Low-complexity feedforward symbol timing estimator using Conditional Maximum Likelihood principle

Yik-Chung Wu and Erchin Serpedin*

Abstract

A low complexity feedforward symbol-timing estimator based on the Conditional Maximum Likelihood principle is proposed. An approximation is applied to the Fourier series expansion of the conditional maximum likelihood function such that implementation complexity is greatly reduced. It is shown that the proposed estimator can be viewed as a generalization of the well-known square nonlinearity estimator proposed by Oerder and Meyr in [8]. Simulation results show that the performance of the proposed estimator is very close to the conditional Cramer-Rao bound and is better than that of the square nonlinearity estimator.

Index Terms

Feedforward, Symbol Timing Estimator, Conditional Maximum Likelihood, Low Complexity

EDICS number : CL1.5.1, Synchronization Techniques

The authors are with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128, USA. Phone: (979) 458 2287, Fax: (979) 862 3128, email: {ycwu, serpedin}@ee.tamu.edu.

*The corresponding author is Dr. Erchin Serpedin.

I. Introduction

In [1], the symbol timing estimation problem is formulated in matrix form and solved using Conditional Maximum Likelihood (CML) principle. Specifically, the complex envelope of a received linear modulation is given by

$$r(t) = e^{j\theta_o} \sum_i d_i g(t - iT - \varepsilon_o T) + n(t), \tag{1}$$

where θ_o is the unknown phase offset; d_i is the complex valued symbol being sent; $g(t)$ is the transmit pulse shape with unit energy, which is assumed to be a root raised cosine filter with roll-off factor α; T is the symbol period; $\varepsilon_o \in [0, 1]$ is the unknown symbol timing delay to be estimated and $n(t)$ is the complex-valued circularly distributed white Gaussian noise with power density N_o. The received signal is filtered by anti-aliasing filter and then sampled at rate $1/T_s$, where $T_s \triangleq T/Q$. The received vector $\mathbf{r}$, which consists of $L_o Q$ consecutive received samples (where L_o is the observation length), can be expressed as (without loss of generality, we consider the received sequence start at $t = 0$)

$$\mathbf{r} = [r(0)\ r(T_s)\ \dots\ r((L_o Q - 1)T_s)]^T = \mathbf{A}_{\varepsilon_o} \mathbf{d}_o + \mathbf{n}, \tag{2}$$

where[1]

$$\begin{aligned}
\mathbf{A}_\varepsilon &\triangleq [\mathbf{a}_{-L}(\varepsilon)\ \mathbf{a}_{-L+1}(\varepsilon)\ \dots\ \mathbf{a}_{L_o+L-1}(\varepsilon)] && (3)\\
\mathbf{a}_i(\varepsilon) &\triangleq [g(-iT - \varepsilon T)\ g(T_s - iT - \varepsilon T) \\
&\qquad \dots\ g((L_o Q - 1)T_s - iT - \varepsilon T)]^T && (4)\\
\mathbf{d}_o &\triangleq e^{j\theta_o}[d_{-L}\ d_{-L+1}\ \dots\ d_{L_o+L-1}]^T && (5)\\
\mathbf{n} &\triangleq [n(0)\ n(T_s)\ \dots\ n((L_o Q - 1)T_s)]^T. && (6)
\end{aligned}$$

In the above expressions, L is the number of symbols affected by the Inter-Symbol Interference (ISI) introduced by one side of $g(t)$.

The joint maximum likelihood estimate of ε_o and $\mathbf{d}_o$ is given by maximizing (ε and $\mathbf{d}$ are the trial values for the timing delay and the data vector respectively)

$$p(\mathbf{r}|\varepsilon, \mathbf{d}) = \frac{1}{(\pi N_o)^{L_o Q}} \exp\left[-\frac{(\mathbf{r} - \mathbf{A}_\varepsilon \mathbf{d})^H (\mathbf{r} - \mathbf{A}_\varepsilon \mathbf{d})}{N_o} \right], \tag{7}$$

[1]Notation $\mathbf{x}^T$ denotes the transpose of $\mathbf{x}$ and $\mathbf{x}^H$ denotes the transpose conjugate of $\mathbf{x}$.

or equivalently minimizing

$$J(\mathbf{r}|\varepsilon,\mathbf{d}) = (\mathbf{r}-\mathbf{A}_\varepsilon\mathbf{d})^H(\mathbf{r}-\mathbf{A}_\varepsilon\mathbf{d}). \tag{8}$$

In the CML approach, the nuisance parameters $\mathbf{d}$ are modelled as deterministic and estimated from the received vector $\mathbf{r}$. From the linear signal model given in (2), the minimum variance unbiased estimate for $\mathbf{d}$ (when ε is fixed) is [2]

$$\hat{\mathbf{d}} = (\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon)^{-1}\mathbf{A}_\varepsilon^H\mathbf{r}. \tag{9}$$

Plugging (9) into (8), after some straightforward manipulations and dropping the terms irrelevant to optimization, the timing delay is estimated by maximizing the following CML function [1]

$$\Lambda(\varepsilon) = \mathbf{r}^H\mathbf{A}_\varepsilon(\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon)^{-1}\mathbf{A}_\varepsilon^H\mathbf{r}. \tag{10}$$

In general, the maximum of the CML function can be found by plugging different values of ε into (10). The value that provides the maximum value of $\Lambda(\varepsilon)$ is the CML estimate. Since ε is a continuous variable, this exhaustive search method requires a lot of computation and is impractical. Alternatively, a timing error detector (TED) [1] can be used in a feedback configuration. However, in burst mode transmission, a feedforward timing delay estimator [3]-[5], [8] is preferred since feedback synchronizers usually require a relatively long acquisition time. In this letter, a new method for optimizing (10) is proposed so that a low complexity feedforward symbol-timing estimator results.

II. Proposed Estimator

Suppose we calculated K uniformly spaced values of $\Lambda(\varepsilon)$ using (10) such that a sequence $\Lambda(k) \triangleq \Lambda(k/K)$ for $k = 0, 1, ..., K$-1 is obtained (without loss of generality, we only consider K is even). Let construct a periodic sequence $\tilde{\Lambda}(m)$ by periodically extending $\Lambda(k)$. Further, denote $\tilde{\Lambda}(\varepsilon)$ as the continuous and periodic function with its samples given by $\tilde{\Lambda}(m)$. According to sampling theorem, as long as the sampling frequency K is higher than twice the highest frequency of $\tilde{\Lambda}(\varepsilon)$, then $\tilde{\Lambda}(\varepsilon)$ can be represented by its samples $\tilde{\Lambda}(m)$ without loss of information. Being a periodic function, $\tilde{\Lambda}(\varepsilon)$ can be represented in terms of its Fourier series expansion:

$$\tilde{\Lambda}(\varepsilon) = \sum_{p=-\infty}^{\infty} A_p e^{j2\pi p\varepsilon}, \tag{11}$$

where

$$\begin{aligned} A_p &= \int_0^1 \tilde{\Lambda}(\varepsilon) e^{-j2\pi p\varepsilon} d\varepsilon \ , \\ &= \begin{cases} \frac{1}{K}\sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi pk/K}, & p = -\frac{K}{2}, \ldots, \frac{K}{2} \\ 0 & \text{otherwise} \ , \end{cases} \end{aligned} \tag{12}$$

and (12) follows from the standard relationship between the discrete Fourier transform (DFT) and the Fourier series expansion of a periodic sequence [9]. From (11), it can be seen that once the coefficients A_p are determined, $\tilde{\Lambda}(\varepsilon)$ can be calculated for any $\varepsilon \in [0, 1]$. Then the problem of maximizing (10) can now be replaced by maximizing (11). For efficient implementation, $\tilde{\Lambda}(\varepsilon)$ for $0 \leq \varepsilon \leq 1$ can be approximated by a K'-point sequence, denoted as $\Lambda(k')$ for $0 \leq k' \leq K' - 1$, by zero padding the high frequencies coefficients of A_p and performing a K'-point inverse DFT. For sufficiently large value of K', $\Lambda(k')$ becomes very close to $\tilde{\Lambda}(\varepsilon)$ for $0 \leq \varepsilon \leq 1$, and the index with the maximum amplitude can be viewed as an estimate of the unknown timing parameter ε_o.

To avoid the complexity in performing the K'-point IDFT, an approximation is applied to (11). More precisely, extensive simulations show that it is sufficient to approximate (11) as follows

$$\tilde{\Lambda}(\varepsilon) \approx A_0 + 2\mathfrak{Re}\{A_1 e^{j2\pi\varepsilon}\} \qquad \text{for } 0 \leq \varepsilon \leq 1, \tag{13}$$

where $\mathfrak{Re}\{x\}$ stands for real part of x. In order to maximize $\widetilde{\Lambda}(\varepsilon)$, we notice

$$\arg(A_1) = -2\pi\varepsilon, \tag{14}$$

where $\arg(x)$ denotes the phase of x. Or equivalently,

$$\hat{\varepsilon} = -\frac{1}{2\pi} \arg\{\sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi k/K}\}. \tag{15}$$

The estimated delay $\hat{\varepsilon}$ is the normalized time difference (with respect to T) between the first sample of the received vector $\mathbf{r}$ and the nearest optimum sampling instant. The calculation within the $\arg(.)$ operation is actually the 2^{nd} output of a K-point DFT of the sequence $\Lambda(k)$ (or the Fourier coefficient at symbol rate $f = 1/T$).

The proposed estimator in (15) only involves the calculation of K samples of the CML function using (10), a K-point DFT, and an $\arg(.)$ operation. From the simulation results to be presented, it is found that K = 4 is sufficient to yield good estimates in practical applications.[2] Therefore, the 4-point DFT in (15) can be computed easily without any multiplications. The main complexity comes from the calculation of the 4 samples of $\Lambda(\varepsilon)$ using (10). However, notice that the matrix $\mathbf{A}_\varepsilon(\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon)^{-1}\mathbf{A}_\varepsilon^H$ can be pre-computed for $\varepsilon = k/4$ with $0 \leq k \leq 3$. This greatly reduces the arithmetic complexity of implementation. Complexity can be further reduced by approximating the pre-computed $\mathbf{A}_\varepsilon(\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon)^{-1}\mathbf{A}_\varepsilon^H$ using Sum-of-Power-of-Two (SOPOT) expressions [6], [7].

III. Relationship with the Square nonlinearity estimator

In this section, we will show that the proposed estimator in (15) can be viewed as a generalization of the square nonlinearity estimator [8]. First consider the $(i,j)^{th}$ element of $\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon$ $(i,j = -L, -L+1, ..., L_o+L-1)$,

$$\begin{aligned}[\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon]_{ij} &= \sum_{n=0}^{L_oQ-1} g(nT_s - iT - \varepsilon T)g(nT_s - jT - \varepsilon T)\\ &= \sum_{n=-\infty}^{\infty} g(nT_s - iT - \varepsilon T)g(nT_s - jT - \varepsilon T) - z_\varepsilon(i,j)\\ &= \delta_{ij} - z_\varepsilon(i,j),\end{aligned}$$

where $z_\varepsilon(i,j) \triangleq \sum_{n<0, n\geq L_o} g(nT_s - iT - \varepsilon T)g(nT_s - jT - \varepsilon T)$ and $\delta_{ij} = 1$ if $i = j$ and zero otherwise. Therefore, with the matrix inversion lemma, $(\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon)^{-1} = (\mathbf{I}-\mathbf{Z}_\varepsilon)^{-1} = \mathbf{I}-(\mathbf{I}-\mathbf{Z}_\varepsilon^{-1})^{-1}$ where $\mathbf{I}$ is the identity matrix and $[\mathbf{Z}_\varepsilon]_{ij} \triangleq z_\varepsilon(i,j)$. Plugging this result into (10), it follows that

$$\Lambda(\varepsilon) = \|\mathbf{A}_\varepsilon^H\mathbf{r}\|^2 - \mathbf{r}^H\mathbf{A}_\varepsilon(\mathbf{I}-\mathbf{Z}_\varepsilon^{-1})^{-1}\mathbf{A}_\varepsilon^H\mathbf{r}. \tag{16}$$

Now consider the i^{th} element of $\mathbf{A}_\varepsilon^H\mathbf{r}$ $(i = -L, -L+1, ..., L_o+L-1)$,

$$\begin{aligned}[\mathbf{A}_\varepsilon^H\mathbf{r}]_i &= \sum_{n=0}^{L_oQ-1} g(nT_s - iT - \varepsilon T)r(nT_s)\\ &= \sum_{n=-\infty}^{\infty} g((i+\varepsilon)T - nT_s)r(nT_s).\end{aligned} \tag{17}$$

The last equality holds if $r(nT_s)$ is considered to be an infinite length sequence with its values equal to zero for $n < 0$ and $n \geq L_o$. It is recognized that the summation in (17) is just the

[2]Note that the miniimum K for the proposed estimator in (15) to work is $K = 3$, but $K = 4$ is preferred since 4-point DFT is more easily computed.

filtering of $\tilde{r}(t) \triangleq r(t)w(t)$, where $w(t)$ is a rectangular window with length L_oT, by $g(t)$ and then sampled at $t = (i + \varepsilon)T$. If we define $x(t) \triangleq g(t) \otimes \tilde{r}(t)$ where $\otimes$ denotes convolution, we have $[\mathbf{A}_\varepsilon^H \mathbf{r}]_i = x((i + \varepsilon)T)$. Plugging this result into (16) gives

$$\Lambda(k) = \sum_{i=0}^{L_o-1} |x(iT + kT/K)|^2 + \Gamma(k), \tag{18}$$

with $\Gamma(k) \triangleq \sum_{i=-L}^{-1} |x(iT+kT/K)|^2+\sum_{i=L_o}^{L_o+L-1} |x(iT+kT/K)|^2-\mathbf{r}^H\mathbf{A}_\varepsilon(\mathbf{I}-\mathbf{Z}_\varepsilon^{-1})^{-1}\mathbf{A}_\varepsilon^H\mathbf{r}|_{\varepsilon=k/K}$. Therefore, the proposed CML feedforward timing delay estimator in (15) can be rewritten as

$$\hat{\varepsilon} = -\frac{1}{2\pi} \arg\{ \sum_{l=0}^{KL_o-1} |x(lT/K)|^2 e^{-j2\pi l/K} + \sum_{k=0}^{K-1} \Gamma(k) e^{-j2\pi k/K} \}. \tag{19}$$

If the second term in the $arg(.)$ operation is ignored and K=4, we have the well known squaring algorithm [8]. Note that only Q=2 is required in the CML estimator while the original squaring algorithm in [8] requires $Q \geq 3$. Although the squaring estimator might be extended to work at the sampling rate $Q = 2$ by computing intermediate samples through interpolation before symbol timing estimation, the simulation results presented in the next section show that the proposed estimator outperforms the squaring algorithm.

IV. Simulation results and discussions

Figure 1 shows the Mean Square Error (MSE) against E_s/N_o for the proposed algorithm in (15) and the IDFT interpolation method with various K'. The conditional Cramer-Rao bound (CRB) [1] is also shown as a reference. In the simulations, the modulation is QPSK, α=0.5, Q=2, K=4, L_o=100 and L=3. ε_o is uniformly distributed in the range $[0, 1]$. θ_o is generated as uniformly distributed random variable in the range $[-\pi, \pi]$ and is constant in each estimation. Each point is obtained by averaging 10^4 simulation runs. It can be seen that the proposed algorithm (15) has a performance similar to that of IDFT interpolation method with K'=2048. This justifies the approximation in (13). Furthermore, the performance of the algorithm (15) is very close to the conditional CRB. This means that the estimator (15) almost reaches the ultimate performance of the CML principle.

The performances of the estimator (15) and that of the square nonlinearity estimator are compared in Figure 2 for $L_o = 100$. The parameters used for the proposed estimator (15) are the same as that in Figure 1. Note that $Q = 4$ is assumed for the square nonlinearity estimator. It is apparent that while both estimators have similar MSE performances at low E_s/N_o, the estimator (15) has smaller MSEs at high E_s/N_o. It is the additional term in (19) that improve the performance of the CML algorithm with respect to that of square nonlinearity estimator.

V. Conclusions

A new and low complexity feedforward symbol-timing estimator based on the conditional maximum likelihood principle is proposed. Simulation results show that the performance is good and very close to the conditional CRB. Furthermore, the proposed estimator (15) can be viewed as a generalization of the well-known square nonlinearity estimator, and it is shown that the estimator (15) performs better.

References

[1] J. Riba et al., "Conditional Maximum Likelihood Timing Recovery: Estimators and Bounds," *IEEE Tran. Signal Processing*, Vol. 49, pp. 835-850, Apr., 2001.

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing - Estimation Theory*. Prentice Hall, 1993.

[3] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat-fading channels: A cyclostationary approach." *IEEE Trans. Commun.*, vol. 46, pp. 400-411, Mar. 1998.

[4] K. E. Scott and E. B. Olasz, "Simultaneous clock phase and frequency offset estimation," *IEEE Trans. Commun.*, vol. 43, pp. 2263-2270, July 1995.

[5] M. Morelli et al., "Feedforward ML-based timing estimation with PSK signals," *IEEE Commun. Letters*, vol. 1, pp. 80-82, May 1997.

[6] H. Samueli, "An improved search algorithm for the design of multiplierless FIR filters with powers-of-two coefficients," *IEEE Trans. Circuits Syst. - II*, vol. 36, pp. 1044-1047, Jul. 1989.

[7] C.-K.-S. Pun et al., "On the design and efficient implementation of the Farrow structure," *IEEE Signal Processing Letters*, vol. 10, no. 7, pp. 189-192, Jul. 2003.

[8] M. Oerder and H. Meyr, "Digital filter and squaring timing recovery," *IEEE Trans. Commun.*, vol. 36, pp. 605-612, May 1988.

[9] A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing*, Prentice-Hall, 1989.

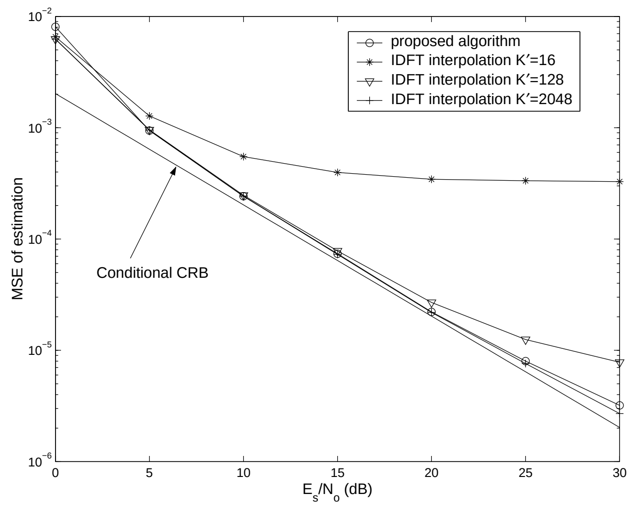

Fig. 1. MSE for the proposed algorithm and the IDFT interpolation method with K'=16, 128, 2048 for QPSK, Q=2, K=4, α=0.5, L_o=100 and L=3.

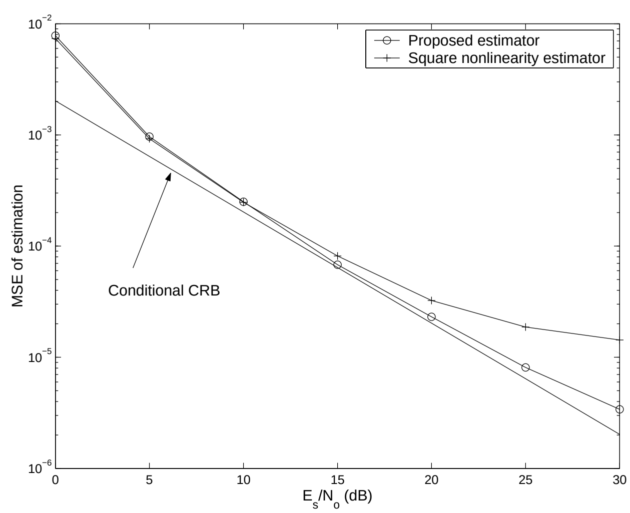

Fig. 2. Comparison of the MSE for the proposed algorithm and the square nonlinearity estimator with QPSK and $L_o = 100$.