A New Importance-Sampling-Based Non-Data-Aided Maximum Likelihood Time Delay Estimator

Ahmed Masmoudi, Faouzi Bellili, Sofiène Affes, and Alex Stéphenne

INRS-EMT, 800, de la Gauchetière Ouest, Bureau 6900, Montreal, Qc, H5A 1K6, Canada
Emails: {masmoudi, bellili, affes}@emt.inrs.ca, stephenne@ieee.org

***Abstract*—In this paper, we present a new non-data-aided (NDA) maximum likelihood (ML) time delay estimator based on importance sampling (IS). We show that a grid search and lack of convergence from which most iterative estimators suffer can be avoided. It is assumed that the transmitted data are completely unknown at the receiver. Moreover the carrier phase is considered as an unknown nuisance parameter. The time delay remains constant over the observation interval and the received signal is corrupted by additive white Gaussian noise (AWGN). We use importance sampling to find the global maximum of the compressed likelihood function. Based on a global optimization procedure, the main idea of the new estimator is to generate realizations of a random variable using an importance function, which approximates the actual compressed likelihood function. We will see that the algorithm parameters affect the estimation performance and that with an appropriate parameter choice, even over a small observation interval, the time delay can be accurately estimated at far lower computational cost than with classical iterative methods.**

***Index Terms*—Symbol timing recovery, Monte-Carlo methods, optimization methods, Cramér-Rao lower bound (CRLB), non-data-aided (NDA) estimation.**

I. INTRODUCTION

Estimation of the symbol timing from the received data is a fundamental task for any digital receiver. Typically, in network communications, the time delay is usually assumed to be confined within the symbol duration [1] and recovered to allow for sampling the signal at accurate time instants. However, in many other applications such as radar [2] or sonar systems [3], where it can exceed the symbol duration, the time delay is used to localize targets.

Many references on this topic are available from the open literature where many estimators are derived trying to achieve the well-known Cramér-Rao lower bound (CRLB) [1-4]. Roughly speaking, the time delay estimation techniques are classified as [4]: minimum mean square error (MMSE) schemes, zero-forcing (ZF) schemes, early-late schemes and maximum-likelihood schemes. Typically, the maximum-likelihood (ML) principle is the most popular approach to obtain more reliable estimators. Indeed, it is well known that the ML estimator performs very close to the CRLB for large enough data records [5]. From another point of view, depending on whether the transmitted symbols are $a\ priori$ known or not, the timing recovery algorithms are also categorized into data-aided (DA) and non-data-aided (NDA) methods, respectively. Here, we focus on the NDA scenario in which the log-likelihood function is extremely non-linear with respect to the parameter of interest and hence it is difficult (if not impossible) to analytically find its global maximum. In this case, numerical methods are usually envisioned to find the ML estimates. In fact, iterative maximization procedures are widely used in this kind of problems, but their performances depend strictly on the initial guess of the parameter to be estimated. They may also suffer from a heavy complexity burden due to an inacceptable number of iterations.

In this paper, we introduce a novel non-iterative implementation of the conditional ML time delay estimator. We extend previous results for frequency estimation [6], Direction of Arrival (DOA) estimation [7] and joint DOA-Doppler estimation [8] to the problem of time delay estimation. We adopt the same model that has been widely applied in the context of sensor array processing [9] and which has been recently formulated in the special case of time-delay estimation [10].

The rest of this paper is organized as follows. In section II, we present the discrete-time signal model that will be used throughout this article. We introduce the importance sampling method that will be used to find the maximum of the compressed likelihood function in section III. Section IV deals with the choice of the importance function and discusses the impact of some parameters on the estimator performance. The newly proposed algorithm is developed in section V. Simulation results are discussed in section VI and, finally, some concluding remarks are drawn out in section VII.

II. SYSTEM MODEL AND COMPRESSED LIKELIHOOD FUNCTION

Consider a traditional communication system where the channel delays the transmitted signal and an AWGN with an overall power of N_0 corrupts the received signal which is expressed as follows:

$$y(t) = \sqrt{E_s}\ x(t-\tau^*)e^{j\theta} + w(t), \tag{1}$$

where τ^* is the unknown time delay of interest, θ is the unknown but deterministic channel distortion phase and $w(t)$ is an additive white Gaussian noise with independent real and imaginary parts, each of variance $N_0/2$. The transmitted signal

Work supported by a Canada Research Chair in Wireless Communications and a Discovery Accelerator Supplement from NSERC.

$x(t)$ is:

$$x(t) = \sum_{i=0}^{K-1} c_i \ h(t - i\ T), \tag{2}$$

where K is the total number of transmitted symbols in the observation interval, $\{c_i\}_{i=0}^{K-1}$ are the unknown complex-valued symbols, $h(t)$ is the shaping pulse and T stands for the symbol's duration.

In the sequel, the signal $y(t)$ is passed through an ideal lowpass filter of bandwidth $F_s/2$ and sampled at a rate $F_s = 1/T_s = k/T$, where k is a given integer which guarantees that F_s is above the Nyquist rate. Then, the discrete-time received signal can be more conveniently modeled in a matrix form as follows:

$$\begin{aligned} \boldsymbol{y} &= [y(0),\ y(T_s),\ ...,\ y((M-1)T_s)]^T \\ &= \boldsymbol{A}_{\tau^*}\boldsymbol{x} + \boldsymbol{w}, \end{aligned} \tag{3}$$

where M is the number of samples taken from $y(t)$ and:

$$\boldsymbol{w} = [w(0),\ w(T_s),\ ...,\ w((M-1)T_s)]^T, \tag{4}$$

$$\boldsymbol{A}_{\tau^*} = [\boldsymbol{a}_0(\tau^*),\ \boldsymbol{a}_1(\tau^*),\ ...,\ \boldsymbol{a}_{K-1}(\tau^*)], \tag{5}$$

with

$$\begin{aligned} \boldsymbol{a}_i(\tau^*) = [\ & h(-iT - \tau^*),\ h(T_s - iT - \tau^*),\ ..., \\ & h((M-1)T_s - iT - \tau^*)]^T. \end{aligned} \tag{6}$$

The covariance matrix of $\boldsymbol{w}$ after sampling is given by:

$$\boldsymbol{C_w} = \sigma^2 \boldsymbol{I}_M = 2N_0 F_s \boldsymbol{I}_M, \tag{7}$$

where $\boldsymbol{I}_M$ refers to the $(M \times M)$ identity matrix. The set of unknown parameters, including the data and the signal phase, are denoted by:

$$\boldsymbol{x} = \boldsymbol{a}e^{j\theta} = [a_0,\ a_1\ ...\ a_{K-1}]^T\ e^{j\theta}. \tag{8}$$

Considering the discrete-time signal model, the useful compressed likelihood function, parametrized by τ, was expressed in [10] as:

$$L_c(\boldsymbol{y};\tau) = \boldsymbol{y}^H \boldsymbol{A}_\tau (\boldsymbol{A}_\tau^T \boldsymbol{A}_\tau)^{-1} \boldsymbol{A}_\tau^T \boldsymbol{y}. \tag{9}$$

Note here that τ is any possible value of the time delay parameter τ^*.

III. Global Maximization of the Compressed Likelihood Function

A ML estimate of the time delay can be obtained by maximizing the compressed likelihood function in (9) with respect to τ. However, this optimization problem is not analytically tractable since the considered objective function is non-linear with respect to τ. Many alternative methods were developed to numerically solve this problem. Most of the approaches are iterative and require a good initial guess of the unknown parameter in order to converge to the global maximum. More so, the need of having a good initial guess usually requires the implementation of a suboptimal algorithm before applying the iterative ML technique. As far as we know, the CML-TED estimator [10] achieves the best performance among all the existing iterative estimators, but as it will be seen in Section VI, its performance depends strictly on its initial guess.

To overcome the drawback of iterative techniques, we proceed in this paper to an entirely different approach. In fact, for such problems, Pincus [11] showed that it is possible to obtain the global maximum that yields the maximum likelihood estimate of the desired parameter in a non-iterative way. Indeed, the theorem of Pincus states that the global maximum of the compressed likelihood function $L_c(\boldsymbol{y};\tau)$ is given by:

$$\widehat{\tau^*} = \lim_{\rho \to \infty} \int_J \tau L'_{c,\rho}(\tau) d\tau, \tag{10}$$

where

$$L'_{c,\rho}(\tau) = \frac{\exp\{\rho L_c(\boldsymbol{y};\tau)\}}{\int_J \exp\{\rho L_c(\boldsymbol{y};u)\} du}, \tag{11}$$

is the normalized function of $\exp\{\rho L_c(\boldsymbol{y};\tau)\}$. Note that the integration interval, J, in (10) and (11) is the interval in which τ^* is supposed to be confined. The function $L'_{c,\rho}(\tau)$ verifies all the properties of a pdf, but since τ is not random, $L'_{c,\rho}(\tau)$ can be seen as a pseudo-pdf. Moreover, we define the ML estimator for the time delay parameter as the one obtained from (10), for some large value of ρ_1 (that should be adequately chosen), as follows:

$$\widehat{\tau^*} = \int_J \tau L'_{c,\rho_1}(\tau) d\tau. \tag{12}$$

In practice, it is also difficult to implement the integral in (12). However, this integral can be viewed as the mean value of a random variable distributed according to the pseudo-pdf $L'_{c,\rho_1}(.)$. In fact, as shown in [12], this type of integral can be approximated by Monte-Carlo techniques as follows:

$$\widehat{\tau^*} = \frac{1}{R}\sum_{k=1}^{R} \tau_k, \tag{13}$$

where $\{\tau_k\}_{k=1}^R$ are realizations of τ distributed according to the pseudo-pdf, $L'_{c,\rho_1}(\tau)$, and R is the number of generated realizations. In this way, the ML estimator can be implemented as a simple averaging which, as $R \to \infty$, converges to the global maximum of $L'_{c,\rho_1}(\tau)$. But, it is not easy to generate realizations using $L'_{c,\rho_1}(\tau)$, since it is a non linear function of τ. Thus, we use the importance sampling technique which was succefully applied to the estimation of the carrier frequency [6], the DOAs [7] and the joint DOA-Doppler parameters [8].

IV. The Importance Sampling Technique

It has been recognized that Importance Sampling (IS) is a powerful tool to compute multiple integrals; in particular the one given in (12). This approach is based on the following simple observation that:

$$\int_J f(\tau) L'_{c,\rho_1}(\tau) d\tau = \int_J f(\tau) \frac{L'_{c,\rho_1}(\tau)}{g'(\tau)} g'(\tau) d\tau, \tag{14}$$

where $g'(.)$ is the normalized importance function which can also be seen as a pseudo-pdf and $f(\tau)$ is a given function.

Then, the right-hand-side of (14) can also be expressed as a summation according to Monte-Carlo methods as follows[1]:

$$\int_J f(\tau)L'_{c,\rho_1}(\tau)d\tau = \frac{1}{R}\sum_{i=1}^{R} f(\tau_i)\frac{L'_{c,\rho_1}(\tau_i)}{g'(\tau_i)}, \quad (15)$$

where τ_i is the ith realization of τ according to the new normalized importance function $g'(.)$. In our case, note that we simply have $f(\tau) = \tau$. It is of interest to choose $g'(.)$ similar to $L'_{c,\rho_1}(.)$ in order to reduce the variance of the estimates. In counter part, $g'(.)$ should be as simple as possible to guarantee easy generation of the realizations. Therefore, we need to choose an appropriate importance function $g'(.)$. In fact, we notice that the inverse matrix $(\boldsymbol{A}_\tau^T\boldsymbol{A}_\tau)^{-1}$ makes the actual compressed likelihood function $L_c(\boldsymbol{y};\tau)$, and therefore, the pseudo-pdf $L'_{c,\rho_1}(.)$ very difficult to evaluate. Hence, one can reasonably replace $(\boldsymbol{A}_\tau^T\boldsymbol{A}_\tau)^{-1}$ by the diagonal matrix $\frac{Ts}{E_s}\boldsymbol{I}_K$ and obtain the following approximation of the actual compressed likelihood function:

$$L_c(\boldsymbol{y};\tau) \approx \frac{T_s}{E_s}\boldsymbol{y}^H\boldsymbol{A}_\tau\boldsymbol{A}_\tau^T\boldsymbol{y}. \quad (16)$$

The approximation of $(\boldsymbol{A}_\tau^T\boldsymbol{A}_\tau)^{-1}$ with $\frac{T_s}{E_s}\boldsymbol{I}_K$ remains valid for most of the conventional shaping functions. In particular, it can be easily verified that for the widely used root raised-cosine shaping pulse, the off-diagonal elements of $(\boldsymbol{A}_\tau^T\boldsymbol{A}_\tau)^{-1}$ are negligible compared to its diagonal elements.
Then, the following importance function can be considered:

$$\begin{aligned} g(\tau) &= \exp\left\{\rho_1\frac{T_s}{E_s}\boldsymbol{y}^H\boldsymbol{A}_\tau\boldsymbol{A}_\tau^T\boldsymbol{y}\right\} \\ &= \prod_{k=0}^{K-1}\exp\left\{\rho_1\frac{T_s}{E_s}I_k(\tau)\right\}, \quad (17) \end{aligned}$$

where for $k = 0, 1, 2, \ldots, K-1$, we have

$$I_k(\tau) = \left|\sum_{i=0}^{M-1} y^*(iT_s)h(iT_s - kT - \tau)\right|^2. \quad (18)$$

Note that the normalization of $g(.)$ yields the normalized importance function[2] $g'(.)$ [i.e., $g'(\tau) = \frac{g(\tau)}{\int_J g(x)dx}$] which will be used in the generation of the useful realizations. To simplify the notations, we will henceforth use $\rho'_1 = \frac{T_s}{E_s}\rho_1$ whose value strictly affects the performance of the new ML estimator. In fact, recall that our ultimate objective is to find the global maximum of $L_c(\tau)$. However, this function exhibits many local maxima and it is difficult to distinguish the global maximum from a local one. This is exactly the role of the parameter ρ'_1 which, as it increases, makes the function of interest more peaked around its global maximum compared to the local maxima. We can verify this behavior in Fig. 1 which plots the function $g'(\tau)$ for $\rho'_1 = 10$ and $\rho'_1 = 20$. However, very high values of ρ'_1 may result in an overflow in

[1]To do so, we generate u_i uniformly distributed in $[0, 1]$ and we use a golden search to find τ_i that satisfies $\tau_i = G^{-1}(u_i)$, where $G(.)$ is the cumulative distribution function of $g(.)$ (i.e., $G(u_i) = \int_0^{u_i} g'(u)du$).

[2]For multiple parameters estimation, ρ_1 in $g(.)$ is different from the constant in $L'_{c,\rho_1}(.)$. In fact, if ρ_1 is too large, then the importance function may become too narrow. As a result, not all possible parameters will be generated [7].

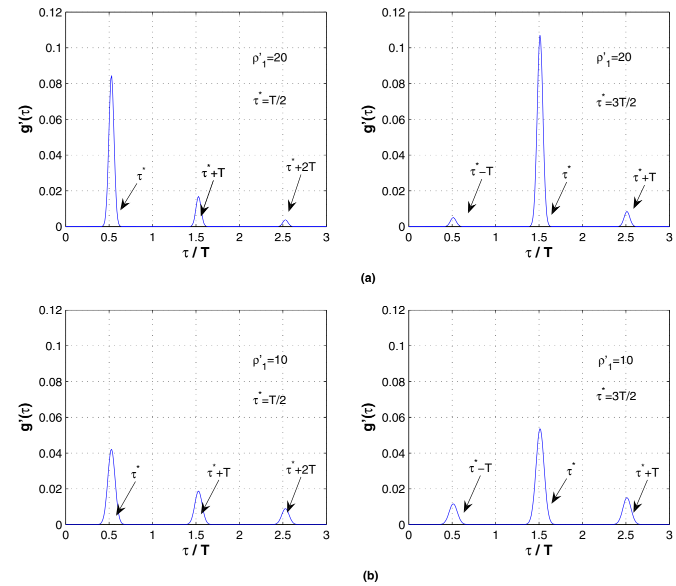

Fig. 1. Plot of $g'(\tau)$ for $\rho'_1 = 20$ and $\rho'_1 = 10$ using a root-raised cosine pulse and for $K = 100$, (a) $\tau^* = T/2$, (b) $\tau^* = 3T/2$.

the computation of $g'(\tau)$ and an optimal choice of ρ'_1 is also of extreme importance as we will see in Section VI. We note also that the matrix-by-vector product operation $\boldsymbol{A}_\tau^T\boldsymbol{y}$ in (17) can be viewed as a simple filtering operation of the received samples $\boldsymbol{y}$ with a filter whose coefficients are the central row of $\boldsymbol{A}_\tau^T$. Therefore, the computation of $g(\tau)$ is quite simple and realizations distributed according to $g'(.)$ can be easier generated than according to the actual normalized $L'_{c,\rho_1}(.)$ in (11).

V. Estimation of the Time Delay

We consider here two scenarios. In the first scenario, the time delay parameter is assumed to take values within[3] $[0, PT]$, where P is a strictly positive integer superior or equal to one. In fact, in many applications such as radar or sonar transmissions, the actual time delay introduced by the channel may exceed the symbol's duration. In the second scenario, we assume that the time delay parameter does not exceed the symbol's duration.

A. First scenario: $\tau^* \in [0, PT]$

First, we mention that the maxima of $g'(.)$, illustrated in Fig. 1, are periodic, with period T. Therefore, even in the total absence of noise, many secondary peaks appear and affect the estimate $\widehat{\tau^*}$. To illustrate the effect of the secondary peaks, suppose that $g'(.)$ has only two maxima located at τ^* and $T + \tau^*$. Then the realizations take values around τ^* as well as around $T + \tau^*$. Since the estimate $\widehat{\tau^*}$ is the mean of all these realizations, then $\widehat{\tau^*}$ will be shifted toward $\tau^* + T$ and does not therefore reflect the real value of the time delay. To circumvent this problem, $g'(.)$ should be made quasi-symmetric around

[3]This limitation of the interval will not reduce the applicability of the new estimator since we always have an *a priori* idea about the range of τ^*.

τ^*. To that end, a simple solution is to suppose that τ^* takes virtually some negative values. Other peaks will appear before τ^* as well, rendering therefore the estimates more accurate. Hence, we extend the interval from $[0, PT]$ to $[-QT, PT]$, where Q is another positive integer smaller than P. This makes other virtual lobes appear before τ^* as well, thereby compensating for the secondary lobes located after τ^*. In fact, using the new extended interval, we see from Fig. 2 that the probability of generating realizations around $\tau^* + iT$ is the same as the one of generating realizations around $\tau^* - iT$. As a result, the estimates become unbiased and more accurate. At this stage, our new algorithm works correctly for

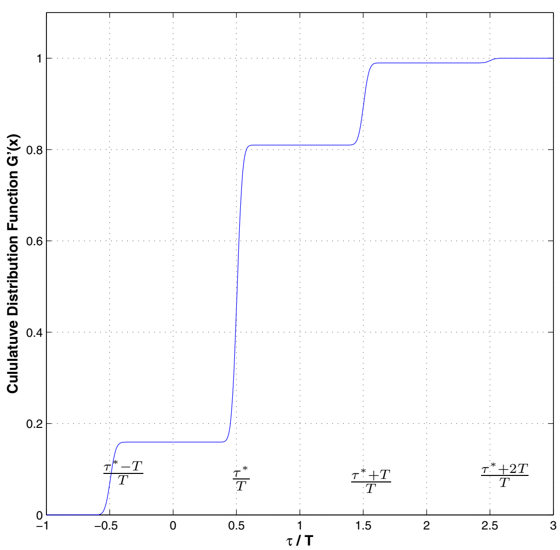

Fig. 2. Plot of the cumulative distribution function (CDF), $G'(\tau)$ whose pdf is $g'(\tau)$, SNR = 5 dB.

constant-envelope constellations. But, we have noticed that its performance degrades for non-constant envelope constellations such as 16-QAM, 64-QAM, etc, a problem that we will cope with subsequently. Consider that the first symbol[4] is of a minimal amplitude and the second one is of a maximal amplitude among all the constellation points, then the value of $g'(\tau^*)$ and $g'(\tau^* + T)$ will be almost the same. However, to have accurate estimates, $g'(\tau^*)$ should be much higher than $g'(\tau^* + T)$. Indeed, $g'(\tau^*)$ can be written as follows:

$$g'(\tau^*) = g'(\tau^* + T)\ \exp\{\rho_1' I_0(\tau^*)\}, \tag{19}$$

where $I_0(\tau^*)$ depends on the amplitude of the first symbol. If its amplitude is minimal, the term $\exp\{\rho_1' I_0(\tau^*)\}$ can be neglected with respect to the other terms, and $g'(\tau^*)$ will be closer to $g'(\tau^* + T)$, which results in a local maximum. To avoid these problems, we can force the first and the last symbols to be of maximum amplitude, so that the local maxima remain very small compared to the global maximum in terms of amplitude.

[4]The same problem occurs with the last received symbol.

B. Second scenario: $\tau^* \in [0, T]$

In many practical cases, the time delay does not exceed the symbol period T. Hence, we can look for the global maximum within $[0, T]$. As previously explained, the maxima of the importance function are periodically located, with a period equal to T. Moreover, since we know *a priori* that τ^* does not exceed T, then we can more conveniently use the circular mean instead of the linear mean[5] in (15). In our case, the time delay is transposed into $[0, 1]$ by normalizing τ^* by T. Then, we have to inverse the resulting transposed estimate to obtain an estimate in the original interval. Therefore, the IS formulation considering the circular mean is:

$$\widehat{\tau^*} = \frac{1}{2\pi T} \angle \frac{1}{R} \sum_{i=1}^{R} \frac{L'_{c,\rho_1}(\tau_i)}{g'(\tau_i)} \exp\left\{ j \frac{2\pi \tau_i}{T} \right\}, \tag{20}$$

or finally:

$$\widehat{\tau^*} = \frac{1}{2\pi T} \angle \frac{1}{R} \sum_{i=1}^{R} F(\tau_i) \exp\left\{ j \frac{2\pi \tau_i}{T} \right\}. \tag{21}$$

where

$$F(x) = \frac{L'_{c,\rho_1}(x)}{g'(x)}. \tag{22}$$

Some simplifications can be considered to make the computational load more manageable. First, we need to determine the angle of the expression in (21), thus it is reasonable to remove the strictly positive normalizing factors in $L'_{c,\rho_1}(.)$ and $g'(.)$ without affecting the final result. Moreover, we can multiply $F(x)$ by a positive scalar to avoid computation overflows as follows[6]:

$$\begin{aligned} F'(\tau_i) \;=\; & \exp\left\{ \rho_1 L_c(\tau_i) - \rho_1' \sum_{k=0}^{K-1} I_k(\tau_i) - \right. \\ & \left. \max_{1 \le l \le R} \left(\rho_1 L_c(\tau_l) - \rho_1' \sum_{k=0}^{K-1} I_k(\tau_l) \right) \right\} \end{aligned} \tag{23}$$

C. Summary of steps

1) Based on the sampled data $y(iT)$, $i = 0,\ 1, \ldots,\ M-1$, evaluate the periodogram $I_k(\tau)$ according to (18).
2) Compute the normalized importance function.
3) Generate R realizations of the parameter $\{(\tau_i)\}_{i=1}^{R}$ using the inverse probability integration.
4) Evaluate the weight coefficient $F(\tau_i)$ defined in (22) (or $F'(\tau_i)$ defined in (23) if we consider that τ is in $[0, T]$) for each generated value τ_i.
5) Compute the mean of the generated variables multiplied by the weight coefficients to find the ML estimate.

Note that most of the complexity burden is focused on how to generate the realizations. Therefore the computation complexity increases with R and a tradeoff between complexity and performance must be made.

[5]Note that the circular mean can not be used in the first senario when τ^* may exceed T since it always returns an estimate in $[0, T]$ by virtually bringing, into this interval, all the secondary lobes of the normalized imoprtance function.

[6]Note that the same simplifications have been used in [7] to estimate the signal DOA.

VI. Simulation Results

In this section, numerical results are presented to substantiate the performance of the new ML estimator as a function of the SNR. The normalized (by T^2) mean square error (NMSE) will be used as a performance measure. A number of 500 realizations (i.e., $R = 500$) provides good results. First, we plot in Fig. 3 the performance of the estimator as a function of ρ_1'. As it was expected, the mean square error decreases as ρ_1' increases, but starting from a certain value, the performance deteriorates considerably due to numerical overflows. Next, through simulations, we verified that, for a root raised-cosine filter, the ratio $L'_{c,\rho_1}(\tau)/g'(\tau)$ is very close to 1. Then to reduce the computational complexity, this ratio is set to 1. As shown in Fig. 4, this simplification does not degrade the performance of the estimator since the NMSE is almost the same. Moreover, we compare our algorithm to the tracking performance of the iterative CML-TED presented in [10]. As far as we know, among all the existing synchronization techniques, the CML-TED algorithm achieves the best performance which depends, however, on its initial guess. To illustrate this point, we consider two initial values of τ^* for the CML-TED. QPSK modulation is used and the roll-off factor is set to 0.5. Fig. 5 represents the performance of our IS-based estimator and the CML-TED for two different initializations. With a good initialization, i.e., $|\widehat{\tau_0^*} - \tau^*| = T/10$, with τ^* being the true time delay value to be estimated and $\widehat{\tau_0^*}$ the initial guess, we can see that our IS-based estimator outperforms the CML-TED in the high SNR region. However, in low SNR region, the CML-TED performs better than our IS-based estimator. The CML-TDE is able to converge to the true time delay value over the entire SNR region while the IS-based depends on the variance of the noise. However, if we choose $\widehat{\tau_0^*}$ such that $|\widehat{\tau_0^*} - \tau^*| = T/2$, the performance of the CML-TED is no longer reliable over the entire SNR range. This illustrates the fact that an iterative algorithm fails to estimate the time delay parameter if the initialization is not appropriately made. In contrast, our IS-based estimator does not require any initialization and hence does not suffer from this major drawback.

The second variant of the algorithm, namely considering the parameter as a circular variable, is also depicted in Fig. 5. We see that the variance error is reduced, especially in the low SNR region, and that the proposed IS-based technique outperforms the CML-TED starting from an SNR value of about 3 dB. Moreover, in contrast to our new ML estimator, we mention that the CML-TED suffers from an increasing variance penalty as the roll-off factor decreases. In fact, as shown in Fig .6, the performance of the CML-TED does not approach asymptotically the MCRLB in contrast to our IS-based algorithm which always reaches the MCRLB in the high SNR regime. Fig. 7 depicts the NMSE of our algorithm for a 16-QAM constellation. As explained in section V, the algorithm fails to estimate the time delay for non-constant-

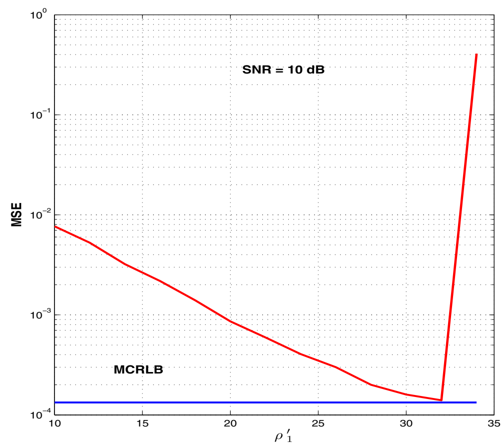

Fig. 3. Performance versus ρ_1' for SNR=10 dB.

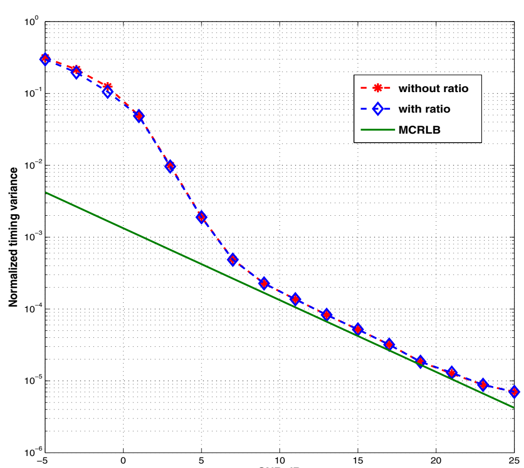

Fig. 4. Estimation performance considering $L'_{c,\rho_1}(\tau)/g'(\tau)$ and setting $L'_{c,\rho_1}(\tau)/g'(\tau)$ to 1, for a roll-off factor of 0.5.

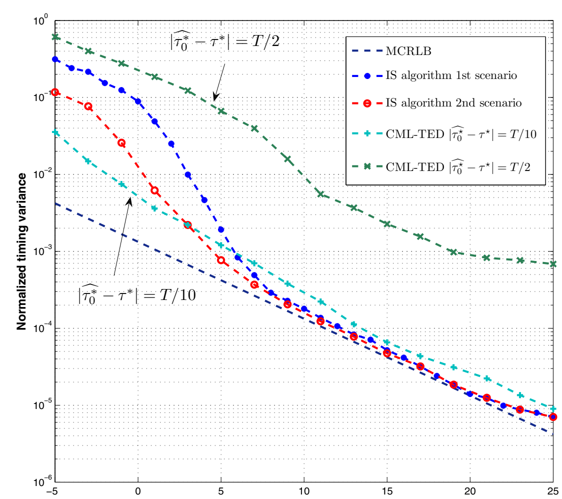

Fig. 5. Comparison between the estimation performance of the IS-based algorithm for the two scenarios and the tracking performance of the CML-TED using QPSK.

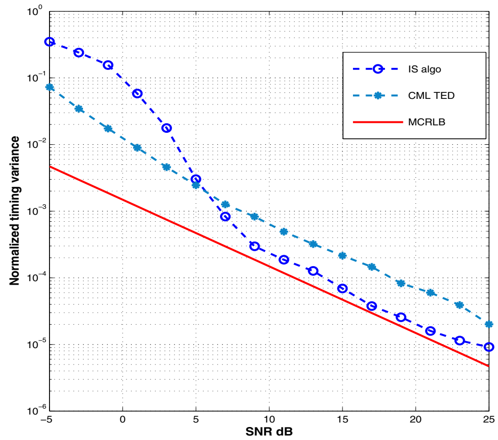

Fig. 6. Comparison between the estimation performance of the IS-based algorithm and the tracking performance of CML-TED using QPSK and a roll-off factor of 0.2.

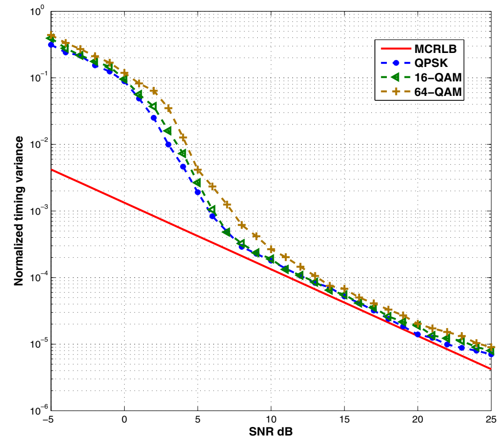

Fig. 8. Normalized MSE of the time delay estimate for different QAM modulation orders, using a root-raised cosine filter with a roll-off factor of 0.5.

envelope modulated signals without forcing the first and the last transmitted symbols to have a maximum energy. However, by performing the maximum-energy forcing technique over the first and the last transmitted symbols, the resulting performance is noticeably improved. In fact, as we see from Fig. 8, the NMSE for QPSK and higher-order QAM modulations (16-QAM, 64-QAM) holds almost the same, with a small performance improvement for the QPSK modulation.

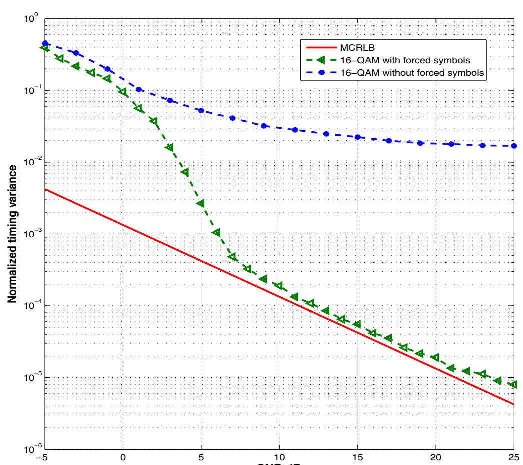

Fig. 7. Comparison of the estimation performance with and without forced symbols using 16-QAM modulation and a roll-off factor of 0.5.

VII. Conclution

In this work, we have proposed a computationally moderate technique to implement the CML estimator of the time delay parameter based on the importance sampling procedure. The CML formulation has been adopted on the basis of a discrete-time signal model. We showed that the global maximum of the compressed likelihood function can be easily found without any initial parameter guess. Moreover, iterative steps are avoided, which guarantees the convergence to the global maximum. These two points illustrate the main advantages of the IS-based algorithm compared to classical iterative approaches. Moreover, the new algorithm performs very close to the CRLB over a wide range of practical SNR values.

References

[1] U. Mengali and A. N. D'Andrea, *Sychronization Techniques for Digital Receivers,* New York: Plenum, 1997.
[2] M.I. Skolnik, *Introduction to Radar Systems*, third edition, McGraw-Hill, New York: 2001.
[3] G.C. Carter, *Coherence and Time Delay Estimation: An Applied Tutorial for Research, Development, Test and Evaluation Engineers*, IEEE Press, New York, 1993.
[4] J. W. M. Bergmans, *Digital Baseband Transmission and Recording*, Boston: Kluwer Academic Publishers, 1996.
[5] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory,* Prentice Hall, Englewood Cliffs, NJ, 1993.
[6] S. Saha and S. Kay, "Mean Likelihood Frequency Estimation," *IEEE Trans. Sig. Process.*, vol. 48, pp. 1937-1946, July 2000.
[7] S. Saha and S. Kay, "An Exact Maximum Likelihood Narrowband Direction of Arrival Estimator," *IEEE Trans. Sig. Process.*, vol. 56, pp. 1082-1092, Oct. 2008.
[8] H. Wang and S. Kay, "A Maximum Likelihood Angle-Doppler Estimator using Importance Sampling," *IEEE Trans. Aerospace and Electro. Syst.*, vol. 46,pp. 610-622, Apr. 2010.
[9] P. Stoica and A. Nehorai, "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation," *IEEE Trans. Acoust., Speech, Sig. Process.*, vol. 38, pp. 1783-1795, Oct. 1990.
[10] J. Riba, J. Sala and G. Vazquez, "Conditional Maximum Likelihood Timing Recovery: Estimators and Bounds," *IEEE Trans. on Sig. Process.*, vol. 49, pp. 835-850, Apr. 2001.
[11] M. Pincus, "A Closed Form Solution for Certain Programming Problems," *Oper. Research*, pp. 690-694, 1962.
[12] L. Stewart, "Bayesian Analysis using Monte-Carlo Integration: a Powerful Methodology for Handling Some Difficult Problems," *American Statistician*, pp. 195-200, 1983.
[13] K. V. Mardia, *Statistics of Directional Data*, New York: Academic, 1972.