# Maximum Likelihood Time Delay Estimation for Direct-Spread CDMA Multipath Transmissions Using Importance Sampling

Ahmed Masmoudi, Faouzi Bellili, and Sofiène Affes

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

***Abstract*—In this paper, we address the problem of time delay estimation for Direct-Spread CDMA (DS-CDMA) multipath transmissions. We observe that the attractive post-correlation model (PCM) of the despread data allows for exploiting results provided in array signal processing and we introduce a new time delay estimator based on the maximum likelihood (ML) criterion. The new technique finds the global maximum of the compressed likelihood function in an efficient way using the importance sampling (IS) technique. Simulation results show that the new estimator provides very good performance in challenging cases of very closely-spaced delays even with a single antenna and transmitted symbol. The Cramer-Rao lower bound (CRLB) for multipath delays is also provided.**

***Index Terms*—Timing synchronization, post-correlation model, optimization methods, Cramér-Rao lower bound (CRLB), maximum likelihood, importance sampling.**

## I. Introduction

Parameter estimation is a crucial operation for any digital receiver; in particular the recovery of time delays introduced by the effects of multipath propagation. In this work, we focus on the CDMA array-receiver that provides very accurate channel estimates with a low complexity cost [1] and the post-correlation model (PCM) of the despread data developed there. Actually, it is well known that incorrect multipath timing degrades considerably the performance of CDMA systems. Therefore, time delay estimation was investigated in [1] by applying the powerful Root-MUSIC algorithm and revised in [2] with a complexity reduction. In this paper, we avoid eigen-decomposition operation widely used in conventional high-resolution methods by adapting the maximum likelihood (ML) criterion to the PCM.

According to estimation theory, the ML estimator is asymptotically efficient for high signal to noise ratio (SNR) or large numbers of received samples. Its performance always outperforms other estimators especially for closely-spaced delays. Unfortunately, in our problem, a closed-form solution for the ML estimates is intractable and a direct maximization of the likelihood function requires a multi-dimensional grid search, whose complexity increases with the number of unknown parameters. Therefore, we resort to the concept of importance sampling (IS), as investigated in the case of frequency [3], DOA [4] and single-path time delay estimation [5].

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

It will be shown through simulations that the new method outperforms the classical Root-MUSIC method for moderate SNR values and presents noticeably better performance in the special case of a single receiving antenna.

The remainder of this paper is organized as follows. In section II, we briefly introduce the DS-CDMA model. In section III, we derive the compressed likelihood function that we aim to maximize. Then, in section IV, we introduce the importance sampling technique with application to our problem of time delay estimation for CDMA systems. The analytical expression of the corresponding CRLB will be derived in section V. Simulation results are discussed in Section VI and some concluding remarks are drawn out in Section VII.

## II. System Model and Background

We consider a CDMA transmission system where the receiver is equipped with $M$ antenna elements. The symbol duration is denoted by $T$ and the processing gain is defined as $L = T/T_c$, where $T_c$ is the chip pulse duration. We consider $P$ propagation paths in a Rayleigh fading channel where the largest delay is small compared to the symbol duration $T$. The post-correlation model of spatio-temporal observation of the $n^{th}$ received symbol on the antenna array is given by [1]:

$$\boldsymbol{Z}_n = \boldsymbol{G}_n \boldsymbol{\Upsilon}_n \boldsymbol{D}^T(\boldsymbol{\tau}) s_n + \boldsymbol{N}_n, \tag{1}$$

where $s_n = b_n \psi_n$ is the product of the transmitted symbol $b_n$ and the total received power $\psi_n^2$. The matrix $\boldsymbol{D}(\boldsymbol{\tau})$ is the time response matrix whose $p^{th}$ column is:

$$\mathbf{d}_p = [\rho_c(-\tau_p),\ \rho_c(T_c - \tau_p), \ldots,\ \rho_c((L-1)T_c - \tau_p)]^T, \tag{2}$$

and $\rho_c(.)$ is the correlation function of the spreading code. $\boldsymbol{G}_n$ is the $M \times P$ propagation matrix and $\boldsymbol{\Upsilon}_n$ is a $P \times P$ diagonal matrix representing the power partition over the different paths [1]. These two matrices can be gathered in one matrix $\boldsymbol{J}_n$ (i.e., $\boldsymbol{J}_n = \boldsymbol{G}_n \boldsymbol{\Upsilon}_n$). Then, a compact form of $\boldsymbol{Z}_n$ is given by:

$$\boldsymbol{Z}_n = \boldsymbol{H}_n s_n + \boldsymbol{N}_n \tag{3}$$

with $\boldsymbol{H}_n = \boldsymbol{J}_n \boldsymbol{D}^T(\boldsymbol{\tau})$ represents the spatio-temporal propagation matrix. Finally, $\boldsymbol{N}_n$ is the $M \times L$ noise matrix. The spatio-temporal model in (3) provides a powerful tool for estimating the channel response and the transmitted symbol. In fact, a transformation of $\boldsymbol{Z}_n$, $\boldsymbol{H}_n$ and $\boldsymbol{N}_n$, using the vect(.) operator[1], into $ML$-dimensional vectors yields:

$$\boldsymbol{z}_n = \boldsymbol{h}_n s_n + \boldsymbol{n}_n, \tag{4}$$

where $\boldsymbol{z}_n$, $\boldsymbol{h}_n$ and $\boldsymbol{n}_n$ are the resulting $ML \times 1$ vectors. At this stage, many blind channel estimators that were derived in the literature can be used to obtain an estimate of $\widehat{\boldsymbol{h}}_n$. In particular, we refer the reader to the channel estimator that was derived within the framework of STAR receiver in [1].

Then, once an estimate $\widehat{\boldsymbol{h}}_n$ is at hand, it will be used to estimate the time delays. Additionally, the PCM space time formulation of $\boldsymbol{H}_n$ offers an interesting structure since it gathers all the parameters of interest in one matrix $\boldsymbol{D}(\boldsymbol{\tau})$ (i.e., $\boldsymbol{H}_n = \boldsymbol{J}_n \boldsymbol{D}^T(\boldsymbol{\tau})$). This formulation will be exploited to develop a ML time delay estimator.

## III. Likelihood Function

Taking into account the estimation error of $\boldsymbol{h}_n$, the estimate, $\widehat{\boldsymbol{H}}_n$, of the spatio-temporal propagation matrix, $\boldsymbol{H}_n$, is given by:

$$\widehat{\boldsymbol{H}}_n^T = \boldsymbol{D}(\boldsymbol{\tau})\boldsymbol{J}_n^T + \boldsymbol{E}_n^T, \tag{5}$$

where $\boldsymbol{E}_n^T$ is the corresponding error matrix whose elements' variance can be expressed as a function of the noise variance in the received signal [6]. According to this representation, the problem can be thought of as the estimation of the unknown delays from $M$ symbols received by $L$ antenna elements. The observation vectors are the columns of $\boldsymbol{H}_n$ and the columns of $\boldsymbol{J}_n$ represent the transmitted signals from $P$ different sources. Now, using the fact that a delay in the time domain becomes a phase shift in the frequency domain, we perform a column by column fast Fourier transform (FFT) of $\widehat{\boldsymbol{H}}_n^T$ and we obtain:

$$\widehat{\boldsymbol{\mathcal{H}}}_n = \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{J}_n^T + \boldsymbol{\mathcal{E}}_n, \tag{6}$$

in which the matrix $\boldsymbol{\mathcal{E}}_n$ is the resulting noise matrix and $\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})$ depends only on the unknown delays and is given by:

$$\boldsymbol{\mathcal{D}}(\boldsymbol{\tau}) = [\boldsymbol{d}(\tau_1),\ \boldsymbol{d}(\tau_2), \ldots,\ \boldsymbol{d}(\tau_P)], \tag{7}$$

whose columns $\{\boldsymbol{d}(\tau_p)\}_{p=1}^P$ are defined as:

$$\boldsymbol{d}(\tau_p) = [c_0,\ c_1 e^{-\frac{j2\pi\tau_p}{L}}, \ldots,\ c_{L-1}e^{-\frac{j2\pi(L-1)\tau_p}{L}}]^T, \tag{8}$$

and $\boldsymbol{c} = [c_0,\ c_1, \ldots,\ c_{L-1}]^T$ is the $(L \times 1)$ vector containing the FFT coefficients of the correlation function of the spreading code. In practice, this function has a very narrow lobe, which makes the FFT coefficients quasi-constant in amplitude. Now, suppose that the columns of $\widehat{\boldsymbol{\mathcal{H}}}_n$, denoted $\widehat{\boldsymbol{\mathcal{H}}}_n^{(i)}$ for $i = 1, \ldots, M$, are independent and the columns of the error matrix $\boldsymbol{\mathcal{E}}_n$ are Gaussian. Each column of $\widehat{\boldsymbol{\mathcal{H}}}_n$ is hence distributed according to a complex Gaussian probability density function (pdf) with zero mean and a covariance matrix $\boldsymbol{R} = \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{P}\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau}) + \sigma^2 \boldsymbol{I}_L$, where $\boldsymbol{I}_L$ is an $(L \times L)$ identity matrix and $\boldsymbol{P}$ is the covariance matrix of the columns of $\boldsymbol{J}_n^T$, supposed to be the same for all the columns. Consequently, the pdf of $\widehat{\boldsymbol{\mathcal{H}}}_n$, parameterized by $\boldsymbol{\tau}$ and $\boldsymbol{P}$, is given by:

$$p(\widehat{\boldsymbol{\mathcal{H}}}_n; \boldsymbol{\tau}, \boldsymbol{P}) = \frac{1}{\pi^{ML}} \frac{1}{\Big(\det(\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{P}\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau}) + \sigma^2\boldsymbol{I}_L)\Big)^M}$$
$$\exp\left\{-\sum_{i=1}^M (\widehat{\boldsymbol{\mathcal{H}}}_n^{(i)})^H \Big(\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{P}\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau}) + \sigma^2\boldsymbol{I}_L\Big)^{-1} \widehat{\boldsymbol{\mathcal{H}}}_n^{(i)}\right\}, \tag{9}$$

where $\det(.)$ refers to the determinant of a matrix and $\boldsymbol{\tau} = [\tau_1,\ \tau_2, \ldots,\ \tau_P]^T$ are the unknown delays to be estimated. The resulting log-likelihood function $L(\boldsymbol{\tau}, \boldsymbol{P})$, after omitting the constant terms, is given by:

$$\begin{aligned} L(\boldsymbol{\tau}, \boldsymbol{P}) &= -\log\Big\{\det(\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{P}\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau}) + \sigma^2\boldsymbol{I}_L)\Big\} - \\ &\quad \frac{1}{M}\sum_{i=1}^M \widehat{\boldsymbol{\mathcal{H}}}_n^{(i)} \Big(\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{P}\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau}) + \sigma^2\boldsymbol{I}_L\Big)^{-1} \widehat{\boldsymbol{\mathcal{H}}}_n^{(i)} \\ &= -\log\{\det(\boldsymbol{R})\} - \text{trace}\Big\{\boldsymbol{R}^{-1}\widehat{\boldsymbol{R}}\Big\}, \end{aligned} \tag{10}$$

with $\widehat{\boldsymbol{R}}$ beeing the estimate of the sample covariance matrix of the columns of $\boldsymbol{\mathcal{H}}_n$ computed from $\widehat{\boldsymbol{\mathcal{H}}}_n$ ($\widehat{\boldsymbol{R}} = \frac{1}{M}\sum_{i=1}^M \widehat{\boldsymbol{\mathcal{H}}}_n^{(i)}(\widehat{\boldsymbol{\mathcal{H}}}_n^{(i)})^H$). The log-likelihood function in (10) depends on the parameter vector of interest $\boldsymbol{\tau}$ and the covariance matrix $\boldsymbol{P}$. Therefore, we need to maximize this function over $\boldsymbol{\tau}$ and $\boldsymbol{P}$. To that end, it can be shown, using similar developments used in the context of DOA estimation [7], that the value of $\boldsymbol{P}$ that maximizes (10) for a given value of $\boldsymbol{\tau}$ is:

$$\widehat{\boldsymbol{P}}_{ML} = \Big(\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\Big)^{-1} \boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\widehat{\boldsymbol{R}}\boldsymbol{\mathcal{D}}(\boldsymbol{\tau}) \Big(\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\Big)^{-1}$$
$$\sigma^2 \Big(\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\Big)^{-1}. \tag{11}$$

Injecting $\widehat{\boldsymbol{P}}_{ML}$ in the expression of the log-likelihood function in (10), we obtain the so-called compressed likelihood function that depends only on $\boldsymbol{\tau}$:

$$L_c(\boldsymbol{\tau}) = \frac{1}{\sigma^2}\text{trace}\Big(\boldsymbol{\Pi}\widehat{\boldsymbol{R}}\Big) - \log\Big(\det\Big(\boldsymbol{\Pi}\widehat{\boldsymbol{R}}\boldsymbol{\Pi} + \sigma^2(\boldsymbol{I}_L - \boldsymbol{\Pi})\Big)\Big), \tag{12}$$

where $\boldsymbol{\Pi}$ is an orthogonal projector matrix ($\boldsymbol{\Pi} = \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\Big(\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\Big)^{-1}\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})$). While $\sigma^2$ is usually unknown, it can be easily estimated either by averaging the $L - M$ smallest eigen-values of $\widehat{\boldsymbol{R}}$ or by exploiting the estimated power carried out in a previous stage of STAR [1].

## IV. The Importance Sampling Technique

The maximum likelihood estimates of the delays are obtained by maximizing the compressed likelihood function $L_c(\boldsymbol{\tau})$ with respect to $\boldsymbol{\tau}$. The most obvious method consists in performing a $P$-dimensional grid search, whose complexity increases with the number of paths. And since a closed-form expression of the solution is analytically intractable, we adopt in this paper an entirely different technique. In fact, based on the theorem of Pincus [8], the global maximum of $L_c(\boldsymbol{\tau})$ is given by:

[1]The vect(.) operator arranges the columns of a matrix as one column vector.

$$\widehat{\tau}_p = \lim_{\rho\to\infty} \frac{\int_J \dots \int_J \tau_p \exp\{\rho L_c(\boldsymbol{\tau})\}\, d\boldsymbol{\tau}}{\int_J \dots \int_J \exp\{\rho L_c(\boldsymbol{u})\}\, d\boldsymbol{u}}, \tag{13}$$

where $J = [0, T]$ is the interval in which the delays are confined. Then, if we define the pseudo-pdf, for a sufficiently large value $\rho_0$, as follows:

$$L'_{c,\rho_0}(\boldsymbol{\tau}) = \frac{\exp\{\rho_0 L_c(\boldsymbol{\tau})\}}{\int_J \dots \int_J \exp\{\rho_0 L_c(\boldsymbol{u})\}\, d\boldsymbol{u}}, \tag{14}$$

then the optimal value of $\widehat{\tau}_i$ is given by:

$$\widehat{\tau}_p = \int_J \dots \int_J \tau_p L'_{c,\rho_0}(\boldsymbol{\tau}) d\boldsymbol{\tau},\ p = 1,\ 2,\ \dots,\ P. \tag{15}$$

From the expression of $L'_{c,\rho_0}(.)$, it can be seen that as $\rho_0$ tends to infinity, $L'_{c,\rho_0}(.)$ tends to a $P$-dimensional Dirac function centered at the location of the global maximum of $L_c(.)$. In fact, the use of the exponential operator makes the value of $L'_{c,\rho_0}(.)$ at the maximum of $L_c(.)$ increases faster than the other values. Therefore, the estimated values $\{\widehat{\tau}_p\}_{p=1}^P$ using (15) are indeed the desired maximum likelihood estimates of the delays. However, the implementation of the estimator in (15) requires the computation of a multidimensional integral, which is usually difficult to perform. Yet, $\widehat{\tau}_p$ can be seen as the expected value of $\tau_p$, when the vector $\boldsymbol{\tau}$ is distributed according to the pseudo-pdf[2] $L'_{c,\rho_0}(.)$. In this case, an estimate of the mean can be simply found using Monte-Carlo methods as follows:

$$\widehat{\boldsymbol{\tau}} = \frac{1}{R}\sum_{k=1}^{R} \boldsymbol{\tau}_k, \tag{16}$$

where $\{\boldsymbol{\tau}_k\}_{k=1}^R$ are $R$ realizations of $\boldsymbol{\tau}$ generated according to $L'_{c,\rho_0}(.)$. For the problem at hand, the pseudo-pdf, $L'_{c,\rho_0}(.)$, depends on the actual compressed likelihood function, $L_c(.)$, which is a complex multidimensional function. Thus generating realizations according to $L'_{c,\rho_0}(.)$ as defined in (15) is a difficult task and hence it will be more interesting to find a simple one-dimensional function and use it instead of $L'_{c,\rho_0}(.)$. To that end, we resort to the concept of importance sampling (IS) as detailed below [9].

The principle of the IS technique is based on the following simple observation:

$$\int_J \dots \int_J f(\boldsymbol{\tau}) L'_{c,\rho_0}(\boldsymbol{\tau}) d\boldsymbol{\tau} = \int_J \dots \int_J f(\boldsymbol{\tau}) \frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})} g'(\boldsymbol{\tau}) d\boldsymbol{\tau}, \tag{17}$$

for any function $f(.)$ and where $g'(.)$ is another pseudo-pdf called normalized importance function (IF). Using (17), the problem can be recast as the computation of the mean of $f(\boldsymbol{\tau})\frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})}$ when $\boldsymbol{\tau}$ is generated according to $g'(\boldsymbol{\tau})$ instead of $L'_{c,\rho_0}(.)$. Then, we use again the Monte-Carlo method to compute the left-hand side of (17) as follows:

$$\int_J \dots \int_J f(\boldsymbol{\tau}) \frac{L'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})} g'(\boldsymbol{\tau}) d\boldsymbol{\tau} = \frac{1}{R}\sum_{k=1}^{R} f(\boldsymbol{\tau}_k) \frac{L'_{c,\rho_0}(\boldsymbol{\tau}_k)}{g'(\boldsymbol{\tau}_k)}, \tag{18}$$

where the $R$ vectors $\{\boldsymbol{\tau}_p\}_{p=1}^R$ are now generated according to $g'(.)$. Clearly, the estimation performance depends on the choice of $R$ and $g'(.)$. Indeed, a large value of $R$ reduces the estimation variance but increases the computation complexity; although an appropriate choice of $g'(.)$ can reduce the required number of realizations $R$. In fact, $g'(.)$ needs to be similar to $L'_{c,\rho_0}(.)$ and simple enough so that realizations according to $g'(.)$ are easily generated. Clearly, there are two contradictory conditions for the choice of $g'(.)$ which should be satisfied. We discuss this point in the next section.

## V. Time Delay Estimator

From (12), we see that the compressed likelihood function involves the sum of two terms. These two terms depend on the $P$ positive eigenvalues $\lambda_1,\ \lambda_2, \dots,\ \lambda_P$ of the matrix $\boldsymbol{\Pi}\widehat{\boldsymbol{R}}\boldsymbol{\Pi}$ as follows:

$$\begin{aligned}\log\Big(\det\Big(\boldsymbol{\Pi}\widehat{\boldsymbol{R}}\boldsymbol{\Pi} + \sigma^2(\boldsymbol{I}_L - \boldsymbol{\Pi})\Big)\Big) &= \log\left((\sigma^2)^{L-P}\prod_{p=1}^{P}\lambda_p\right)\\ &= \sum_{p=1}^{P}\log\left(\frac{\lambda_p}{\sigma^2}\right) + L\log\sigma^2,\end{aligned} \tag{19}$$

and

$$\frac{1}{\sigma^2}\text{trace}\left(\boldsymbol{\Pi}\widehat{\boldsymbol{R}}\right) = \frac{1}{\sigma^2}\text{trace}\left(\boldsymbol{\Pi}\widehat{\boldsymbol{R}}\boldsymbol{\Pi}\right) = \sum_{p=1}^{P}\frac{\lambda_p}{\sigma^2}. \tag{20}$$

Clearly, the term $\sum_{p=1}^{P}\frac{\lambda_p}{\sigma^2}$ is dominant compared to $\sum_{p=1}^{P}\log\left(\frac{\lambda_p}{\sigma^2}\right)$. Therefore, we can neglect the second term in the compressed likelihood function and keep only the first one in the formulation of the IF $g'(.)$. Moreover, to make $g'(.)$ separable with respect to the delays, the inverse matrix $\left(\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\right)^{-1}$ is replaced by the diagonal matrix $\left(\sum_{l=0}^{L-1}|c_l|^2\right)^{-1}\boldsymbol{I}_P$. Note that this approximation is very accurate since the diagonal elements of the matrix $\left(\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\right)^{-1}$ are dominant compared to its off-diagonals ones[3]. Finally, using all these observations, we obtain the following approximation for the actual compressed likelihood function after omitting the constant term $\left(\sum_{l=0}^{L-1}|c_l|^2\right)^{-1}$:

$$\begin{aligned}&\frac{1}{\sigma^2}\ \text{trace}\left(\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{\mathcal{D}}^H(\boldsymbol{\tau})\widehat{\boldsymbol{R}}\right)\\ &\quad= \frac{1}{M\sigma^2}\sum_{i=1}^{M}\sum_{p=1}^{P}\left|\sum_{l=1}^{L} c_{l-1}\exp\left\{-\frac{j2\pi(l-1)\tau_p}{L}\right\}\widehat{\boldsymbol{\mathcal{H}}}_n(i,l)\right|^2\\ &\quad= \frac{1}{M\sigma^2}\sum_{p=1}^{P} I(\tau_p),\end{aligned} \tag{21}$$

[2]We use the term "pseudo-pdf" because $L'_{c,\rho_0}(.)$ has all the properties of a pdf but $\boldsymbol{\tau}$ is not a random variable.

[3]We mention that justifications of this approximation were omitted due to lack of space.

where

$$I(\tau)=\sum_{i=1}^{M}\left|\sum_{l=1}^{L}c_{l-1}\exp\left\{-\frac{j2\pi(l-1)\tau}{L}\right\}\widehat{\boldsymbol{\mathcal{H}}}_n(i,l)\right|^2. \tag{22}$$

Note that we keep the term $1/(M\sigma^2)$ since it depends on the experience's conditions (the SNR and the number of antennas) while $\left(\sum_{k=0}^{L-1}c_l\right)^{-1}$ is always constant.

Hence, the intended normalized importance function (IF) is given by:

$$g'_{\rho_1}(\boldsymbol{\tau})=\frac{\prod_{p=1}^{P}\exp\{\rho_1 I(\tau_p)\}}{\left(\int_J\exp\{\rho_1 I(u)\}\,du\right)^P}, \tag{23}$$

where $\rho_1$ is another constant[4] different from $\rho_0$. Note that it is more interesting to use two different constants $\rho_0$ and $\rho_1$. In fact, the normalized IF is built upon an approximation of the actual compressed likelihood function and this results in a biased estimation of the delays. This bias is alleviated by the weighting factor $L'_{c,\rho_0}(.)/g'(.)$. Therefore, it is of interest to maximize the contribution of the compressed likelihood function in this ratio rather than the normalized IF. Hence, we need to make $\rho_0$ larger than $\rho_1$.

Clearly, this choice of the normalized IF facilitates the task substantially since it leads to a separable optimization problem. Therefore, one realization of the vector $\boldsymbol{\tau}$ is simply obtained by generating independent $P$ realizations of $P$ scalar random variables according to the same pdf:

$$p(\tau)=\frac{\exp\{\rho_1 I(\tau)\}}{\int_J\exp\{\rho_1 I(u)\}\,du}. \tag{24}$$

Note that the construction of $p(.)$ is made up such that it exhibits $P$ lobes centered at the location of the different delays. The generated variables will be in the vicinity of these $P$ lobes. However, it should be kept in mind that the estimation error on $\boldsymbol{H}_n$ makes other lobes appear and thus some generated variables will take values around these undesired lobes thereby increasing the estimation error. To circumvent this problem, we may increase $\rho_1$ so that the undesired lobes disappear. Yet, we should also keep in mind that increasing $\rho_1$ too much may destroy some useful lobes and some delays may not be estimated properly. Thus, the optimal value of $\rho_1$ is the highest one for which the pdf $p(.)$ has at least $P$ lobes. Moreover, this appropriate choice of $\rho_1$ reduces the computation complexity by reducing the required number of realizations $R$ since it decreases the probability of generating undesired realizations. To summarize, the proposed estimator requires the generation of $R$ realizations of a random vector according to $g'(.)$ which is made by finding a simple separable function instead of a $P$-dimensional function. Then we evaluate the mean of $\tau_1,\ \tau_2,\dots\ \tau_P$ as follows:

$$\hat{\tau}_p=\frac{1}{R}\sum_{k=1}^{R}\boldsymbol{\tau}_k(p)\frac{L'_{c,\rho_0}(\boldsymbol{\tau}_k)}{g'_{\rho_1}(\boldsymbol{\tau}_k)}, \tag{25}$$

[4]Note that the term $\left(M\sigma^2\right)^{-1}$ is inserted in $\rho_1$ to symplify the notation.

where $\boldsymbol{\tau}_k(i)$ is the i$^{th}$ element of the vector $\boldsymbol{\tau}_k$. Some other modifications are adopted to further minimize the computational burden. In fact, noting that the delays are often confined in the interval $[0,LT_c]$ [1], it is more attractive to use the circular mean instead of the linear mean presented in (25). To do so, the delays are transposed into the interval $[0,1]$ before applying the definition of the circular mean [10] to obtain:

$$\widehat{\tau}_p=\frac{LT_c}{2\pi}\angle\frac{1}{R}\sum_{k=1}^{R}F(\boldsymbol{\tau}_k)\exp\left\{\frac{j2\pi\boldsymbol{\tau}_k(p)}{LT_c}\right\}, \tag{26}$$

where $F(.)$ is the weighting factor (i.e., $F(\boldsymbol{\tau})=L'_{c,\rho_0}(\boldsymbol{\tau})/g'(\boldsymbol{\tau})$). From the formulation in (26), we are only interested in finding the angle of a complex number. Therefore, we drop the two positives factors $\int_J\dots\int_J\exp\{\rho L_c(\boldsymbol{u})\}\,d\boldsymbol{u}$ and $\left(\int_J\exp\{\rho_1 I(u)\}\,du\right)^P$, used to normalize $L'_{c,\rho_0}(\boldsymbol{\tau})$ and $g'(\boldsymbol{\tau})$, respectively. Moreover, an overflow may occur since both the numerator and the denominator are exponential. To reduce this overflow, we multiply the weighting factor by a positive number:

$$F'(\boldsymbol{\tau}_k)=\exp\left\{\rho_0 L_c(\boldsymbol{\tau}_k)-\rho_1\sum_{p=1}^{P}I(\boldsymbol{\tau}_k(p))-\max_{1\le l\le R}\left(\rho_0 L_c(\boldsymbol{\tau}_l)-\rho_1\sum_{p=1}^{P}I(\boldsymbol{\tau}_l(p))\right)\right\}. \tag{27}$$

## VI. THE CRAMÈR-RAO LOWER BOUND

It is well known that the CRLB is a lower bound on the variance of any unbiased estimator and this bound will be henceforth derived and used as a benchmark to evaluate the performance of the proposed estimator. In the following, the multipath fading coefficients are modeled as random variables with a known distribution. Therefore, the unknown parameters are $\boldsymbol{\alpha}=[\boldsymbol{\tau},\ \{\Re\{\boldsymbol{P}(l,m)\},\ \Im\{\boldsymbol{P}(l,m)\}\}_{l,m=1}^{P},\ \sigma^2]$. The CRLB is the inverse of the Fisher information matrix (FIM), given by:

$$\boldsymbol{I}(m,n)=M\mathrm{trace}\left\{\boldsymbol{R}^{-1}\frac{\partial\boldsymbol{R}}{\partial\boldsymbol{\alpha}(m)}\boldsymbol{R}^{-1}\frac{\partial\boldsymbol{R}}{\partial\boldsymbol{\alpha}(n)}\right\}, \tag{28}$$

and the lower bound of the parameters of interest are the diagonal elements of the CRLB. However, the derivation of (28) appears to be intractable. Alternatively, the CRLB is asymptotically equivalent to the error covariance matrix, $\boldsymbol{C}_{\mathrm{ML}}(\boldsymbol{\tau})=E\{(\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau})(\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau})^T\}$, of the maximum likelihood estimate, as $M$ tends to infinity [11]. Recall that the ML estimate $\widehat{\boldsymbol{\tau}}$ of $\boldsymbol{\tau}$ verifies the following equation:

$$\frac{\partial L(\widehat{\boldsymbol{\tau}},\boldsymbol{P})}{\partial\boldsymbol{\tau}}=\boldsymbol{0}, \tag{29}$$

where $\partial L(.)/\partial\boldsymbol{\tau}$ is the gradient of $L(.)$ with respect to $\boldsymbol{\tau}$. Using the first-order Taylor series expansion to the left-hand side of (29), we obtain the following expression of the estimation error:

$$\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau}=-\left[\frac{\partial^2 L_0(\boldsymbol{\tau},\boldsymbol{P})}{\partial\boldsymbol{\tau}^2}\right]^{-1}\frac{\partial L(\boldsymbol{\tau},\boldsymbol{P})}{\partial\boldsymbol{\tau}}. \tag{30}$$

When injected back in the covariance matrix $\boldsymbol{C}_{\mathrm{ML}}(\boldsymbol{\tau})$, it leads to the following analytical expression for the CRLB of time delay estimates:

$$\mathrm{CRLB}(\boldsymbol{\tau}) = \frac{\sigma^2}{2M}\left[\Re\left\{ \left(\boldsymbol{U}^H\left[\boldsymbol{I}-\boldsymbol{\Pi}\right]\boldsymbol{U}\right) * \left(\boldsymbol{P}\boldsymbol{D}^H\boldsymbol{R}^{-1}\boldsymbol{D}\boldsymbol{P}\right)^T \right\}\right]^{-1}, \quad (31)$$

where $*$ stands for the element-wise product and the matrix $\boldsymbol{U}$ is defined as follows:

$$\boldsymbol{U} = [\boldsymbol{u}_1,\ \boldsymbol{u}_2, \ldots,\ \boldsymbol{u}_P], \quad (32)$$

$$\boldsymbol{u}_i = \frac{\partial \boldsymbol{d}(\tau_i)}{\partial \tau_i}. \quad (33)$$

To the best of our knowledge, this is the first CRLB expression for DS-CDMA multipath time-delay estimation in closed-form.

## VII. Simulation Results

In this section, we assess the performance of the proposed algorithm. As a benchmark for comparisons with our estimator, we will use the Root-MUSIC algorithm implemented by the STAR [1], [2] receiver to estimate the unknown delays. We consider 3 propagation paths with closely-spaced delays equal to $0.12\ T$, $0.15\ T$ and $0.18\ T$ and a processing gain $L = 64$. The number of realizations $R$ is fixed at $R = 100$. We also assume that the power is equally distributed between the three paths. The mean square error (MSE) — used as performance measure — of the two estimators is compared to the CRLB
First, in Fig. 1, we evaluate the performance for $M = 4$ antennas. As expected, the two methods attain the CRLB at high SNR values. However, the new ML IS-based estimator performs better over the entire SNR range. The difference between these two methods is more perceptible in the medium SNR range. For a relatively small number of realizations $R = 100$, please notice that our new estimator offers very good performance with a reduced complexity burden. If we now use only one receiving antenna branch, the performance of the Root-MUSIC algorithm degrades considerably. Indeed, we can see from Fig. 2 that the gap between the two methods becomes very important. This is hardly surprising since Root-MUSIC needs the estimation of the covariance matrix of the columns of $\boldsymbol{J}_n^T$ from the columns of the matrix $\widehat{\boldsymbol{\mathcal{H}}}_n$, reduced to one column if $M = 1$.
To further investigate this issue, we fix the SNR value at 10 dB and vary the number of antenna branches $M$ from 1 to 8. The MSE of the two algorithms versus $M$ is plotted in Fig. 3. As expected, the IS-based method attains the CRLB starting from a small value of $M$ contrarily to the Root-MUSIC algorithm. This means that the IS-based estimator is well geared toward situations of reduced antenna array sizes.

## VIII. Conclusion

In this paper, we derived a new ML estimator of the time delays for direct-spread CDMA multipath transmissions. We avoided the brute multidimensional grid search to find the global maximum of the likelihood function by recurring to a simple approximation of the compressed likelihood function and adopting the concept of importance sampling. We also provided an analytical expression for the corresponding CRLB on multipath time delay estimation. We have shown through simulations that the new estimator performs better than the classical Root-MUSIC method at a very much reduced complexity cost and that it achieves the CRLB even for a small number of antennas.

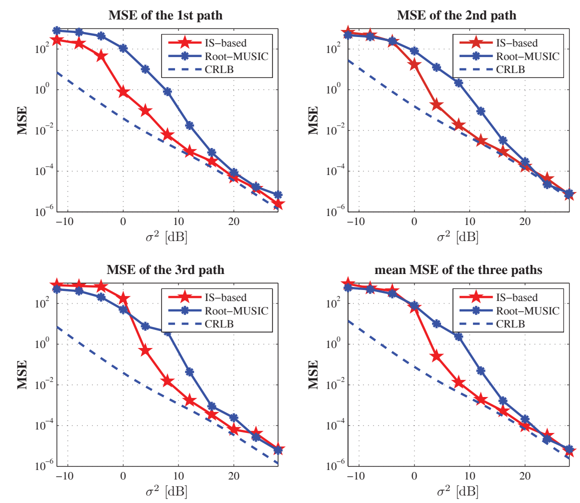


Fig. 1. Estimation performance of the IS-based and the Root-MUSIC algorithms vs. channel estimation error for closely separated delays and $M = 4$.

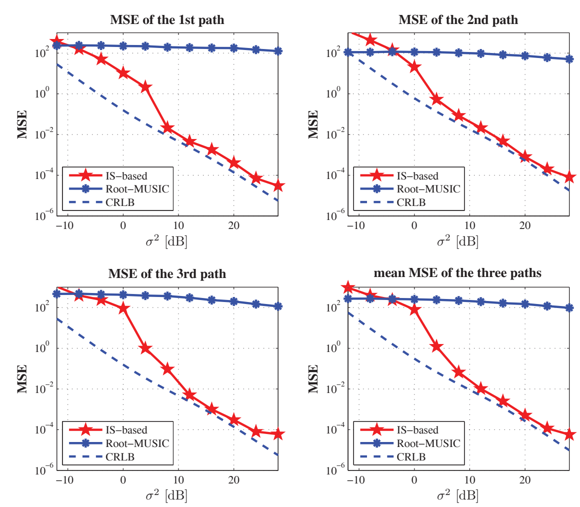


Fig. 2. Estimation performance of the IS-based and the Root-MUSIC algorithms vs. channel estimation error for closely separated delays and $M = 1$

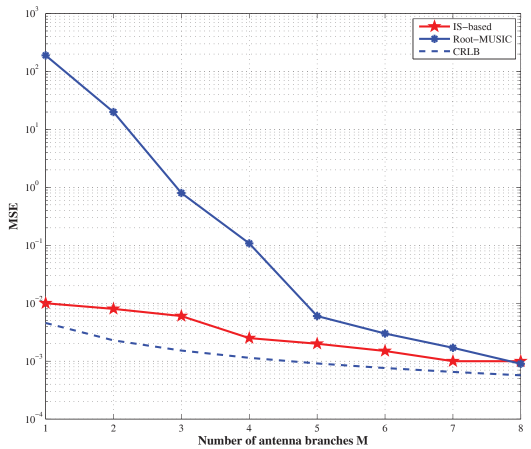


Fig. 3. MSE vs. number of antenna branches for the two algorithms at SNR = 10 dB.

## REFERENCES


[1] S. Affes, P. Mermelstein, “A new receiver structure for asynchronous CDMA: STAR–the spatio-temporal array-receiver”, *IEEE J. Select. Areas in Comm.*, vol. 16, no. 8, pp. 1411-1422, Oct. 1998.

[2] K. Cheikhrouhou, S. Affes, and P. Mermelstein, “Impact of synchronization on performance of enhanced array-receivers in wideband CDMA networks”, *IEEE J. Select. Areas in Comm.*, vol. 19, no 12, pp. 2462-2476, Dec. 2001.

[3] S. Saha, and S. Kay, “Mean likelihood frequency estimation,” *IEEE Trans. Sig. Process.*, vol. 48, pp. 1937-1946, Jul. 2000.

[4] S. Saha, and S. Kay, “An exact maximum likelihood narrowband direction of arrival estimator,” *IEEE Trans. Sign. Process.*, vol. 56, pp. 1082-1092, Oct. 2008.

[5] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, “A non-data-aided maximum likelihood time delay estimator using importance sampling”, IEEE Trans. on Sign. Process., vol. 59, pp. 4505-4515, Oct. 2011

[6] S. Affes and P. Mermelstein, ”Adaptive space-time processing for wireless CDMA”, *Book Chapter: Adaptive Signal Processing: Application to Real-World Promblems*, J. Benesty and A.H. Huang, ms., Springer, Berlin, Jan. 2003.

[7] A. G. Jaffer, ”Maximum likelihood direction finding for stochastic sources: A separable solution,” *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 2893-2896, Apr. 1988,.

[8] M. Pincus, “A closed form solution for certain programming problems,” *Oper. Research*, pp. 690-694, 1962.

[9] W. B. Bishop and P. M. Djuric, “Model order selection of damped sinusoids in noise by predictive densities,” *IEEE Trans. Sign. Process.,* vol. 44, pp. 611-619, Mar. 1996.

[10] K. V. Mardia, *Statistics of directional data*, New York: Academic, 1972.

[11] P. Stoica, A. Nehorai, “Performance study of conditional and unconditional direction-of-arrival estimation” *IEEE Trans. on Acoustics, Speech and Sign. Process.*, pp. 1783-1795, Oct 1990.