)$ of the form:

$$\Pi(z) = \frac{G(z)G^*(z^{*-1})}{H(z)H^*(z^{*-1})}\,, \tag{8}$$

where $G(z)$ and $H(z)$ are polynomials over $\mathbb{C}$ such that $\Pi(z)$ does not have poles on the unit circle. The *penalized* objective can then be put in the form

$$\phi(z) = F(z) + F^*(z^{*-1}) + \log\, \Pi(z), \tag{9}$$

where $F(z)$ is a polynomial over $\mathbb{C}$. Notice that $\phi(z)$ in (9) is a real, differentiable function on the unit circle $z = e^{i\omega}$. It follows that the maximization step can be realized by differentiation of (9) with respect to z, numerically finding the roots which lie on the unit circle, and selecting the one which maximizes the penalized objective (9) to be the next parameter estimate.

5. RECURSIVE IMPLEMENTATION

In this section an efficient recursive implementation of the single-user SAGE algorithm is given. The basic assumption is that both $p(t)$ and $\psi(t)$ have compact support. We define the following variables:

$$\mu_m^{(i)}(n) = \sum_{jk \in S^{(i)}} y_{jk} c_{jk,m}^*(n) \quad m \in \mathbb{Z}, n \in I_N$$

$$\nu_m^{(i)}(n_1, n_2) = \sum_{jk \in S^{(i)}} d_{jk,m}(n_1, n_2) \quad m \in \mathbb{Z}, n_1, n_2 \in I_N\,, \tag{10}$$

where $d_{jk,m}(n_1, n_2) = \sum_{l \in \mathbb{Z}} c_{jk,l}(n_1) c_{jk,l-m}^*(n_2)$. We also define the following symbol index sets:

$$\begin{aligned} T^{(i)} &= \{n : m^{S^{(i-1)}} - B_q < n < n^{S^{(i)}} + B_q\} \\ U^{(i)} &= \{n : m^{S^{(i-1)}} - B_q < n \le m^{S^{(i)}} - B_q\}\,, \end{aligned} \tag{11}$$

where n^S (m^S) denotes the smallest (largest) symbol index localized outside (inside) the index set S. B_q is a constant which depends on T_0, the supports of $p(t)$ and $\psi(t)$, and on q which is the largest scale used in the wavelet decomposition. The algorithm is stated below, assuming that the wavelet coefficients parameters $\{w_{jk}(n)\}$ are updated first, then the data symbols $\{b_n\}$ in increasing order, and finally the gain a.

Case A: Update the objective of $\{w_{jk}(n)\}$. Compute recursively the following:

$$\begin{aligned} \alpha_m^{(i+1)} &= \alpha_m^{(i)} + \sum_{\substack{(n_1,n_2) \in V^{(i)} \\ |n_1 - n_2| < 2B_q}} b_{n_1}^{(i)} b_{n_2}^{(i)*} \nu_m^{(i)}(n_1, n_2) \\ \beta_m^{(i+1)} &= \beta_m^{(i)} + \sum_{n \in U^{(i)}} b_n^{(i)*} \mu_m^{(i)}(n)\,, \end{aligned} \tag{12}$$

where $V^{(i)} = U^{(i)} \times T^{(i)} \cup T^{(i)} \times U^{(i)}$. Next the objective is adjusted by including additional terms:

$$\begin{aligned} \gamma_m^{(i+1)} &= |a^{(i)}|^2 \Big[\alpha_m^{(i+1)} + \sum_{\substack{n_1,n_2 \in T^{(i+1)} \\ |n_1 - n_2| < 2B_q}} b_{n_1}^{(i)} b_{n_2}^{(i)*} \nu_m^{(i+1)}(n_1, n_2)\Big] \\ \delta_m^{(i+1)} &= a^{(i)*} \Big[\beta_m^{(i+1)} + \sum_{n \in T^{(i+1)}} b_n^{(i)*} \mu_m^{(i+1)}(n)\Big]. \end{aligned} \tag{13}$$

Finally the objective for z is given by:

$$\phi^{(i+1)}(z) = \frac{2}{N_0} \sum_{m \in \mathbb{Z}} (\delta_{-m}^{(i+1)} + \delta_m^{(i+1)*} - \gamma_m^{(i+1)}) z^m\,. \tag{14}$$

Case B: Update the objective of $\{b_n : n \in T^{(i+1)}\}$. Compute recursively the following for $n \in T^{(i+1)}$:

$$\epsilon_m^{(i+1)}(n) = \epsilon_m^{(i)}(n) + \sum_{\substack{n_1 \in U^{(i)} \\ |n - n_1| < 2B_q}} b_{n_1}^{(i)*} \nu_m^{(i)}(n, n_1)\,. \tag{15}$$

Next additional terms are included:

$$\begin{aligned} \varepsilon^{(i+1)}(n) = a^{(i)} \sum_{m \in \mathbb{Z}} z^{(i+1)m} \Big(\mu_m^{(i+1)*}(n) - a^{(i)*}[\epsilon_m^{(i+1)}(n) + \\ \sum_{\substack{n_1 \in T^{(i+1)} \\ n_1 < n}} b_{n_1}^{(i+1)*} \nu_m^{(i+1)}(n, n_1) + \sum_{\substack{n_1 \in T^{(i+1)} \\ n_1 > n}} b_{n_1}^{(i)*} \nu_m^{(i+1)}(n, n_1)]\Big) \end{aligned} \tag{16}$$

Notice that data symbols with index smaller than n have already assumed the updated value. Finally the objective for b_n is given by:

$$\phi^{(i+1)}(b_n) = \frac{2}{N_0}[b_n \varepsilon^{(i+1)}(n) + b_n^{-1} \varepsilon^{(i+1)*}(n)]. \tag{17}$$

Case C: Update the objective of a. Compute:

$$\begin{aligned} \zeta^{(i+1)} &= \frac{2}{N_0} \sum_{m \in \mathbb{Z}} (\beta_{-m}^{(i+1)} + \sum_{n \in T^{(i+1)}} b_n^{(i+1)*} \mu_{-m}^{(i+1)}(n)) z^{(i+1)m} \\ \eta^{(i+1)} &= \frac{2}{N_0} \sum_{m \in \mathbb{Z}} \Big(\alpha_m^{(i+1)} + \\ &\quad \sum_{\substack{n_1,n_2 \in T^{(i+1)} \\ |n_1 - n_2| < 2B_q}} b_{n_1}^{(i+1)} b_{n_2}^{(i+1)*} \nu_m^{(i+1)}(n_1, n_2)\Big) z^{(i+1)m}. \end{aligned} \tag{18}$$

The objective for a is given by:

$$\phi^{(i+1)}(a) = a^* \zeta^{(i+1)} + a \zeta^{(i+1)*} - |a|^2 \eta^{(i+1)}. \tag{19}$$

6. SIMULATIONS

The SAGE algorithm has been evaluated by means of a simulation program written in MATLAB. The results summarized below were obtained from 100 Monte-Carlo simulations at each SNR value. In each run the signal parameters were chosen randomly and independently of each other. The data bits $\{b_n\}$ were selected as either +1 or −1. The gain a was of a fixed amplitude and random phase uniformly distributed in $[-\pi, \pi]$, and the time delay d was uniformly distributed in $[-T_p/2, T_p/2]$. In each run a 7 chip PN code was phase modulated by a 63 bit long sequence. The transmitted signal was then passed through a band limited AWGN channel, $|f| \le 1/2T_c$, where T_c is the chip time. The data bits were differentially decoded, due to phase ambiguity. The algorithm was allowed to synchronize during the first

31 bits. The Daubechies wavelets [1] of length 4 have been used in the wavelet decomposition of the received signal.

Figure 1 shows the bit error probability. We observe that the simulation results closely match the theoretical bound for a DPSK decoder. The performance of a phase coherent DLL with comparable response time is also shown. The parameter $\delta = 2/W_L T$ is the normalized response time of the DLL, where W_L is the two-sided loop bandwidth, and T is the signaling period, so a value of $\delta = 31$ was used. The loop damping of the equivalent PLL was taken as $\zeta = 0.707$, and zero detuning was assumed. The DLL error probability was found by numerical integration [4]. We observe that the DLL performance has degraded by more than 1dB with respect to the ideal PSK error bound, while the SAGE performance is essentially optimal. Figure 2 shows the synchronization performance of the SAGE algorithm. The SAGE algorithm achieves a very small synchronization error at low SNR which explains its nearly optimal detection performance. The synchronization error of the DLL is seen to be much larger. The Cramer-Rao (CR) bound on time delay estimation is also shown for reference. Finally, Figure 3 compares the phase error variance of the SAGE algorithm with that of a Data Aided Loop (DAL). The CR bound on phase error variance is shown for reference. The DAL performance was calculated with the same value of δ as above, for the case of a suppressed carrier ($m = 0$, see [5]). The two systems have a similar performance, but it should be noted that the DAL relies on a perfectly synchronized reference.

7. CONCLUSIONS

In this paper we have introduced a new method for joint timing and phase synchronization and optimal ML detection of transmitted symbols. This method was derived by applying the principles of the SAGE algorithm. The use of the discrete wavelet and Fourier transforms and a choice of parameter update strategy is the key to its simple recursive form. The flexibility in choosing several system parameters such as the wavelet basis, the step size of the algorithm, and the penalty functions, makes it a useful solution for a variety of applications.

REFERENCES

[1] I. Daubechies, "Orthonormal bases of compactly supported wavelets", Comm. Pure Appl. Math., 41 (1988), no. 7, 909–996.

[2] A.P. Dempster, N.M. Laird, D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", J. Roy Stat. Soc. series B, 39(1):1-38, 1977.

[3] J.A. Fessler, A.O. Hero, "Space-alternating generalized expectation-maximization algorithm", IEEE Trans. ASSP, 42(10):2664-2677, Oct. 1994.

[4] W.C. Lindsey, M.K. Simon, *Telecommunication Systems Engineering*, Englewood Cliffs, NJ: Prentice-Hall, 1973.

[5] W.C. Lindsey, M.K. Simon, "Data-aided carrier tracking loops", IEEE Trans. COM, 19(2):157-168, Apr. 1971.

[6] H. Meyr, M. Oerder, A. Polydoros, "On sampling rate, analog prefiltering, and sufficient statistics for digital receivers", IEEE Trans. COM, 42:3208-14, Dec. 1994.

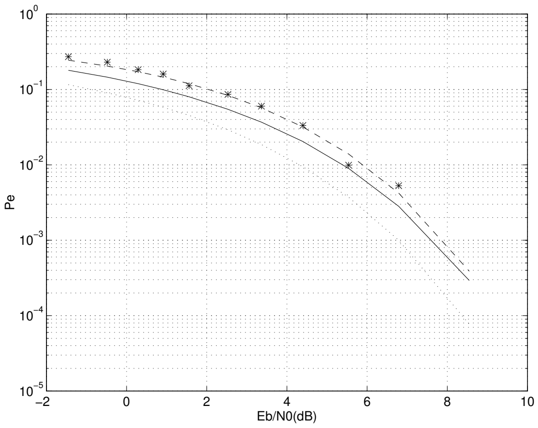

Figure 1. Bit error probability vs. SNR, stars: simulation results; solid: DLL performance; dashed: DPSK error bound; dotted: PSK error bound.

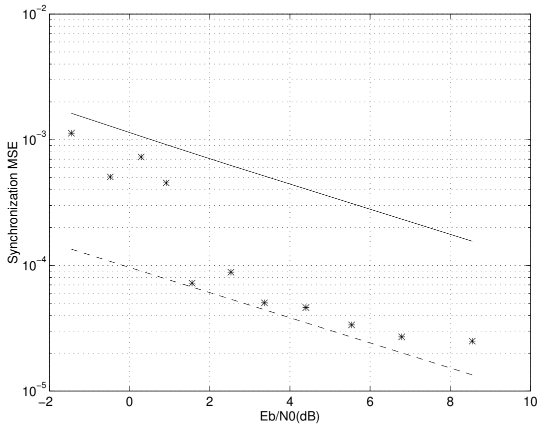

Figure 2. Synchronization mean squared error vs. SNR, stars: simulation results; solid: DLL performance; dashed: CR bound.

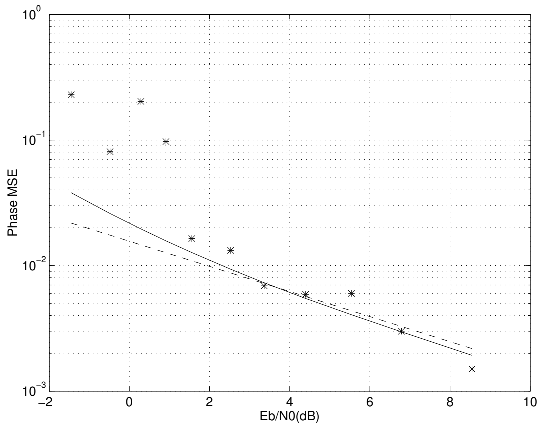

Figure 3. Phase mean squared error, stars: simulation results; solid: DAL performance; dashed: CR bound.