

Design and Analysis of Feedforward Symbol Timing Estimators Based on the Conditional Maximum Likelihood Principle

Yik-Chung Wu and Erchin Serpedin*

Abstract

This paper presents a general feedforward symbol-timing estimation framework based on the conditional maximum likelihood principle. The proposed timing estimator presents reduced implementation complexity and is obtained by performing an approximation on the Fourier series expansion of the conditional maximum likelihood function. The proposed algorithm is applied to linear modulations and two commonly used continuous phase modulations: MSK and GMSK. For the linear modulations, it is shown both analytically and via simulations that the performance of the proposed estimator is very close to the conditional CRB and modified CRB for signal-to-noise ratios (SNR's) in the range $\text{SNR} \leq 30\text{dB}$. Furthermore, the proposed estimator is shown to be asymptotically equivalent to the classic square-law nonlinearity estimator [3] under certain conditions. In the case of MSK and GMSK modulations, although the proposed algorithm reaches the conditional CRB at certain SNRs, however, the conditional CRB is quite far away from the modified CRB, and there exists an alternative algorithm whose performance comes closer to the modified CRB. Therefore, the proposed estimator is more suitable for linear modulations rather than for MSK and GMSK modulations.

EDICS 3-SYNC: Synchronization and Timing Recovery

Index Terms

Feedforward, Symbol Timing Estimation, Conditional Maximum Likelihood, Linear modulations, MSK, GMSK, Cramer-Rao bound

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. This work was supported by the NSF Award No. CCR-0092901 and the Croucher Foundation.

*The corresponding author is Dr. Erchin Serpedin.

I. INTRODUCTION

In digital receivers, symbol timing synchronization can be implemented either in a feedforward or feedback mode. Although feedback schemes exhibit good tracking performances, they require a relatively long acquisition time. Therefore, for burst-mode transmissions, feedforward timing recovery schemes are more suitable. An all-digital feedforward symbol timing recovery scheme consists of first estimating the timing delay from the received samples, which is the focus of this paper, and then adjusting the timing using some sort of interpolation [1], [2].

Due to bandwidth efficiency considerations, non-data aided or blind symbol timing estimation schemes have attracted much attention during the last decade. Most of the feedforward timing estimators proposed in the literature exploit the cyclostationarity induced by oversampling the received signal [3]-[8]. In [3], Oerder and Meyr proposed the well-known square nonlinearity estimator. Several extensions of this square nonlinearity estimator can be found in [5]-[7]. In [8], a low-SNR approximation was applied to the maximum likelihood function in order to derive a logarithmic nonlinearity. Reference [4] reported for the first time a detailed performance analysis of the estimators based on various types of nonlinearities.

Recently, the conditional maximum likelihood (CML) principle was introduced for designing digital timing delay synchronizers by Riba, Sala and Vazquez [9], [10]. The CML solution is especially important for symbol timing synchronization because it yields self-noise free timing estimates at medium and high SNRs. However, [9], [10] concentrate on deriving a CML timing error detector (TED) so that the timing delay can only be tracked using a feedback loop. A first example of a feedforward symbol timing estimator based on the CML principle was reported in [11]. However, no insightful analysis pertaining to the design principles and performances has been reported. The purpose of this paper is to fill in this gap in the literature. The main design and performance characteristics of CML-based feedforward symbol timing delay estimators are established for general linear modulations and two commonly used continuous phase modulations, namely, minimum shift keying (MSK) and Gaussian MSK (GMSK) [12], [13]. The performance of the timing estimators is analyzed analytically and through simulations, and compared with the conditional Cramer-Rao bound (CCRB) [9], [10], the modified Cramer-Rao bound (MCRB) [14] and other existing state-of-the-art feedforward timing delay estimators [3], [14], [16], [17], and [19].

In the proposed algorithm, an approximation is applied to the Fourier series expansion of the CML function so that the complexity of the proposed estimator is greatly reduced. Although the resulting estimator is not completely self-noise free (due to the approximation), the performances of the proposed estimator (for both linear and nonlinear modulations) are in general very close to the CCRB for signal-to-noise ratios (SNR's) smaller than 30 dB. For higher SNRs, mean square error (MSE) floor occurs, but notice that at that high SNRs, the estimation MSE achieved by the proposed estimator is already very small, so the effect of MSE floors becomes relatively less critical.

For linear modulations, it is shown that the proposed estimator is asymptotically equivalent to the well-known square nonlinearity estimator [3]. However, the proposed estimator exhibits better performance (less self-noise/jitter) than [3] when a reduced number of data samples are available. Furthermore, it is shown that the performances of the proposed estimator for linear modulations are also very close to MCRB for $\text{SNR} \leq 30\text{dB}$. For MSK and GMSK modulations, although the performances of the proposed estimator come very close to the CCRB at certain SNR ranges, however, the CCRB is quite far away from the MCRB, and there exists an alternative algorithm whose performance comes closer to the MCRB. Therefore, it is concluded that the proposed estimator is more suitable for linear modulations than MSK and GMSK modulations.

The rest of the paper is organized as follows. The signal model and the CML function are first described in Section II. The proposed feedforward timing estimator is derived in Section III. The relationship between the proposed estimator and the well-known square nonlinearity estimator [3] is addressed in Section IV. The MSE expressions are derived in Section V. Simulation results and discussions are then presented in Section VI, and finally conclusions are drawn in Section VII.

II. SIGNAL MODEL AND THE CML FUNCTION

A. Signal Model

The complex envelope of a received linear modulation is given by

$$r(t) = e^{j\theta_o} \sqrt{\frac{E_s}{T}} \sum_i d_i g(t - iT - \varepsilon_o T) + \eta(t), \quad (1)$$

where θ_o is the unknown phase offset; E_s is the symbol energy; d_i stands for the zero-mean unit variance, independently and identically distributed (i.i.d.) complex valued symbols being sent;

$g(t)$ is the transmit pulse with unit energy; T is the symbol period; $\varepsilon_o \in [0, 1)$ is the unknown symbol timing delay to be estimated and $\eta(t)$ is the complex-valued circularly distributed white Gaussian noise with power density N_o . After passing through the anti-aliasing filter, the received signal is then sampled at the rate $1/T_s$, where $T_s \triangleq T/Q$. Note that the oversampling factor Q is determined by the frequency span of $g(t)$; if $g(t)$ is bandlimited to $f = \pm 1/T$ (an example of which is the square-root raised cosine pulse), $Q = 2$ is sufficient. 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 $t = 0$)

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

where¹

$$\mathbf{A}_{\varepsilon} \triangleq [\mathbf{a}_{-L_g}(\varepsilon), \mathbf{a}_{-L_g+1}(\varepsilon), \dots, \mathbf{a}_{L_o+L_g-1}(\varepsilon)], \quad (3)$$

$$\mathbf{a}_i(\varepsilon) \triangleq [g(-iT - \varepsilon T), g(T_s - iT - \varepsilon T), \dots, g((L_o Q - 1)T_s - iT - \varepsilon T)]^T, \quad (4)$$

$$\mathbf{d}_o \triangleq e^{j\theta_o} \sqrt{\frac{E_s}{T}} [d_{-L_g}, d_{-L_g+1}, \dots, d_{L_o+L_g-1}]^T, \quad (5)$$

$$\boldsymbol{\eta} \triangleq [\eta(0), \eta(1), \dots, \eta(L_o Q - 1)]^T, \quad (6)$$

$\eta(i) \triangleq \eta(iT/Q)$, and L_g denotes the number of symbols affected by the inter-symbol interference (ISI) introduced by one side of $g(t)$.

For MSK and GMSK modulations, the complex envelope of the received signal is given by

$$r(t) = \sqrt{\frac{E_s}{T}} \exp[j\pi \sum_n a_n q(t - nT - \varepsilon_o T) + j\theta_o] + \eta(t), \quad (7)$$

where a_n stands for the i.i.d. binary transmitted symbols, and $q(t)$ is the phase response of the modulator with length L and satisfies

$$q(t) = \begin{cases} 0 & t \leq 0 \\ 1/2 & t \geq LT. \end{cases} \quad (8)$$

The derivative of $q(t)$ is referred to as the frequency response of the modulator, and takes the form of a rectangular pulse or a convolution between a rectangular pulse and a Gaussian shaped pulse for MSK and GMSK modulations, respectively. According to the Laurent's expansion

¹Notation \mathbf{x}^T denotes the transpose of \mathbf{x} , and \mathbf{x}^H stands for the transpose conjugate of \mathbf{x} .

(LE) [18] and the fact that most of the energy of the GMSK modulation is concentrated in the first component of the expansion [18]-[20] (the MSK signal has only one component in the expansion), MSK and GMSK received signals can be approximated by

$$r(t) \cong e^{j\theta_o} \sqrt{\frac{E_s}{T}} \sum_i \bar{d}_i \bar{g}(t - iT - \varepsilon_o T) + \eta(t), \quad (9)$$

where

$$\bar{d}_i \triangleq \exp[j\frac{\pi}{2} \sum_{n=1}^i a_n], \quad (10)$$

$$\bar{g}(t) \triangleq \prod_{n=0}^{L-1} p(t + nT), \quad (11)$$

and

$$p(t) \triangleq \begin{cases} \sin[\pi q(t)] & 0 \leq t \leq LT \\ p(2LT - t) & LT < t \leq 2LT \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

Therefore, the sampled MSK and GMSK modulations can also be expressed in a form similar to (2). Since the pseudo-symbols (or equivalent data) \bar{d}_i are zero mean, unit variance, a single system model is sufficient to treat the linear modulations, MSK and GMSK signals within a common framework.

Remark 1. Notice that another formulation for the GMSK signal is to express the signal using all the 2^{L-1} terms of the LE as is done in [10]. However, there is a disadvantage for doing this: including more LE terms in the formulation would significantly increase the number of pseudo-symbols. Since in CML method, the pseudo-symbols and the unknown timing delay are jointly estimated from an observation vector of certain length, increasing the number of pseudo-symbols to be estimated would definitely degrade the overall estimation accuracy (of both pseudo-symbols and timing delay). Of course, neglecting some small LE terms (as is done in this paper) would introduce interference and degrade the performance for the resulting estimator, but from the simulation examples to be presented in Section VI, the effect of the system model approximation (9) is very small and occurs only at a very high SNR region (at $\text{SNR} \geq 50\text{dB}$).

Remark 2. MSK and GMSK modulations belong to a broader class of modulation, called MSK-type modulation [12], [13]. The system model, the subsequent proposed estimator and the

MSE analysis can also be applied to other members of this MSK-type modulation as long as the approximation in (9) is tight (e.g., 1RC, 2RC modulations). However, in this paper, we only concentrate on two commonly used members: MSK and GMSK.

B. The CML Function

From (2), the joint maximum likelihood estimate of ε_o and \mathbf{d}_o is given by maximizing

$$p(\mathbf{r}|\varepsilon, \mathbf{d}) = \frac{1}{(\pi N_o)^{L_o Q}} \exp \left[-\frac{(\mathbf{r} - \mathbf{A}_\varepsilon \mathbf{d})^H (\mathbf{r} - \mathbf{A}_\varepsilon \mathbf{d})}{N_o} \right], \quad (13)$$

or equivalently minimizing

$$J(\mathbf{r}|\varepsilon, \mathbf{d}) = (\mathbf{r} - \mathbf{A}_\varepsilon \mathbf{d})^H (\mathbf{r} - \mathbf{A}_\varepsilon \mathbf{d}), \quad (14)$$

where ε and \mathbf{d} are the trial values for ε_o and \mathbf{d}_o , respectively.

In the CML approach, the nuisance parameters \mathbf{d}_o are modeled as deterministic and estimated from the received vector \mathbf{r} . From the linear signal model given in (2), if no constraint is imposed on the possible value of \mathbf{d}_o , the maximum likelihood estimate for \mathbf{d}_o (when ε is fixed) is [15]

$$\hat{\mathbf{d}} = (\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon)^{-1} \mathbf{A}_\varepsilon^H \mathbf{r}. \quad (15)$$

Plugging (15) into (14), after some straightforward manipulations and dropping the irrelevant terms, the timing delay is estimated by maximizing the following CML function [9]

$$\Lambda(\varepsilon) = \mathbf{r}^H \mathbf{A}_\varepsilon (\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon)^{-1} \mathbf{A}_\varepsilon^H \mathbf{r}. \quad (16)$$

In general, the maximum of the CML function can be found by plugging different values of ε into (16). 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 computations and is impractical. Alternatively, a timing error detector (TED) [9] can be used in a feedback configuration. However, in burst mode transmissions, feedforward timing delay estimators [3]-[8] are preferred since they avoid the relatively long acquisition time and hang-up problem in feedback schemes. In the following, a new method for optimizing (16) is proposed so that an efficient implementation of the feedforward symbol-timing estimator results.

III. PROPOSED ESTIMATOR

Figure 1 shows some realizations of the CML function calculated using (16), where the true timing delay is $\varepsilon_o = 0.25$ (for the linear modulation, $g(t)$ is a square-root raised cosine filter with roll-off factor 0.5). It can be seen that the CML function has only one maximum. Since the CML function is *smooth*, we expect that it is not necessary to calculate the CML function for all the values of ε . It is possible that the CML function is first calculated for some ε 's and the values in between can be found by interpolation.

More specifically, suppose we calculated K uniformly spaced values of $\Lambda(\varepsilon)$ using (16) such that a sequence $\Lambda(k) \triangleq \Lambda(k/K)$ for $k = 0, 1, \dots, K-1$ is obtained (without loss of generality, we consider K is even). Let construct a periodic sequence $\tilde{\Lambda}(m)$ by periodically extending $\Lambda(k)$. Further, denotes $\tilde{\Lambda}(\varepsilon)$ as the continuous and periodic function with its samples given by $\tilde{\Lambda}(m)$. According to the sampling theorem, as long as the sampling frequency $1/K$ 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. The relationship between $\tilde{\Lambda}(\varepsilon)$ and $\tilde{\Lambda}(m)$ is then given by

$$\tilde{\Lambda}(\varepsilon) = \sum_{m=-\infty}^{\infty} \tilde{\Lambda}(m) \text{sinc}(\pi K(\varepsilon - m/K)). \quad (17)$$

where $\text{sinc}(x) \triangleq \sin(x)/x$. Now, expand $\tilde{\Lambda}(\varepsilon)$ into a Fourier series

$$\tilde{\Lambda}(\varepsilon) = \sum_{p=-\infty}^{\infty} A_p e^{j2\pi p\varepsilon}, \quad (18)$$

where

$$A_p = \int_0^1 \tilde{\Lambda}(\varepsilon) e^{-j2\pi p\varepsilon} d\varepsilon. \quad (19)$$

Substituting (17) into (19) yields

$$\begin{aligned} A_p &= \sum_{m=-\infty}^{\infty} \tilde{\Lambda}(m) \int_0^1 \text{sinc}(\pi K(\varepsilon - m/K)) e^{-j2\pi p\varepsilon} d\varepsilon \\ &= \sum_{k=0}^{K-1} \Lambda(k) \sum_{l=-\infty}^{\infty} \int_0^1 \text{sinc}(\pi K(\varepsilon - l - k/K)) e^{-j2\pi p\varepsilon} d\varepsilon \\ &= \sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi pk/K} \int_{-\infty}^{\infty} \text{sinc}(\pi K\varepsilon) e^{-j2\pi p\varepsilon} d\varepsilon \\ &= \sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi pk/K} \cdot \frac{1}{K} \mathfrak{F}\{\text{sinc}(\pi\varepsilon)\}_{f=p/K}, \end{aligned} \quad (20)$$

where $\mathfrak{F}\{\}$ denote the Fourier transform. It is clear that,

$$A_p = \begin{cases} \frac{1}{K} \sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi pk/K}, & p = -\frac{K}{2}, \dots, \frac{K}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

From (18), 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 (16) can now be replaced by maximizing (18). For efficient implementation, the function $\tilde{\Lambda}(\varepsilon)$ for $0 \leq \varepsilon < 1$ can be approximated by a K' -point sequence ($K' > K$) as follow

$$\Lambda(k') = \sum_{p=-K'/2}^{K'/2-1} A_p e^{j2\pi pk'/K'} \quad \text{for } k' = 0, 1, \dots, K' - 1. \quad (22)$$

This is equivalent to first calculating A_p using (21), then zero padding the high frequency coefficients (A_p) and finally performing a K' -point inverse discrete Fourier transform (IDFT). For sufficiently large value of K' , $\Lambda(k')$ becomes very close to $\tilde{\Lambda}(\varepsilon)$ for $0 \leq \varepsilon < 1$, and the index with the maximum amplitude can be viewed as an estimate of the unknown timing parameter ε_o . Fig. 2a shows the block diagram for this algorithm when $K=4$. For the rest of the paper, we refer to this estimator as the IDFT-based CML estimator.

To avoid the complexity in performing the K' -point IDFT, an approximation is applied to (18). More precisely, it can be seen from Fig. 1 that the CML function for symbol timing estimation resembles a sine function with one period in the interval $0 \leq \varepsilon < 1$. It is expected that the Fourier coefficient A_1 is much larger than the Fourier coefficients associated with higher frequencies. Therefore, it is sufficient to approximate (18) as follows

$$\tilde{\Lambda}(\varepsilon) \approx A_0 + 2\Re\{A_1 e^{j2\pi\varepsilon}\} \quad \text{for } 0 \leq \varepsilon < 1, \quad (23)$$

where $\Re\{x\}$ stands for real part of x . In order to maximize $\tilde{\Lambda}(\varepsilon)$, the following equation must hold

$$\arg(A_1) = -2\pi\varepsilon, \quad (24)$$

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

$$\hat{\varepsilon} = -\frac{1}{2\pi} \arg\left\{ \sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi k/K} \right\}. \quad (25)$$

The estimated delay $\hat{\varepsilon}$ is the normalized (with respect to T) time difference between the first sample of the received vector \mathbf{r} and the nearest optimum sampling instant. The calculation within the $\arg(\cdot)$ operation is actually the first bin (i.e., 2^{nd} output) of a K -point discrete Fourier transform (DFT) of the sequence $\Lambda(k)$ (or the Fourier coefficient at symbol rate $f = 1/T$). Based on (24), it is not difficult to check that the proposed estimator (25) is asymptotically unbiased, a result which is independent of the approximation used in (23).

From a computational viewpoint, it is worth mentioning that the proposed estimator only involves the calculation of K samples of the CML function using (16), a K -point DFT, and an $\arg(\cdot)$ operation. From the results to be presented, it is found that $K = 4$ is sufficient to yield good estimates in practical applications. Therefore, the 4-point DFT in (25) can be computed easily without requiring any multiplications. The main complexity comes from the calculation of the 4 samples of $\Lambda(\varepsilon)$ using (16). 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 [21], [22].

IV. RELATIONSHIP WITH THE SQUARE NONLINEARITY ESTIMATOR

In this section, we will show that, if $g(t)$ is a square-root raised cosine pulse, the proposed estimator in (25) asymptotically reduces to the well-known square nonlinearity estimator [3]. First notice that when $g(t)$ is a square-root raised cosine pulse and in the asymptotic case (as $L_o \rightarrow \infty$), $[\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon]_{ij} \approx \delta_{ij}$ [10], where $\delta_{ij} = 1$ if $i = j$ and zero otherwise. Notice that the matrix $\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon$ is of dimension $(L_o + 2L_g) \times (L_o + 2L_g)$. The approximation $[\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon]_{ij} \approx \delta_{ij}$ holds very well for the central portion (of dimension $L_o \times L_o$) of $\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon$. For the boundary of $\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon$, the values are smaller than 1. As $L_o \rightarrow \infty$, the boundary of $\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon$ becomes insignificant and can be ignored. Putting $[\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon]_{ij} \approx \delta_{ij}$ into (16), it follows that

$$\Lambda(\varepsilon) \approx \|\mathbf{A}_\varepsilon^H \mathbf{r}\|^2. \quad (26)$$

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

$$\begin{aligned} [\mathbf{A}_\varepsilon^H \mathbf{r}]_i &= \sum_{n=0}^{L_o Q - 1} g(nT_s - iT - \varepsilon T) r(nT_s) \\ &= \sum_{n=-\infty}^{\infty} g((i + \varepsilon)T - nT_s) \tilde{r}(nT_s) \end{aligned} \quad (27)$$

where $\tilde{r}(t) \triangleq r(t)w(t)$, with $w(t)$ is a rectangular window of length $L_o T$. It is recognized that the summation in (27) is just the filtering of $\tilde{r}(t)$, through $g(t)$, followed then by sampling at $t = (i + \varepsilon)T$. Notice that since $g(t)$ is a square-root raised cosine filter, $g(t) = g(-t)$ and (27) actually corresponds to the sampled matched filter output. 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 (26) and noting that asymptotically, the range of i can be approximated by $i = 0, \dots, L_o - 1$, we have

$$\Lambda(k) \approx \sum_{i=0}^{L_o-1} |x(iT + kT/K)|^2. \quad (28)$$

The proposed CML feedforward timing delay estimator in (25) can then be rewritten as

$$\hat{\varepsilon} = -\frac{1}{2\pi} \arg\left\{ \sum_{k=0}^{K-1} \left(\sum_{i=0}^{L_o-1} |x(iT + kT/K)|^2 \right) e^{-j2\pi k/K} \right\} \quad (29)$$

$$= -\frac{1}{2\pi} \arg\left\{ \sum_{l=0}^{KL_o-1} |x(lT/K)|^2 e^{-j2\pi l/K} \right\}. \quad (30)$$

Therefore, when $L_o \rightarrow \infty$ and $K=4$, we have the well known squaring algorithm [3]. Figs. 2b and 2c show the block diagrams for the proposed estimator (25) with $K = 4$ and the squaring algorithm. It can be seen that the structures of the proposed algorithm and the squaring algorithm are very alike. Note that both the proposed algorithm and the squaring algorithm require four samples per symbol period to form the timing estimate. For the proposed estimator, the received signal is first sampled with minimum oversampling ratio $Q=2$ and then samples with $K = 4$ different phases are generated by filtering (see (27)). For the squaring algorithm, the four different samples per symbol period are directly obtained through sampling. Notice that the squaring algorithm might work also by first sampling at $Q = 2$ and then the intermediate (additional two) samples are computed by interpolation before symbol timing estimation. Although the proposed estimator and the squaring algorithm have many characteristics in common, simulation results presented in Section VI show that the proposed estimator outperforms the squaring algorithm for reduced length observation records.

V. ANALYTICAL PERFORMANCE ANALYSIS

In this section, we derive the mean square error (MSE) expressions for the proposed estimator as a function of E_s/N_o . First, express the true timing delay ε_o as follows

$$\varepsilon_o = -\frac{1}{2\pi} \arg(e^{-j2\pi\varepsilon_o}). \quad (31)$$

From (25) and (31), the MSE for a specific delay is given by²

$$\mathbf{E}[(\hat{\varepsilon} - \varepsilon_o)^2] = \left(\frac{1}{2\pi}\right)^2 \mathbf{E} \left[\left(\arctan \left\{ \frac{\Im\{\phi\}}{\Re\{\phi\}} \right\} \right)^2 \right] \quad (32)$$

where

$$\phi \triangleq e^{j2\pi\varepsilon_o} \sum_{k=0}^{K-1} \Lambda(k) e^{-j2\pi k/K}. \quad (33)$$

Applying the approximation $\arctan(x) \approx x$ for small x , we have

$$\begin{aligned} \mathbf{E}[(\hat{\varepsilon} - \varepsilon_o)^2] &\approx \left(\frac{1}{2\pi}\right)^2 \mathbf{E} \left[\left(\frac{\phi - \phi^*}{j(\phi + \phi^*)} \right)^2 \right] \\ &\approx - \left(\frac{1}{2\pi}\right)^2 \frac{\mathbf{E}[\phi^2] - 2\mathbf{E}[\phi\phi^*] + \mathbf{E}[(\phi^*)^2]}{\mathbf{E}[\phi^2] + 2\mathbf{E}[\phi\phi^*] + \mathbf{E}[(\phi^*)^2]} \\ &= - \left(\frac{1}{2\pi}\right)^2 \frac{\Re\{\mathbf{E}[\phi^2]\} - \mathbf{E}[\phi\phi^*]}{\Re\{\mathbf{E}[\phi^2]\} + \mathbf{E}[\phi\phi^*]}. \end{aligned} \quad (34)$$

The last equality in (34) comes from the fact that $\mathbf{E}[(\phi^*)^2] = (\mathbf{E}[\phi^2])^*$. The second approximation in (34) can be justified using similar arguments as in [23]. A close examination of the fraction $(\phi - \phi^*)^2/(\phi + \phi^*)^2$ in (34) illustrates that the mean of its denominator is much larger than the mean of its numerator, and the standard deviations of its numerator and denominator are in general much smaller than the mean of denominator. Figure 3 plots $\mathbf{E}[(\phi - \phi^*)^2]$ and $\mathbf{E}[(\phi + \phi^*)^2]$ as a function of E_s/N_o when $g(t)$ is a square-root raised cosine pulse with $\alpha = 0.5$, $Q = 2$, $K = 4$, $L_o=100$ and $L_g = 3$ for $\varepsilon_o = 0, 0.25, 0.5$ and 0.75 . Note that all curves for different values of ε_o overlap. It can be seen that for $E_s/N_o \geq 10\text{dB}$, $\mathbf{E}[(\phi - \phi^*)^2]$ is much smaller than $\mathbf{E}[(\phi + \phi^*)^2]$. The same result can be obtained for different other pulse shapes $g(t)$. Also, one can check that at medium and high E_s/N_o , the standard deviations of $(\phi - \phi^*)^2$ and $(\phi + \phi^*)^2$ are small relative to $\mathbf{E}[(\phi + \phi^*)^2]$. All these considerations justify the second approximation made in (34).

From (33), we note that

$$\mathbf{E}[\phi^2] = e^{j4\pi\varepsilon_o} \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \mathbf{E}[\Lambda(k_1)\Lambda(k_2)] e^{-j2\pi k_1/K} e^{-j2\pi k_2/K}, \quad (35)$$

²Notation $\mathbf{E}[z]$ denotes the mean of random variable z , while $\Im\{\phi\}$ stands for the imaginary part of ϕ .

$$\mathbf{E}[\phi\phi^*] = \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} \mathbf{E}[\Lambda(k_1)\Lambda^*(k_2)]e^{-j2\pi k_1/K}e^{j2\pi k_2/K}. \quad (36)$$

It is proved in Appendix I that

$$\begin{aligned} \mathbf{E}[\Lambda(k_1)\Lambda(k_2)] &= \frac{E_s^2}{T^2} \left\{ \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T]\text{tr}[\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}^T] + \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}^T] + c(k_1, k_2) \right. \\ &+ \left(\frac{E_s}{N_o}\right)^{-1} Q \left\{ \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T]\text{tr}[\mathbf{B}_{k_2}] + \text{tr}[\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}^T\mathbf{B}_{k_1}] + \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T\mathbf{B}_{k_2}] + \text{tr}[\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}^T]\text{tr}[\mathbf{B}_{k_1}] \right\} \\ &\left. + \left(\frac{E_s}{N_o}\right)^{-2} Q^2 \left\{ \text{tr}[\mathbf{B}_{k_1}]\text{tr}[\mathbf{B}_{k_2}] + \text{tr}[\mathbf{B}_{k_1}\mathbf{B}_{k_2}] \right\} \right\}, \quad (37) \end{aligned}$$

$$\begin{aligned} \mathbf{E}[\Lambda(k_1)\Lambda^*(k_2)] &= \frac{E_s^2}{T^2} \left\{ \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T]\text{tr}[\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}] + \text{tr}[\mathbf{B}_{k_1}^T\mathbf{P}_{\varepsilon_o}\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}] + c(k_1, k_2) \right. \\ &+ \left(\frac{E_s}{N_o}\right)^{-1} Q \left\{ \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T]\text{tr}[\mathbf{B}_{k_2}] + \text{tr}[\mathbf{P}_{\varepsilon_o}\mathbf{B}_{k_2}\mathbf{B}_{k_1}^T] + \text{tr}[\mathbf{B}_{k_1}\mathbf{P}_{\varepsilon_o}^T\mathbf{B}_{k_2}^T] + \text{tr}[\mathbf{B}_{k_2}\mathbf{P}_{\varepsilon_o}]\text{tr}[\mathbf{B}_{k_1}] \right\} \\ &\left. + \left(\frac{E_s}{N_o}\right)^{-2} Q^2 \left\{ \text{tr}[\mathbf{B}_{k_1}]\text{tr}[\mathbf{B}_{k_2}] + \text{tr}[\mathbf{B}_{k_1}^T\mathbf{B}_{k_2}] \right\} \right\}, \quad (38) \end{aligned}$$

where $\text{tr}[\cdot]$ denotes the trace of a matrix. In (37) and (38),

$$\mathbf{B}_k \triangleq \mathbf{A}_\varepsilon(\mathbf{A}_\varepsilon^H\mathbf{A}_\varepsilon)^{-1}\mathbf{A}_\varepsilon^H|_{\varepsilon=k/K}, \quad (39)$$

\mathbf{P}_ε is an $L_oQ \times L_oQ$ matrix with the $(i, j)^{\text{th}}$ element ($i, j = 0, 1, \dots, L_oQ - 1$) given by

$$[\mathbf{P}_\varepsilon]_{ij} \triangleq \sum_{n=-\infty}^{\infty} g^*(iT/Q - nT - \varepsilon T)g(jT/Q - nT - \varepsilon T), \quad (40)$$

and

$$c(k_1, k_2) \triangleq \begin{cases} (m_4 - 2) \sum_{n=-\infty}^{\infty} [\mathbf{a}_n(\varepsilon_o)^H\mathbf{B}_{k_1}\mathbf{a}_n(\varepsilon_o)][\mathbf{a}_n(\varepsilon_o)^H\mathbf{B}_{k_2}\mathbf{a}_n(\varepsilon_o)] & \text{for linear modulations} \\ \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} (-1)^{|n_1-n_2|} [\mathbf{a}_{n_1}(\varepsilon_o)^H\mathbf{B}_{k_1}\mathbf{a}_{n_2}(\varepsilon_o)][\mathbf{a}_{n_1}(\varepsilon_o)^H\mathbf{B}_{k_2}\mathbf{a}_{n_2}(\varepsilon_o)] & \\ - 2 \sum_{n=-\infty}^{\infty} [\mathbf{a}_n(\varepsilon_o)^H\mathbf{B}_{k_1}\mathbf{a}_n(\varepsilon_o)][\mathbf{a}_n(\varepsilon_o)^H\mathbf{B}_{k_2}\mathbf{a}_n(\varepsilon_o)] & \text{for MSK and GMSK} \end{cases} \quad (41)$$

where $m_4 = \mathbf{E}[|d_i|^4]$ is the fourth order moment of the transmitted symbols, which is fixed for a specific constellation (e.g., $m_4 = 1$ for PSK and $m_4 > 1$ for QAM). Therefore, the MSE for a specific delay ε_o can be found by using (34)-(38). As the symbol timing delay ε_o is assumed to be uniformly distributed in $[0, 1)$, the average MSE is calculated by numerical integration of (34).

Notice that the MSE expressions in this section can only be regarded as an approximated analysis for GMSK since only the principle component of LE is taken into consideration. However, from the results to be presented in next section, excellent agreement between analytical expressions and simulations can be observed (see Fig. 9), only a small deviation occurs at very high SNRs.

VI. SIMULATION RESULTS AND DISCUSSIONS

In this section, the performance of the proposed algorithm and other existing symbol timing estimators are assessed by Monte Carlo simulations and then compared with the analytical results derived in the last section, the CCRB [10] and the MCRB [14]. In all the simulations, the observation length is fixed to $L_o=100$, and ε_o is uniformly distributed in the range $[0, 1)$. θ_o is generated as a 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. In all figures, the CCRB and the MCRB are plotted as references.

First consider the case of linear modulations. QPSK is chosen as the symbol constellation. The oversampling ratio for the proposed estimator is $Q=2$, $g(t)$ is the square-root raised cosine pulse with roll-off factor $\alpha=0.5$, the number of ISI symbols introduced by one side of $g(t)$ is assumed to be $L_g=3$. Figs. 4a and 4b show the MSE against E_s/N_o for the proposed algorithm and the IDFT-based CML estimator for $K = 4$ and $K = 8$, respectively. It can be seen (from both Figs. 4a and 4b) that the proposed algorithm has a performance similar to that of IDFT-based CML estimator with $K'=2048$. This further justifies the approximation in (23). Note that, for $K = 4$, the self noise is not completely eliminated for both IDFT-based CML estimator and the proposed estimator (as seen from the MSE departure from CCRB at high SNR in Fig. 4a). This can be explained as follows. For the IDFT-based CML estimator, the self-noise is due to the small value of K chosen, since in the derivation, it is assumed that the CML function can be *completely* represented by K samples. However, there is no guarantee that $K = 4$ is sufficient (although $K = 4$ results in pretty good performance). Increasing the value of K to 8 removes the self-noise of the IDFT-based CML estimator (with $K' = 2048$) as shown in Fig. 4b. For the proposed estimator, although it is also required that K should be large enough such that $\Lambda(\varepsilon)$ can be represented by its samples, the self-noise is due to another more critical factor – the approximation (23) in the CML function. This can be seen from the fact that the performance

of the proposed estimator does not improve by increasing K from 4 to 8 (compare Figs. 4a and 4b). As $K = 4$ is good enough for the proposed estimator, $K = 4$ is used for the rest of the paper.

Figure 5 illustrates the very close match between the simulation and the analytical results derived in the last section. It is also clear that, for the SNRs under consideration, the performance of the proposed algorithm is very close to the CCRB, which means that the proposed estimator almost reaches the ultimate performance of the CML principle. Furthermore, the CCRB is close to the MCRB. Since MCRB is a lower bound on the variance of *any* unbiased estimate, this shows that the proposed algorithm is close to optimal for a wide range of E_s/N_o . Notice that at E_s/N_o around 30dB, an MSE floor begins to occur (due to the approximation (23) in the CML function), but at that high SNR, the estimation MSE achieved by the proposed estimator is already very small (on the order of 10^{-6}), so the effect of the MSE floor becomes relatively less critical.

Fig. 6 compares the performance of the proposed estimator with that of the square nonlinearity estimator (with $Q=4$) [3]. It is apparent that the proposed estimator outperforms the square nonlinearity estimator, especially at high E_s/N_o . This is because for finite observation length, $\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon \neq \mathbf{I}$, and the self-noise is better cancelled by the matrix $(\mathbf{A}_\varepsilon^H \mathbf{A}_\varepsilon)^{-1}$ than \mathbf{I} . Fig. 7 compares the performances of the proposed algorithm with the existing state-of-the-art feedforward algorithms that require only two samples per symbol to operate: Mengali [14, pp.401], Zhu *etal.* [16] and Lee [17]. It can be seen that while the performances for different algorithms are similar at low E_s/N_o , the proposed algorithm has the smallest MSE at high E_s/N_o .

Next consider that MSK is the modulation format. Figure 8 shows the performances of the proposed estimator (with $Q = 2$, $Q = 4$ and $Q = 8$) and the low-SNR approximated maximum likelihood (ML) algorithm [19] for MSK. The number of ISI symbols introduced by one side of $\bar{g}(t)$ is assumed to be 1. The following observations can be inferred from Fig. 8. First, it can be seen that, for the proposed algorithm, the higher the oversampling ratio, the better the performance. This is because the pulse $\bar{g}(t)$ is time limited [18], so its frequency response is not bandlimited; a higher oversampling ratio reduces the aliasing and thus provides better performance. Second, the theoretical MSE analysis matches very well the simulation results. Third, although a higher oversampling ratio increases the range of SNR's over which the performance of the proposed estimator comes close to the CCRB, MSE floors still occur at

high SNR's due to the approximation (23) assumed in the derivation of estimator. Furthermore, the CCRB is far away from the MCRB and the simulation results show that the low-SNR approximated ML algorithm [19] approaches the MCRB. Therefore, direct application of the CML principle is not suitable for the MSK modulation.

Now let consider the GMSK modulation. Fig. 9 show the performances of the proposed estimator (with $Q = 2$ and $Q = 4$) and the low-SNR approximated ML algorithm [19] for GMSK with pre-modulator bandwidth $BT = 0.5$. The number of ISI symbols introduced by one side of $\bar{g}(t)$ is assumed to be 2. Notice that although the proposed estimator is based on the approximated linear model (9), the GMSK signal in the simulation is generated according to (7) without approximation. The MCRB for GMSK is exact and its expression can be found in [14]. For the CCRB, it is based on the approximated linear model (9). Although the resulting CCRB is not exact, it is still a valid lower bound for the proposed estimator since when the proposed estimator is applied to the true GMSK signal, the ignored components in LE would become interferences and the performances would be poorer than that predicted by the CCRB, which assumes no interference from other components of LE. Note that the CCRB obtained by expressing the GMSK signal using all the LE components (as done in [10]) is not applicable here, since in that case, the resultant CCRB is *conditioned* on the fact that all the pseudo-symbols are being jointly estimated together with the unknown timing offset, while in the proposed estimator, only the pseudo-symbols related to the first LE component are estimated.

From Fig. 9, it can be seen that, for the proposed estimator, a higher oversampling ratio also results in better performance due to the same reason as in the case of MSK modulation. However, by comparing Figs. 8 and 9, if the same oversampling ratio is used, it is found that the performance of the proposed estimator for GMSK modulation is better than that corresponding to MSK. This is due to the fact that the pulse $\bar{g}(t)$ is longer in GMSK than in MSK (although they both are time-limited), therefore, with the same oversampling ratio, the aliasing introduced in GMSK is smaller than that in MSK. Second, it is obvious that the analytical MSE expressions derived in the last section match very well with the simulation results. Only for the case of $K = 4$ and at SNR=50-60 dB, the analytical MSE expressions predict a slightly better performance than simulations. Third, the performance of the proposed estimator with $Q = 4$ comes very close to the CCRB for $E_s/N_o \leq 40$ dB. The MSE floor, which is caused by the approximation (23) in the CML function, begins to occur only for $E_s/N_o > 40$ dB. Notice that the effect of the

approximation (9) in the system model (which results in the gap between analytical MSE and simulations) is much smaller than that of approximation (23) in the CML function (which causes the MSE floor). Compared to the low-SNR approximated ML algorithm [19], at low SNR's the proposed estimator exhibits poorer performance, but for medium and high SNR's, the proposed estimator performs much better.

Finally, notice that the CCRB is far away from the MCRB as in the case of MSK. Since the CCRB is a valid bound only for estimators that rely on quadratic nonlinearities [10], it is expected that algorithms exploiting higher-order (> 2) nonlinearities might exist with performances closer to the MCRB. An example of such an algorithm is the low-SNR approximated ML algorithm [19], for which we already demonstrated that its performance is closer to the MCRB at low SNRs. The next question is whether there is an estimator whose performance comes close to the MCRB for a larger range of SNRs. This is a subject open to future investigations.

VII. CONCLUSIONS

A new feedforward symbol-timing estimator based on the conditional maximum likelihood principle was proposed. An approximation was applied in the Fourier series expansion of the CML function so that the complexity of the proposed estimator is greatly reduced. It was shown, analytically and via simulations, that the performances of the proposed estimator for linear modulations are in general very close to the CCRB and MCRB for $\text{SNR} \leq 30\text{dB}$. For higher SNRs, MSE floors occur, but notice that at these high SNRs, the MSE achieved by the proposed estimator is already very small, so the effect of MSE floors becomes relatively less critical. Furthermore, for linear modulations with the transmit pulse being a square-root raised cosine pulse, the proposed estimator was shown to be asymptotically equivalent to the well-known square nonlinearity estimator [3]. However, in the presence of reduced number of samples, the proposed estimator presents better performance than [3]. For MSK and GMSK modulations, it was found that although the performances of the proposed estimator come very close to the CCRB at certain SNR ranges, however, the CCRB is quite far away from the MCRB, and there exists an alternative algorithm that come closer to the MCRB. Therefore, it was concluded that the proposed estimator is more suitable for linear modulations than MSK and GMSK modulations.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for carefully reading this manuscript and for their constructive comments, which greatly improved the presentation of this paper.

Appendix I: Proof of (37) and (38)

From the definition of $\Lambda(\varepsilon)$ in (16), we have

$$\begin{aligned} \mathbf{E}[\Lambda(k_1)\Lambda(k_2)] &= \mathbf{E}[\mathbf{r}^H \mathbf{B}_{k_1} \mathbf{r} \mathbf{r}^H \mathbf{B}_{k_2} \mathbf{r}] \\ &= \sum_{l_1=0}^{L_o Q-1} \sum_{i_1=0}^{L_o Q-1} \sum_{l_2=0}^{L_o Q-1} \sum_{i_2=0}^{L_o Q-1} b_{i_1 l_1}^{(k_1)} b_{i_2 l_2}^{(k_2)} \mathbf{E}[r^*(i_1)r(l_1)r^*(i_2)r(l_2)], \end{aligned} \quad (42)$$

where $b_{ij}^{(k)}$ is the $(i, j)^{th}$ element in \mathbf{B}_k . Now we concentrate on $\mathbf{E}[r^*(i_1)r(l_1)r^*(i_2)r(l_2)]$ which is given by

$$\begin{aligned} &\mathbf{E}[r^*(i_1)r(l_1)r^*(i_2)r(l_2)] \\ &= \mathbf{E} \left[\left(e^{-j\theta_o} \sqrt{\frac{E_s}{T}} \sum_{n_1} d_{n_1}^* g^*(i_1 T/Q - n_1 T - \varepsilon_o T) + \eta^*(i_1) \right) \right. \\ &\quad \cdot \left(e^{j\theta_o} \sqrt{\frac{E_s}{T}} \sum_{n_2} d_{n_2} g(l_1 T/Q - n_2 T - \varepsilon_o T) + \eta(l_1) \right) \\ &\quad \cdot \left(e^{-j\theta_o} \sqrt{\frac{E_s}{T}} \sum_{n_3} d_{n_3}^* g^*(i_2 T/Q - n_3 T - \varepsilon_o T) + \eta^*(i_2) \right) \\ &\quad \left. \cdot \left(e^{j\theta_o} \sqrt{\frac{E_s}{T}} \sum_{n_4} d_{n_4} g(l_2 T/Q - n_4 T - \varepsilon_o T) + \eta(l_2) \right) \right]. \end{aligned} \quad (43)$$

First, let consider linear modulations, in which case we have

$$\mathbf{E}[d_{n_1} d_{n_2}] = 0, \quad (44)$$

$$\mathbf{E}[d_{n_1} d_{n_2}^*] = \delta(n_1, n_2), \quad (45)$$

$$\mathbf{E}[\eta(i_1)\eta(i_2)] = 0, \quad (46)$$

$$\mathbf{E}[\eta^*(i_1)\eta(i_2)] = \frac{N_o Q}{T} \delta(i_1, i_2), \quad (47)$$

$$\mathbf{E}[d_{n_1}^* d_{n_2} d_{n_3}^* d_{n_4}] = \begin{cases} 1 & \text{for } n_1 = n_2 \neq n_3 = n_4 \\ 1 & \text{for } n_1 = n_4 \neq n_2 = n_3 \\ m_4 & \text{for } n_1 = n_4 = n_2 = n_3 \\ 0 & \text{otherwise,} \end{cases} \quad (48)$$

and therefore, 10 out of the 16 terms which result from (43) vanish. With the definitions

$$\begin{aligned}
P_\varepsilon(i, j) &\triangleq \sum_n g^*(iT/Q - nT - \varepsilon T)g(jT/Q - nT - \varepsilon T), \\
Z_\varepsilon(i, j, k, l) &\triangleq \sum_n g^*(iT/Q - nT - \varepsilon T)g(jT/Q - nT - \varepsilon T) \\
&\quad \cdot g^*(kT/Q - nT - \varepsilon T)g(lT/Q - nT - \varepsilon T),
\end{aligned} \tag{49}$$

the remaining terms can be expressed as

$$\mathbf{E}[r^*(i_1)r(l_1)r^*(i_2)r(l_2)] = S_1 + S_{12} + S_{14} + S_{23} + S_{34} + S_4, \tag{50}$$

where

$$\begin{aligned}
S_1 &\triangleq \frac{E_s^2}{T^2} \sum_{n_1} \sum_{n_2} \sum_{n_3} \sum_{n_4} \mathbf{E}[d_{n_1}^* d_{n_2} d_{n_3}^* d_{n_4}] g^*(i_1 T/Q - n_1 T - \varepsilon_o T) g(l_1 T/Q - n_2 T - \varepsilon_o T) \\
&\quad \cdot g^*(i_2 T/Q - n_3 T - \varepsilon_o T) g(l_2 T/Q - n_4 T - \varepsilon_o T) \\
&= \frac{E_s^2}{T^2} \{P_{\varepsilon_o}(i_1, l_1)P_{\varepsilon_o}(i_2, l_2) + P_{\varepsilon_o}(i_1, l_2)P_{\varepsilon_o}(i_2, l_1) + (m_4 - 2)Z_{\varepsilon_o}(i_1, l_1, i_2, l_2)\}, \tag{51}
\end{aligned}$$

$$\begin{aligned}
S_{12} &\triangleq \frac{E_s}{T} \sum_{n_1} \sum_{n_2} \mathbf{E}[d_{n_1}^* d_{n_2}] \mathbf{E}[\eta^*(i_2)\eta(l_2)] g^*(i_1 T/Q - n_1 T - \varepsilon_o T) g(l_1 T/Q - n_2 T - \varepsilon_o T) \\
&= \frac{E_s N_o Q}{T^2} P_{\varepsilon_o}(i_1, l_1) \delta(i_2, l_2), \tag{52}
\end{aligned}$$

$$\begin{aligned}
S_{14} &\triangleq \frac{E_s}{T} \sum_{n_1} \sum_{n_4} \mathbf{E}[d_{n_1}^* d_{n_4}] \mathbf{E}[\eta(l_1)\eta^*(i_2)] g^*(i_1 T/Q - n_1 T - \varepsilon_o T) g(l_2 T/Q - n_4 T - \varepsilon_o T) \\
&= \frac{E_s N_o Q}{T^2} P_{\varepsilon_o}(i_1, l_2) \delta(i_2, l_1), \tag{53}
\end{aligned}$$

$$\begin{aligned}
S_{23} &\triangleq \frac{E_s}{T} \sum_{n_2} \sum_{n_3} \mathbf{E}[d_{n_2} d_{n_3}^*] \mathbf{E}[\eta^*(i_1)\eta(l_2)] g(l_1 T/Q - n_2 T - \varepsilon_o T) g^*(i_2 T/Q - n_3 T - \varepsilon_o T) \\
&= \frac{E_s N_o Q}{T^2} P_{\varepsilon_o}(i_2, l_1) \delta(i_1, l_2), \tag{54}
\end{aligned}$$

$$\begin{aligned}
S_{34} &\triangleq \frac{E_s}{T} \sum_{n_3} \sum_{n_4} \mathbf{E}[d_{n_3}^* d_{n_4}] \mathbf{E}[\eta^*(i_1)\eta(l_1)] g^*(i_2 T/Q - n_3 T - \varepsilon_o T) g(l_2 T/Q - n_4 T - \varepsilon_o T) \\
&= \frac{E_s N_o Q}{T^2} P_{\varepsilon_o}(i_2, l_2) \delta(i_1, l_1), \tag{55}
\end{aligned}$$

$$\begin{aligned}
S_4 &\triangleq \mathbf{E}[\eta^*(i_1)\eta(l_1)\eta^*(i_2)\eta(l_2)] \\
&= \mathbf{E}[\eta^*(i_1)\eta(l_1)] \mathbf{E}[\eta^*(i_2)\eta(l_2)] + \mathbf{E}[\eta^*(i_1)\eta^*(i_2)] \mathbf{E}[\eta(l_1)\eta(l_2)] \\
&\quad + \mathbf{E}[\eta^*(i_1)\eta(l_2)] \mathbf{E}[\eta(l_1)\eta^*(i_2)] + \mathbf{E}[\eta^*(i_1)] \mathbf{E}[\eta(l_1)] \mathbf{E}[\eta^*(i_2)] \mathbf{E}[\eta(l_2)] \\
&= \frac{N_o^2 Q^2}{T^2} (\delta(i_1, l_1)\delta(i_2, l_2) + \delta(i_1, l_2)\delta(i_2, l_1)). \tag{56}
\end{aligned}$$

Plugging (50)-(56) back into (42) and expressing the summations using matrices, some straightforward calculations lead to (37). A similar procedure can be used to prove (38).

Now, let consider MSK and GMSK. Since the pseudo-symbols in (10) are not circularly symmetric, (44) and (48) have to be modified accordingly. After some lengthy but straightforward calculations, it is found that

$$\mathbf{E}[d_{n_1}d_{n_2}] = (-1)^{n_1}\delta(n_1, n_2) , \quad (57)$$

$$\mathbf{E}[d_{n_1}^*d_{n_2}d_{n_3}^*d_{n_4}] = \begin{cases} 1 & \text{for } n_1 = n_2 \neq n_3 = n_4 \\ 1 & \text{for } n_1 = n_4 \neq n_2 = n_3 \\ 1 & \text{for } n_1 = n_4 = n_2 = n_3 \\ (-1)^{|n_2-n_1|} & \text{for } n_1 = n_3 \neq n_2 = n_4 \\ 0 & \text{otherwise .} \end{cases} \quad (58)$$

Due to (57), two more cross terms in the expansion of (43) have to be considered. One of them is S_{13} , which is given by

$$S_{13} \triangleq e^{-2j\theta_o} \frac{E_s}{T} \sum_{n_1} \sum_{n_3} \mathbf{E}[d_{n_1}^*d_{n_3}] \mathbf{E}[\eta(l_1)\eta(l_2)] g^*(i_1T/Q - n_1T - \varepsilon_oT) g^*(i_2T/Q - n_3T - \varepsilon_oT).$$

But thanks to the correlation property of noise samples, $S_{13} = 0$. The other extra term is also zero due to the same reason. For the fourth order moment in (58), compared to the corresponding expression for linear modulations (48), we notice that $m_4 = 1$ and there is an extra non-zero fourth-order moment. Therefore, apart from setting $m_4 = 1$ in S_1 , an extra term has to be added to S_1 in (51). The modified S_1 , denoted as S_{1MSK} , can be expressed as

$$S_{1MSK} = S_1|_{m_4=1} + \frac{E_s^2}{T^2} \sum_{\substack{n_1 \\ n_1 \neq n_2}} \sum_{n_2} (-1)^{|n_2-n_1|} g^*(i_1T/Q - n_1T - \varepsilon_oT) g(l_1T/Q - n_2T - \varepsilon_oT) \\ \cdot g^*(i_2T/Q - n_1T - \varepsilon_oT) g(l_2T/Q - n_2T - \varepsilon_oT). \quad (59)$$

Plugging (59) into (42), and then expressing the multiplications using matrix notation, it can be proved that the only change is the definition of $c(k_1, k_2)$, which is given in (41).

REFERENCES

- [1] C. W. Farrow, "A continuously variable digital delay element," *Proc. of ISCAS88*, pp. 2641-2645, 1998.
- [2] L. Erup, F. M. Gardner and R. A. Harris, "Interpolation in digital modems – part II: implementation and performance," *IEEE Trans. Commun.*, vol. 41, pp. 998-1008, June 1993.
- [3] M. Oerder and H. Meyr, "Digital filter and square timing recovery," *IEEE Trans. Commun.*, vol. 36, pp. 605-612, May 1988.
- [4] E. Panayirci and E. Y. Bar-Ness, "A new approach for evaluating the performance of a symbol timing recovery system employing a general type of nonlinearity," *IEEE Trans. Commun.*, vol. 44, pp. 29-33, Jan. 1996.
- [5] 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.
- [6] K. E. Scott and E. B. Olasz, "Simultaneous clock phase and frequency offset estimation," *IEEE Trans. Commun.*, vol. 43, pp. 2263-2270, July 1995.
- [7] Y. Wang, E. Serpedin and P. Ciblat, "Blind Feedforward Cyclostationarity-Based Timing Estimation for Linear Modulations," accepted for publication in *IEEE Trans. on Wireless Communications*, May 2003.
- [8] M. Morelli, A. N. D'Andrea and U. Mengali, "Feedforward ML-based timing estimation with PSK signals," *IEEE Commun. Letters*, vol. 1, pp. 80-82, May 1997.
- [9] J. Riba, J. Sala and G. Vazquez, "Conditional maximum likelihood timing recovery: estimators and bounds," *IEEE Tran. Signal Processing*, Vol. 49, pp. 835-850, Apr. 2001.
- [10] G. Vazquez and J. Riba, "Non-data-aided digital synchronization," in *Signal Processing Advanced in Wireless and Mobile Communications: Volume 2* (edited by G. B. Giannakis, Y. Hua, P. Stoica and L. Tong), Prentice Hall 2001.
- [11] Y.-C. Wu and E. Serpedin, "Low-complexity feedforward symbol timing estimator using conditional maximum likelihood principle," *IEEE Commun. Letters*, vol. 8, pp. 168-170, May 2004.
- [12] P. Galko and S. Pasupathy, "On a class of Generalized MSK," *Proc. of ICC81*, pp. 2.4.1-2.4.5, June 1981.
- [13] O. Andrisano and M. Chiani, "The first Nyquist criterion applied to coherent receiver design for generalized MSK signals," *IEEE Trans. Commun.*, vol. 42, pp. 449-457, Feb/Mar/Apr. 1994.
- [14] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum Press, 1997.
- [15] S. M. Kay, *Fundamentals of Statistical Signal Processing - Estimation Theory*. Prentice Hall, 1993.
- [16] W.-P. Zhu, M. O. Ahmad, and M. N. S. Swamy, "A fully digital timing recovery scheme using two samples per symbol," *Proc. IEEE Int. Symp. on Circuit and Systems*, pp. 421-424, May. 2001.
- [17] S. J. Lee, "A new non-data-aided feedforward symbol timing estimator using two samples per symbol," *IEEE Commun. Letters*, Vol. 6, pp. 205-207, May. 2002.
- [18] P. A. Laurent, "Exact and approximate construction of digital phase modulations by superposition of amplitude modulated pulse (APM)," *IEEE Trans. Commun.*, vol. 34, pp. 150-160, Feb 1986.
- [19] M. Morelli and G. Vitetta, "Joint phase and timing recovery for MSK-type signals," *IEEE Trans. Commun.*, vol. 48, pp. 1997-1999, Dec. 2000.
- [20] Y.-C. Wu and T.-S. Ng, "Symbol timing recovery for GMSK modulations based on the square algorithm," *IEEE Comm. Lett.*, vol. 5, pp. 221-223, May 2001.
- [21] H. Samuelli, "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.

- [22] C.-K.-S. Pun, Y.-C. Wu, S.-C. Chan and K.-L. Ho, "On the design and efficient implementation of the Farrow structure," *IEEE Signal Processing Letters*, vol. 10, no. 7, pp. 189-192, Jul. 2003.
- [23] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. on Communications*, vol. 45, no. 12, pp. 1613-1621, Dec. 1997.

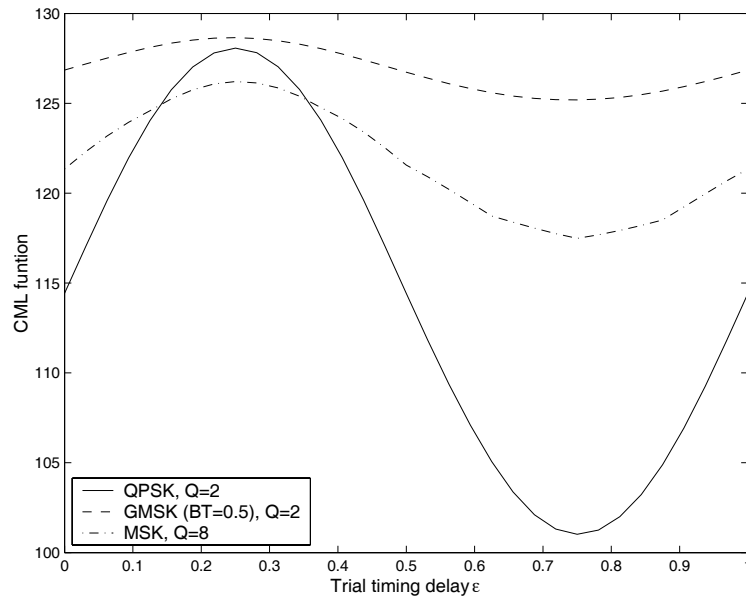


Fig. 1. Examples of CML functions.

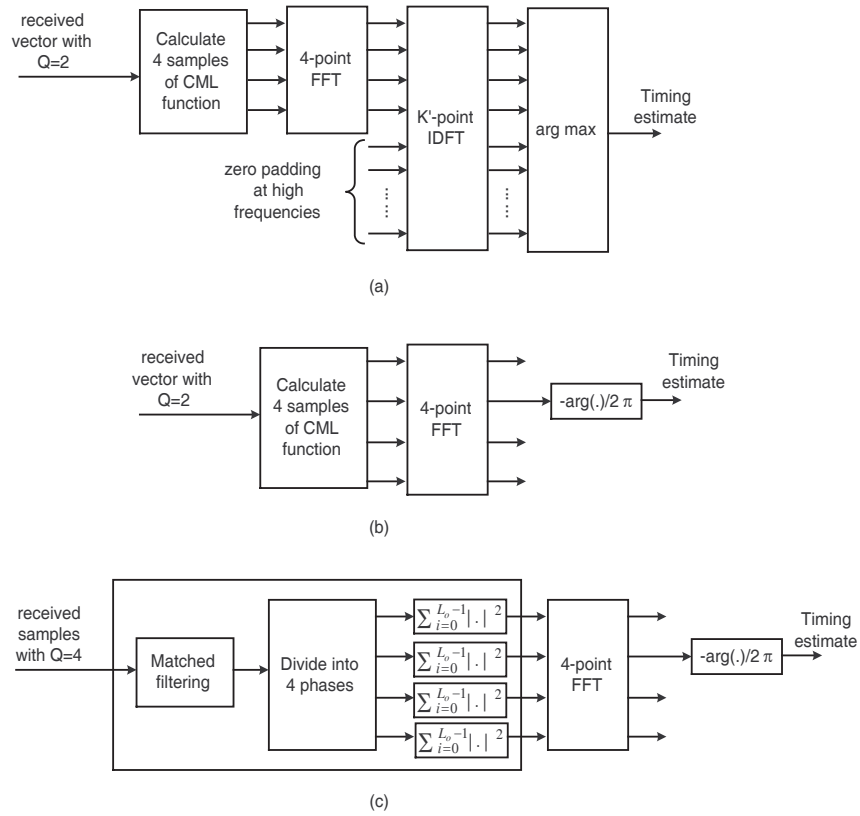


Fig. 2. Block diagrams for (a) the IDFT-based CML estimator ($K=4$), (b) the proposed estimator ($K=4$) and (c) the squaring estimator.

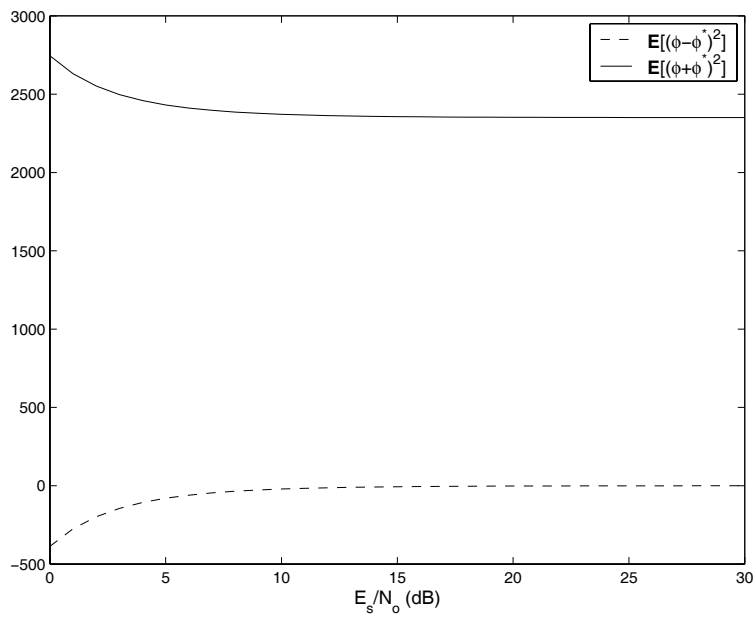
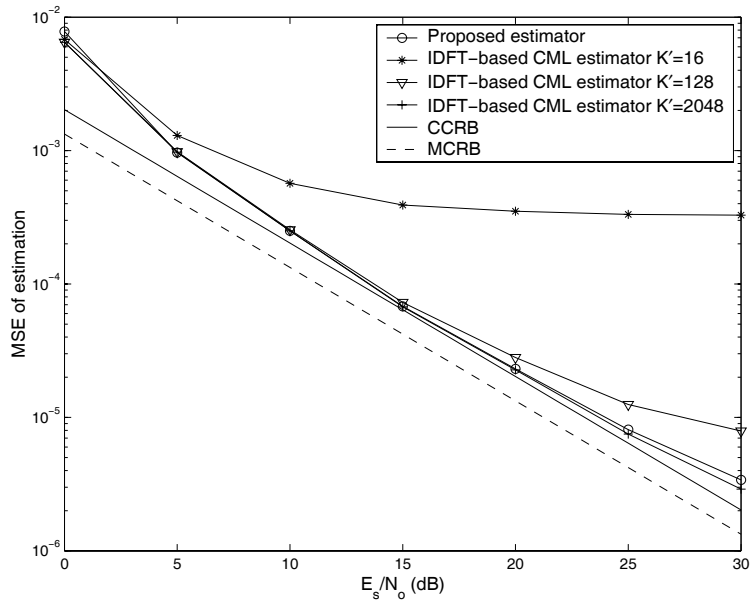
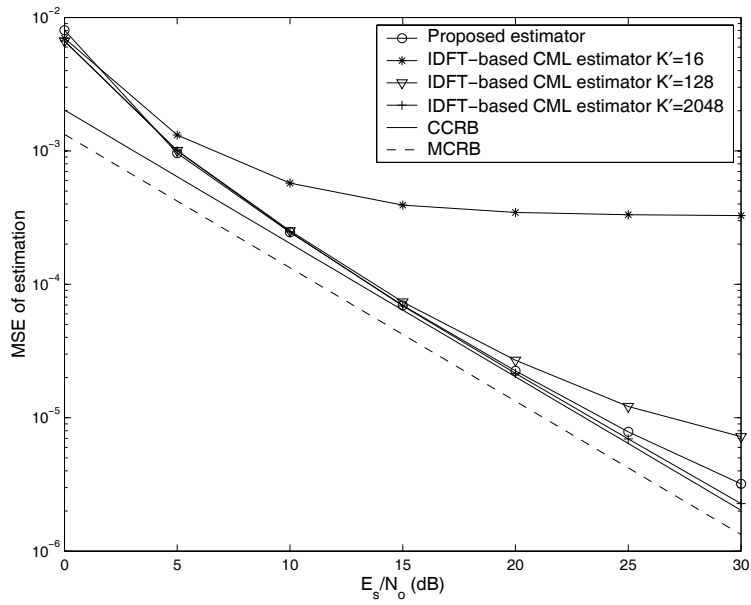


Fig. 3. Plots of $\mathbf{E}[(\phi - \phi^*)^2]$ and $\mathbf{E}[(\phi + \phi^*)^2]$ as a function of E_s/N_o for $\varepsilon_o = 0, 0.25, 0.5$ and 0.75 ($g(t)$ is a square-root raised cosine pulse with $\alpha = 0.5, Q = 2, K = 4, L_o=100$ and $L_g = 3$). Note that all curves for different values of ε_o overlap.



(a)



(b)

Fig. 4. MSE of the proposed estimator and the IDFT-based CML estimator with (a) $K = 4$, (b) $K = 8$ (QPSK, $Q=2$, $\alpha=0.5$, $L_o=100$ and $L_g=3$).

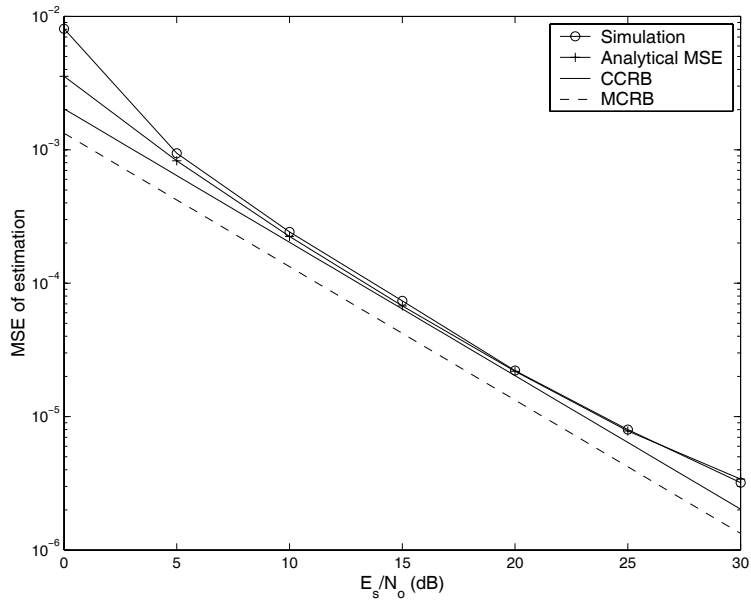


Fig. 5. Comparison between analytical MSE and simulations of the proposed estimator (QPSK, $Q=2$, $K=4$, $\alpha=0.5$, $L_o=100$ and $L_g=3$).

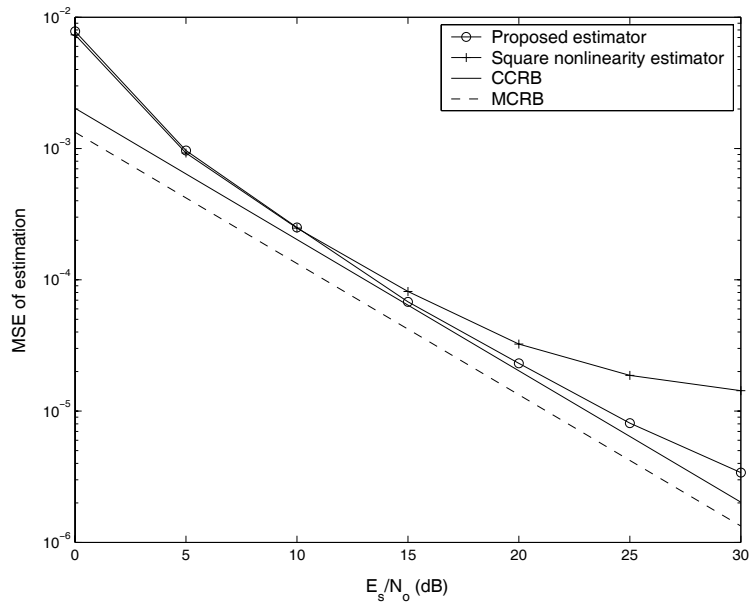


Fig. 6. Comparison of the MSE of the proposed estimator and the square nonlinearity estimator (QPSK, $\alpha=0.5$, $L_o = 100$ and $L_g=3$).

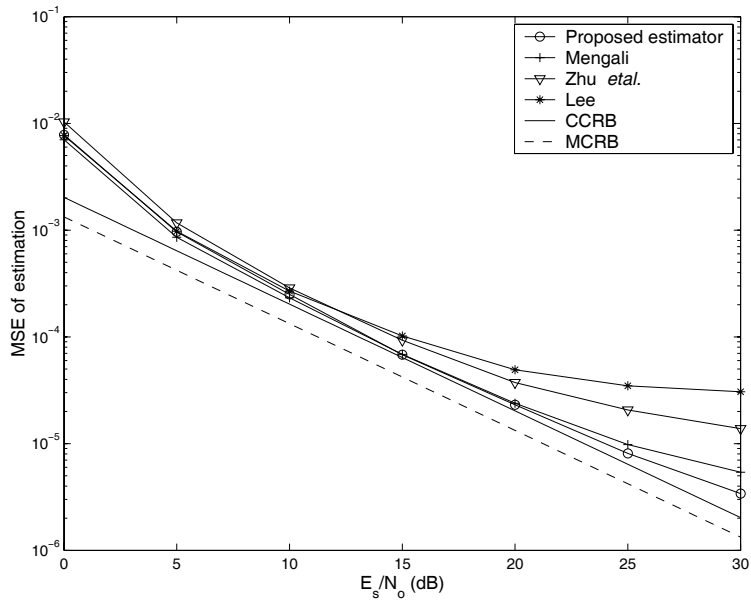


Fig. 7. MSE for the proposed estimator, the algorithms in [14, pp.401], [16] and [17] (QPSK, $Q = 2$, $\alpha=0.5$, $L_o=100$ and $L_g=3$).

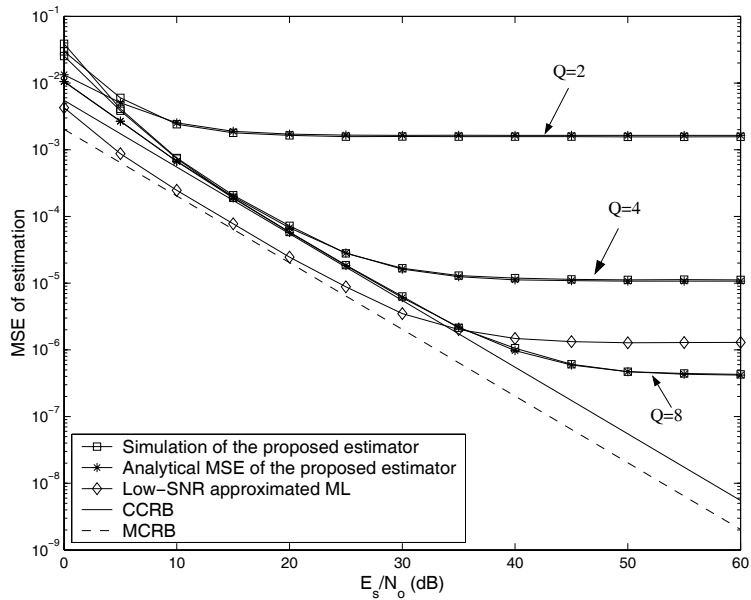


Fig. 8. MSE of the proposed estimator ($Q = 2$, $Q = 4$ and $Q = 8$) and the low-SNR approximated ML algorithm [19] for MSK ($L_{\bar{g}} = 1$ and $L_o=100$).

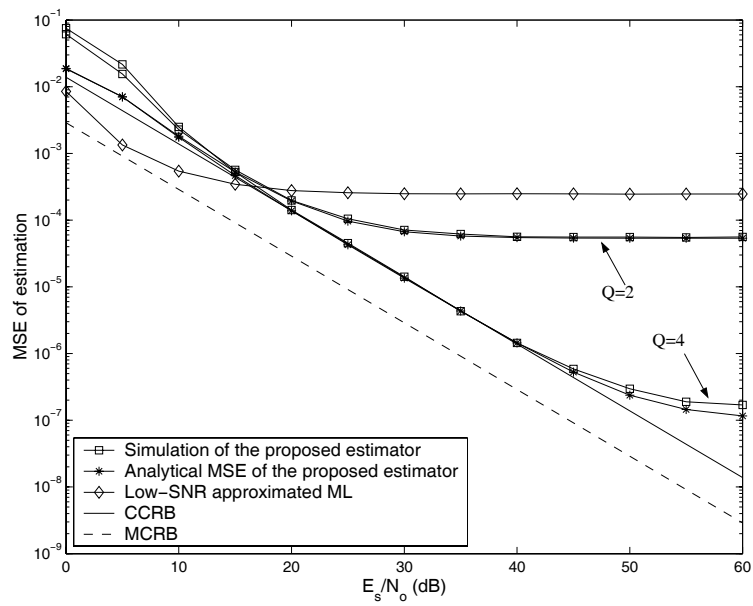


Fig. 9. MSE of the proposed estimator ($Q = 2$ and $Q = 4$) and the low-SNR approximated ML algorithm [19] for GMSK with $BT = 0.5$ ($L_{\bar{g}} = 2$ and $L_o=100$).