Maximum Likelihood Time Delay Estimation From Single- and Multi-Carrier DSSS Multipath MIMO Transmissions for Future 5G Networks

Ahmed Masmoudi, Faouzi Bellili, Sofiène Affes, *Senior Member, IEEE*, and Ali Ghrayeb

Abstract—In this paper, we address the problem of time delay estimation (TDE) from single-carrier (SC) or multi-carrier (MC) direct-sequence spread spectrum (DSSS) multipath transmissions in the presence of multiple transmit and/or receive antennas that will characterize future 5G radio interface technologies (RITs), such as coded-domain nonorthogonal multiple access. We derive for the first time a closed-form expression for the Cramer–Rao lower bound (CRLB) and develop two maximum likelihood (ML) multipath TDEs for SC DSSS single-input multiple-output (SIMO) in the non-data-aided (NDA) case. The first TDE, based on iterative expectation maximization (EM), provides accurate estimates whenever a good initial guess of the parameters is available at the receiver. The second TDE implements the ML criterion in a non-iterative way and finds the global maximum of the *compressed* likelihood function using the importance sampling (IS) technique without requiring any initialization. We also extend both the SC DSSS SIMO CRLB and the two new SC DSSS SIMO ML NDA TDEs to MC DSSS RITs and to multiple-input multiple-output structures with any diversity versus multiplexing pre-coding type before generalizing them all to the data-aided (DA) case. Simulations suggest that the EM TDE is suitable for large observation in space, time, and/or frequency, whereas the IS TDE is preferred in the opposite case of very short data records. Moreover, we show in the NDA case, both analytically and by simulations, that spatial (transmit and receive), temporal, and frequency samples interchangeably have the same impact on estimation accuracy and performance bound regardless of the channel correlation type and amount present in each dimension. Furthermore, we are able to properly cope with such channel correlations that do indeed arise in practice and, hence, become very challenging both in estimation and CRLB derivation in the DA case, but that have been so far overlooked in previous works.

Index Terms—DSSS, 5G, NOMA, SCMA, MUSA, LDS-CDMA, 3G, DS-CDMA, MC-DS-CDMA, MT-CDMA, WCDMA, cdma2000, IEEE 802.11b, 2G, IS-95, TDE, post-correlation model (PCM), maximum likelihood (ML), Cramér-Rao lower bound (CRLB), expectation maximization (EM), importance sampling (IS), data-aided (DA), non-data-aided.

I. INTRODUCTION

THE most important challenge for current and future wireless networks is the development of robust transceivers that are able to transmit at high data rates with high bandwidth efficiency. DSSS systems meet this requirement. As one key SC DSSS RIT, CDMA has indeed already been adopted in 2G IS-95 and 3G WCDMA [3], cdma2000 [4], and IEEE 802.11b [5] standards because of its flexibility in cell planning, user capacity, support for different rates and robustness to multipath effects. In particular, one of the most important motivations behind the use of CDMA is to increase the number of simultaneous users in dense environments with acceptable error performance. Furthermore, MC DSSS RITs such as coded-domain NOMA [6] have already been recognized as potential candidates for future 5G networks due to their improved spectrum efficiency and robustness against more adverse conditions of channel frequency selectivity [7]–[9], more so when implemented in MIMO structures. Therefore, it has the potential to be adopted in future 5G high-density wireless networks.

Yet, to ensure good performance, these systems require large time synchronization capabilities. In digital communications, the output of the demodulator is sampled periodically in order to recover the transmitted information. Since the propagation delay from the transmitter to the receiver is generally unknown at the receiver, this delay must be estimated from the received signal in order to efficiently sample the output of the demodulator. Knowledge of the propagation delay is also a requirement in many other applications related to localization and tracking which both require highly accurate time delay estimation.

In this work, without lack of generality, we focus on the CDMA array receiver which has received much interest sequel to the performance potential it carries [9]–[12]. Roughly speaking, the post correlation model (PCM) of the despread data presents the signal in the manifold structure that has been massively studied in the field of array signal processing. To date, a suboptimal root-MUSIC estimator was initially developed in [13] to recover the time delays and was later refined in [7] to significantly reduce its complexity. As a low-complexity subspace-based method, it relies on the maximization of the spectrum of the received signal to find its peak frequency components. It is not iterative and does not need any initialization, but instead finds the roots of a polynomial obtained from the inverse of the spectrum function. The present work also investigates multiple time delay estimation for CDMA array receivers, but in an optimal way in which the ML criterion is investigated using the PCM [7], [13].

Actually, the problem of high-resolution parameter estimation has been extensively studied in the past few decades. In this context, it is well known that the ML technique always outperforms any sub-optimal method in the severe conditions of low signal-to-noise ratios (SNRs) or a small number of available data snapshots. However, a direct implementation of the ML criterion requires a multi-dimensional grid-search which is impractical since its complexity increases exponentially with the number of unknown parameters. Alternatively, eigen-decomposition methods (which reduce the problem to one-dimensional grid-search) [14] have attracted much interest due to their simplicity and their high-resolution capabilities. Yet, they rely on the sample estimate of the covariance matrix and require therefore a large number of data snapshots. However, as we will see later, the number of snapshots, is proportional to the product of the number of transmit or receive antennas, carriers, and symbols. Thus, applying a traditional covariance-based technique would require a large number of receiving antennas and/or carriers and/or symbols.

Motivated by these facts, our main goal here is to derive efficient implementations of the ML-based TDEs that avoid the trivial yet computationally-prohibitive multidimensional grid-search approach. To that end, iterative methods are usually envisioned in order to find the ML estimates since a closed-form solution is, in most cases, deemed intractable. We first develop an efficient iterative solution for the estimation of the delays using the EM concept. While this concept was previously leveraged to solve the problem of multiple time delays estimation [15], it has not been yet adapted to the context of CDMA systems. As will be seen later, its derivation for CDMA array receivers is far from being straightforward. In a nutshell, the proposed EM TDE virtually decomposes the observed signal into different replicas, each corresponding to a single path, and then treats each component separately. Therefore, the original multidimensional optimization problem is successively recast into multiple one-dimensional optimization tasks, thereby resulting in tremendous computational saving. Under good initialization, the algorithm is guaranteed to converge to the global maximum of the likelihood function.

Alternatively, when a good initialization is not available, we resort to the importance sampling (IS) concept to derive another technique that finds the global maximum of the likelihood function in a non-iterative way. In our case, the likelihood function depends on the time delays and the channel covariance matrix. To obtain a function that depends on the unknown delays only, the channel covariance matrix is replaced by its ML estimate which is a function of the delays themselves. The resulting objective function, called CLF, is then maximized with respect to the unknown time delays. To do so, we rely on the Pincus' theorem [16] which provides an efficient tool for finding the global maximum of multidimensional objective functions. However, it still requires the computation of a multidimensional integral which is itself difficult to compute. To sidestep this problem, we will make use of the powerful IS concept and recast the multi-dimensional integration problem into the computation of expected values of multi-dimensional random variables. The required realizations that are needed to compute the underlying mean values will be easily generated according to a properly designed importance function. Compared to the EM ML estimator, the IS-based approach does not require any initialization and hence does not suffer from lack of global convergence problems.

It should be noted here that the combination of Pincus' theorem and IS concept has been previously applied to many fundamental estimation problems. To the best of our knowledge, however, this elegant combination was first pioneered by S. Kay and S. Saha in [17] in the context of multiple frequencies estimation. There, it was shown for the very first time that joint ML estimation of multiple frequencies boils down to the computation of sample mean estimates from a number of realizations generated according to a carefully designed importance function (or pseudo-PDF). Pincus' theorem along with the IS concept were later on applied to the joint estimation of the angle-Doppler [18]. More recently, these powerful tools were leveraged in the context of time difference of arrival (TDOA)-based source localization [19] and NDA timing recovery [20]. In [21], we also presented a TDE for multipath environments using the IS method. The multipath case in [21] was specifically developed in the DA case for unmodulated signals (i.e., radar or sonar signals) which are different from the DSSS signal structure treated in the current paper. Moreover, the work in [21] dealt only with the IS-based estimator, since the EM ML estimator and the CRLB were already presented in the literature for the specific case of radar and sonar signals.

The contributions of this work cover both SC and MC DSSS with MIMO, SIMO Multiple-Input Single-Output (MISO), or Single-Input Single-Output (SISO) multi-antenna structures. While the SC and MC DSSS show-cases considered herein, without lack of generality, are SC-DS-CMDA and MC-DS-CDMA, respectively, all developments presented in the article remain valid for other SC or MC DSSS RITs such as IS-95, cdma2000, WCDMA, and IEEE 802.11b implemented standards, multitone (MT)-CDMA [22] and prospective 5G coded-domain NOMA [6].

Indeed, many key RIT candidates for adoption in future 5G networks such as LDS-CDMA (low-density-spreading CDMA), SCMA (sparse-code multiple access) and MUSA (multi-user synchronous access) rely on code multiplexing

to accommodate more users simultaneously on the same time-frequency resources [6]. The main difference between these schemes is the way they combine interleaving, modulation and spreading. LDS-CDMA, for example, uses large spreading sequences with few nonzero elements instead of the dense ones usually adopted in DS-CDMA to efficiently reduce interference among multiple users with appropriately designed spreading sequences. SCMA is a multi-dimensional codebook-based spreading technique which can be seen as a generalization of LDS-CDMA. It merges the QAM mapping and the CDMA spreading together to directly map a set of bits to a sparse vector called codeword. No matter how spreading codes are exploited, whether for pure spreading, symbol mapping or modulation, or even both (cf. IS-95 example in [23] where Walsh-Hadamard sequences are used for modulation), the PCM model and hence the proposed TDEs remain valid and applicable to all above-mentioned coded-domain NOMA RITs.

The proposed developments also cover both DA and NDA estimation scenarios. In the NDA case, we prove that any possible impact of channel correlation in space, time, and/or frequency on the observed samples is offset by data modulation due to the observation's randomization with the symbols's quasi-independence across all dimensions. We hence prove that time, frequency, and space (transmit and receive) dimensions merge together into a combined dimension which is the product of the three where spatial, temporal or frequency samples interchangeably have exactly the very same impact on estimation performance regardless of the channel correlation type and amount present in each dimension. In the DA case, we are able to properly cope with such channel correlations that do indeed arise in practice and, hence, become very challenging both in estimation and CRLB derivation, but that have been so far overlooked in previous works.

In [1], we proposed an EM ML TDE for DS-CDMA multipath transmissions. There, the algorithm was developed for SIMO SC-DS-CDMA NDA with the straightforward extension to MC-DS-CDMA by applying the EM algorithm separately on each subcarrier and then averaging the resulting estimates over all subcarriers. Here, we propose 1) a more judicious extension of the algorithm to MC-DS-CDMA by applying it only once over all the subcarriers jointly, thereby reducing considerably the computational cost; that is on the top of 2) extending it to the DA case; and 3) from SIMO to MIMO transceiver structures. A basic version of the IS TDE was also presented in [2] for SIMO SC-DS-CDMA NDA only. Here we 4) extend it to the cases of DA, MC, and MIMO transmissions. We also 5) develop for the first time the most general closed-form CRLB expressions ever that i) cover both SC and MC DSSS transmissions with either SISO, SIMO, MISO, or MIMO transceiver structures with any diversity versus multiplexing pre-coding type, ii) apply in both NDA and DA cases, and iii) account for the challenging impact of channel correlation in space, time, and/or frequency of any type that inevitably arise in the DA case but that have been so far overlooked in previous works.

II. SIMO SC DSSS NDA CASE

A. System Model and Background

Without lack of generality, we consider a SIMO SC-DS-CDMA communication system where the receiver is equipped with M independent receiving antenna elements that capture signals travelling through a multipath propagation environment consisting of P different Rayleigh-fading paths. The signals received on the M antennas are uncorrelated with the spreading code and sampled at the chip rate T_c. Denoting the processing gain by L (i.e., $L = T/T_c$ with T being the symbol duration), the resulting post-correlation data of the spatio-temporal observation during the nth received symbol is modelled by the following matrix form [13]:

$$Z_n = G_n \Upsilon_n D(\tau)^T s_n + N_n, \qquad (1)$$

where the unknown delays are gathered in the parameter vector $\tau = [\tau_1, \ldots, \tau_P]^T$, $s_n = b_n \psi_n$ is the product of the unknown transmitted symbol, b_n, and the square root, ψ_n, of the total received power ψ_n^2 and N_n is the $M \times L$ post-correlation noise matrix. The pth column of the matrix $D(\tau)$ that gathers the time delay parameters, $\tau_1, \ldots, \tau_P$, is given by:

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

where $\rho_c(.)$ is the correlation function of the spreading code. G_n is the $M \times P$ spatial propagation matrix of unit-power Rayleigh-fading random variables and Υ_n is a $P \times P$ diagonal matrix representing the normalized power ratios over the different paths (i.e., trace$\{\Upsilon_n^2 = 1\}$) [13]. These two matrices can be grouped in one single spatial-response matrix J_n (i.e., $J_n = G_n \Upsilon_n$) and by including the scalar term s_n in the matrix[1] J_n, a more compact form of Z_n is given by:

$$Z_n^T = D(\tau) J_n^T + N_n^T. \qquad (3)$$

Using the representation in (3), the original problem can be interpreted as the estimation of the time delays, involved in the matrix $D(\tau)$, from M snapshots observed on L antennas. Each column of Z_n^T represents an observation vector and the columns of J_n^T are interpreted as the transmitted signals from P different sources. Supposing that the delay vector τ remains constant over N consecutive symbols, a compact representation of (3) over the N symbols is given by:

$$Z \triangleq \left[Z_1^T, \ Z_2^T, \ldots, \ Z_N^T\right] = D(\tau) J^T + N^T, \qquad (4)$$

with $J^T \triangleq \left[J_1^T, \ldots, J_N^T\right]$ and $N^T \triangleq \left[N_1^T, \ldots, N_N^T\right]$. In our quest for a computationally tractable ML solution, we first perform a column-by-column fast Fourier transform (FFT) of Z^T to obtain:

$$\mathcal{Z} = \mathcal{D}(\tau) J^T + \mathcal{N}, \qquad (5)$$

where $\mathcal{N}$ is the resulting transformed noise matrix and $\mathcal{D}(\tau)$ depends only on the unknown delays and is given by:

$$\mathcal{D}(\tau) = [d(\tau_1), \ d(\tau_2), \ldots, \ d(\tau_P)], \qquad (6)$$

[1] For the sake of simplicity, we use the same notation J_n for both J_n and $J_n s_n$. Yet the formulation holds, unless specified otherwise, for NDA.

where $d(\tau_p) = \left[c_0, \ c_1 e^{-\frac{j2\pi\tau_p}{LT_c}}, \ldots, \ c_{L-1} e^{-\frac{j2\pi \ (L-1)\tau_p}{LT_c}} \right]^T$ depends on the pth delay and $\{c_l\}_{l=0}^{L-1}$ are the FFT coefficients of the spreading code correlation function. Note that, for CDMA systems, the correlation function of a perfect spreading code is a Dirac function. In this special ideal case, the corresponding FFT coefficients are constant. This feature actually holds true as a very good approximation even with practical spreading codes [7], [13].

B. The CRLB

Before we develop the two new estimators, we derive in this part a closed-form expression for the CRLB for the problem at hand, which will be used as an overall benchmark against which we gauge the performance of the new estimators.

In fact, the CRLB is a well known lower bound for the variance of any unbiased estimator of a given parameter. Many works have so far dealt with the evaluation of the CRLB for the TDE problem but, as far as we know, no contributions have been made yet in the context of multipath DSSS systems. To that end, we assume that the multipath fading coefficients, gathered in J_n^T, are random variables with unknown covariance matrix R_J which is a diagonal matrix.[2] In the following, R_J is assumed to be the same for all receiving antennas and remains constant during N consecutive symbols. Therefore, the vector of unknown parameters involved in the estimation process is:

$$\alpha = \left[\tau, \Re\{R_J(m,m)\}_{m=1}^P, \Im\{R_J(m,m)\}_{m=1}^P, \sigma^2 \right]^T, \quad (7)$$

where $\Im\{.\}$ returns the real and imaginary parts of any complex quantity. In the following, we suppose that the columns of Z, denoted Z_i, are mutually independent and that the columns of $\mathcal{N}$ are also mutually independent and Gaussian distributed. Under these assumptions and starting from (5), the probability density function (PDF) of Z, parametrized by τ and R_J (the covariance matrix of the columns of J_n^T), is given by:

$$\overline{p}(Z; \tau, R_J)$$
$$= \frac{1}{\left(\pi^L \det\{D(\tau)R_J D(\tau)^H + \sigma^2 I_L\} \right)^{MN}}$$
$$\times \exp\left\{ -\sum_{i=1}^{MN} Z_i^H \left(D(\tau)R_J D(\tau)^H + \sigma^2 I_L \right)^{-1} Z_i \right\}, \quad (8)$$

where $\det\{.\}$ returns the determinant of any square matrix and I_L is the $(L \times L)$ identity matrix. Therefore, the log-likelihood function, $\mathcal{L}(\tau, R_J) = \ln\left(\overline{p}(Z; \tau, R_J) \right)$, reduces simply to:

$$\mathcal{L}(\tau, R_J)$$
$$= -\ln\left(\det\left\{ D(\tau)R_J D(\tau)^H + \sigma^2 I_L \right\} \right)$$
$$- \frac{1}{MN} \sum_{i=1}^{MN} Z_i^H \left(D(\tau)R_J D(\tau)^H + \sigma^2 I_L \right)^{-1} Z_i. \quad (9)$$

[2] R_J is a $P \times P$ matrix whose (i,j)th generic term is the correlation factor between the two multipath propagation channels i and j. And since by definition resolvable paths are mutually uncorrelated, R_J should be diagonal.

Usually, the CRLB is computed by deriving and inverting the Fisher information matrix (FIM), denoted here by $\mathfrak{J}$, whose entries are given by:

$$[\mathfrak{J}]_{m,n} = M \ \text{trace}\left\{ R_Z^{-1} \frac{\partial R_Z}{\partial \alpha(m)} R_Z^{-1} \frac{\partial R_Z}{\partial \alpha(n)} \right\}, \quad (10)$$

with $R_Z = D(\tau)R_J D(\tau)^H + \sigma^2 I_L$ being the covariance matrix of Z_i. After tedious algebraic manipulations (see [2] for more details), we obtain the following analytical expression for the CRLB of the time delay estimates:

$$\text{CRLB}(\tau)$$
$$= \frac{\sigma^2}{2MN}\left[\Re\left\{ \left(U^H \Pi^\perp U \right) \odot \left(R_J D(\tau)^H R_Z^{-1} D(\tau) R_J \right)^T \right\} \right]^{-1}, \quad (11)$$

where $\odot$ stands for the element-wise product, $\Pi^\perp = I_L - D(\tau)\left(D(\tau)^H D(\tau) \right)^{-1} D(\tau)^H$ is an orthogonal projection onto the null space of $D(\tau)$, and the matrix U is defined as:

$$U \triangleq [u_1, \ u_2, \ldots, \ u_P] = \left[\frac{\partial d(\tau_1)}{\partial \tau_1}, \ \frac{\partial d(\tau_2)}{\partial \tau_2}, \ldots, \ \frac{\partial d(\tau_P)}{\partial \tau_P} \right]. \quad (12)$$

The analytical expression of the CRLB in (11) reveals that the time delay estimation from SC-DS-CDMA NDA merges both space and time dimensions whereby spatial and temporal samples indistinguishably play identical roles.

C. The EM ML TDE

The EM algorithm is a computationally-modest technique that finds the maximum likelihood estimate whenever a closed-form solution is intractable. Rewriting the LLF in (9) in a more compact form, we obtain:

$$\mathcal{L}(\tau, R_J) = -\ln\left(\det\{R_Z\} \right) - \text{trace}\left\{ R_Z^{-1} \widehat{R}_Z \right\}, \quad (13)$$

with $\widehat{R}_Z$ being an estimate of the actual covariance matrix, R_Z, computed from the columns of Z as follows:

$$\widehat{R}_Z = \frac{1}{MN} \sum_{i=1}^{MN} Z_i Z_i^H. \quad (14)$$

Note here that the LLF, $\mathcal{L}(\tau, R_J)$, depends on the delays vector of interest, τ, and the unknown covariance matrix R_J. The goal is therefore to jointly maximize $\mathcal{L}(\tau, R_J)$ with respect to τ and R_J. Here, while σ^2 is usually unknown, it can be easily estimated either by averaging the $L - M$ smallest eigen-values of $\widehat{R}_Z$ or simply by exploiting the estimated power that could be obtained from a previous processing stage of the receiver (see [13] for more details).

Clearly, the above expression of the likelihood function in (13) cannot be maximized analytically. Thus, we resort as a first option to the well-known EM concept [15], to maximize (13) iteratively. The purpose is to decompose the observation vectors, $\{Z_i\}_{i=1}^{MN}$ into P complete-data from which the P delays are estimated separately. This is equivalent to performing P parallel maximizations over a one-dimensional space. This method reduces considerably the computational

complexity compared to the brute grid-search. Towards this goal, we define the set of complete data as:

$$z^{(p)}(i) = J^T(i, p)d(\tau_p) + n^{(p)}(i),$$
$$p = 1, 2, \ldots, P, \ i = 1, 2, \ldots, MN, \quad (15)$$

where $z^{(p)}(i)$ can be seen as the ith spatio-temporal snapshot from the pth path and $n^{(p)}(i)$ is an arbitrary decomposition of the noise $\left(\text{i.e., } \mathcal{N}_i = \sum_{p=1}^{P} n^{(p)}(i)\right)$. From (15), the covariance of $z^{(p)}(i)$, $\mathbb{E}\left\{z^{(p)}(i)z^{(p)}(i)^H\right\}$ is given by:

$$R_{z^{(p)}} = \varepsilon_p^2 d(\tau_p)d(\tau_p)^H + \frac{\sigma^2}{P}I_L, \quad (16)$$

with $\{\varepsilon_p^2\}_{p=1}^P$ being the diagonal elements of R_J. From (15), any column, z_i, of z can be written as a function of the complete data as follows:

$$z_i = \sum_{p=1}^{P} z^{(p)}(i). \quad (17)$$

Now, we are ready to describe the Expectation-step (E-step) and the Maximization-step (M-step) of the EM algorithm. The E-step consists in finding the conditional expectations of the sample covariance matrices $\{\widehat{R}_{z^{(p)}}\}_{p=1}^P$ of the complete data defined as:

$$\widehat{R}_{z^{(p)}} = \frac{1}{MN} \sum_{i=1}^{MN} z^{(p)}(i)\left(z^{(p)}(i)\right)^H. \quad (18)$$

Given $R_J^{\{q-1\}}$ and $\tau^{\{q-1\}}$ (the estimates of R_J and τ at iteration $(q-1)$) and $\widehat{R}_z$, the expectation of $\widehat{R}_{z^{(p)}}$ can be computed from the classical formulas of the conditional expectation with Gaussian distributed random vectors as follows:

$$\widehat{R}_{z^{(p)}}^{\{q\}} = \mathbb{E}\left\{\widehat{R}_{z^{(p)}}\Big|\widehat{R}_z; R_J^{\{q-1\}}; \tau^{\{q-1\}}\right\}$$
$$= R_{z^{(p)}}^{\{q\}}\left(R_z^{\{q\}}\right)^{-1}\widehat{R}_z\left(R_z^{\{q\}}\right)^{-1}R_{z^{(p)}}^{\{q\}} + R_{z^{(p)}}^{\{q\}}$$
$$- R_{z^{(p)}}^{\{q\}}\left(R_z^{\{q\}}\right)^{-1}R_{z^{(p)}}^{\{q\}}, \quad (19)$$

where the matrices $R_{z^{(p)}}^{\{q\}}$ are computed at the qth iteration from the estimates $\tau_p^{\{q-1\}}$ and $\varepsilon_p^{2\,\{q-1\}}$ already computed at iteration $q-1$. Now, turning to the estimation of $R_z^{\{q\}}$, the procedure is different from the one used in previous EM algorithms in [15] and [25] where the covariance matrix of the received signal is simply diagonal, contrarily to the problem at hand. Therefore, we resort here to another approach to estimate the covariance matrix $R_z^{\{q\}}$. First, assume that $R_z^{\{q\}}$ is a Toeplitz matrix[3] (i.e., stationary processes). Then, we consider the method of estimating Toeplitz-structured matrices proposed in [27] (briefly detailed here), which is also based on the EM principle. We define the $N_s \times N_s$ circulant extended version of R_z, denoted as R_s. The matrix R_s represents the covariance matrix of the extended vectors $\{\widetilde{z}_i\}_{i=1}^{MN}$, where $\widetilde{z}_i$ consists of the vector z_i augmented by (N_s-L)-dimensional zero vectors.

The extended covariance matrix R_s has the following eigen-decomposition:

$$R_s = F^H R_C F, \quad (20)$$

where F is the standard $N_s \times N_s$ discrete Fourier transform (DFT) matrix and R_C is a diagonal matrix constructed from the eigenvalues of R_s. The DFT transform of $\widetilde{z}_i$ yields the rotated vectors $C_i = F\widetilde{z}_i$ for $i = 1, 2, \ldots, MN$. Denoting by $\widehat{R}_C$ the estimate of R_C obtained from $\{C_i\}_{k=1}^{MN}$ (i.e., $\widehat{R}_C = \frac{1}{MN}\sum_{i=1}^{MN}C_iC_i^H$), the expectation of $\widehat{R}_C$ conditioned on R_z and $\widehat{R}_z$ — by applying the same formula used to find (19) — is given by:

$$\mathbb{E}\left\{\widehat{R}_C\Big|R_z, \widehat{R}_z\right\}$$
$$= R_{zC}\left(R_z^{-1}\widehat{R}_z(R_z^{-1})^H - R_z^{-1}\right)R_{zC} + R_C, \quad (21)$$

where R_{zC} is the cross-covariance of C_i and z_i. Noticing that $z_i = \widetilde{F}^H C_i$, with $\widetilde{F} = F[I_L \ 0]^T$ and 0 is the $(L \times N_s - L)$ zero matrix, the cross-covariance matrix R_{zC} is equal to $R_C\widetilde{F}$. Thus, the estimate of R_C at iteration q is given by:

$$R_C^{\{q\}} = \text{diag}\left(R_C^{\{q-1\}}\widetilde{F}\left(\left(R_z^{\{q-1\}}\right)^{-1}\widehat{R}_z\left(R_z^{\{q-1\}}\right)^{-1}\right.\right.$$
$$\left.\left. - \left(R_z^{\{q-1\}}\right)^{-1}\right)\widetilde{F}^H R_C^{\{q-1\}} + R_C^{\{q-1\}}\right), \quad (22)$$

and R_z is obtained using the transformation $z_i = \widetilde{F}^H C_i$ as follows:

$$R_z^{\{q\}} = \widetilde{F}^H R_C^{\{q\}}\widetilde{F}. \quad (23)$$

As it was shown in [27], the stable point of (23) is equal to the ML estimate of R_z.

During the M-step of the EM algorithm, we aim at maximizing the LLF of the complete-data with respect to the parameters of interest $\{\tau_i\}_{i=1}^P$. It is the same objective function given in (9), with the true expectation of the complete data being replaced by the conditional expectation of $z^{(p)}(i)$; in other words R_z is substituted by $R_{z^{(p)}}$ and $\widehat{R}_z$ by $\widehat{R}_{z^{(p)}}^{\{q\}}$. Thus, we obtain the LLF of the complete-data, $\mathcal{L}_p(\tau_p, R_{z^{(p)}})$, as follows:

$$\mathcal{L}_p(\tau_p, R_{z^{(p)}}) = -\ln\left(\det\left\{R_{z^{(p)}}\right\}\right) - \text{trace}\left\{\widehat{R}_{z^{(p)}}^{\{q\}}R_{z^{(p)}}^{-1}\right\}. \quad (24)$$

Then, at iteration q, the estimates $\widehat{\tau}_p^{\{q\}}$ and $\widehat{R}_{z^{(p)}}^{\{q\}}$ are those which jointly maximize $\mathcal{L}_p(\tau_p, R_{z^{(p)}})$. Using the eigen-decomposition[4] of $R_{z^{(p)}}$, the LLF of the complete-data can be expressed as:

$$\mathcal{L}_p(\tau_p, R_{z^{(p)}}) = -\ln\left(\varepsilon_p^2 + \frac{\sigma^2}{P}\right) - (M-1)\ln\left(\frac{\sigma^2}{P}\right)$$
$$\times \left(\frac{1}{\varepsilon_p^2 + \frac{\sigma^2}{P}} - \frac{P}{\sigma^2}\right)d(\tau_p)^H\widehat{R}_{z^{(p)}}^{\{q\}}d(\tau_p)$$
$$- \frac{P}{\sigma^2}\text{trace}\left\{\widehat{R}_{z^{(p)}}^{\{q\}}\right\}, \quad (25)$$

---

[3]This assumption implies that the covariance between z_m and z_n depends only on the difference between m and n, which characterizes stationary processes.

[4]Noticing that the matrix $\varepsilon_p^2 d(\tau_p)d(\tau_p)^H$ is of rank one with $L-1$ zero eigenvalues and one equal to ε_p^2, the eigen-decomposition of $R_{z^{(p)}}$ can therefore be easily performed.

TABLE I
SUMMARY OF STEPS OF THE EM TDE

1) Compute $\boldsymbol{R}_Z^{\{q\}}$ from (22) and (23).   Repeat for $p = 1, \ldots, P$ 2) Evaluate $\widehat{\boldsymbol{R}}_{z(p)}^{\{q\}}$ as defined in (19).   3) Maximize $\boldsymbol{d}(\tau_p)^H \widehat{\boldsymbol{R}}_{z(p)}^{\{q\}} \boldsymbol{d}(\tau_p)$ to find $\tau_p^{\{q\}}$ and evaluate $\varepsilon_p^{2\{q\}}$ from (26).

which emphasizes the dependence of $L_p(\tau_p, \boldsymbol{R}_{z(p)})$ on τ_p and ε_p^2. The closed-form expression of its maximum with respect to ε_p^2, for a given $\tau_p^{\{q\}}$, is:

$$\varepsilon_p^{2\{q\}} = \boldsymbol{d}(\tau_p^{\{q\}})^H \widehat{\boldsymbol{R}}_{z(p)}^{\{q\}} \boldsymbol{d}(\tau_p^{\{q\}}) - \frac{\sigma^2}{P}. \tag{26}$$

Now, plugging (26) in (25) yields the following one-dimensional maximization problem:

$$\tau_p^{\{q\}} = \arg\max_{\tau_p} \left\{ -\ln\left(\boldsymbol{d}(\tau_p)^H \widehat{\boldsymbol{R}}_{z(p)}^{\{q\}} \boldsymbol{d}(\tau_p)\right) \right.$$
$$\left. + \frac{P}{\sigma^2} \boldsymbol{d}(\tau_p)^H \widehat{\boldsymbol{R}}_{z(p)}^{\{q\}} \boldsymbol{d}(\tau_p) \right\}, \tag{27}$$

or simply the problem of maximizing $\boldsymbol{d}(\tau_p)^H \widehat{\boldsymbol{R}}_{z(p)}^{\{q\}} \boldsymbol{d}(\tau_p)$. This follows immediately from the fact that $f(x) = -\ln(x) + \frac{P}{\sigma^2} x$ is a monotonic and increasing function. The main steps of the EM TDE are summarized in Table I.

A variation of the proposed EM algorithm can be obtained following the idea of the space-alternating generalized EM (SAGE) method [29]. The SAGE method is an implementation of the EM algorithm that has a higher convergence rate. Applying it to the problem at hand, we update the estimate $\boldsymbol{R}_Z^{\{q\}}$ after estimating the pth delay $\tau_p^{\{q\}}$ then evaluate $\widehat{\boldsymbol{R}}_{z(p+1)}^{\{q\}}$ with the new $\boldsymbol{R}_Z^{\{q\}}$. In the EM algorithm described in Table I, it is worth mentioning that the estimates $\boldsymbol{R}_Z^{\{q\}}$ and $\boldsymbol{R}_Z^{\{q\}}$ are re-evaluated only when all delays have been updated.

So far, the ML estimate has been found in an iterative way. Clearly, this method needs an initial guess about the unknown parameters to serve as an adequate starting point for the iterative EM TDE. Alternatively, to avoid all initialization hurdles, we develop in the next section a non-iterative algorithm that efficiently finds the ML estimates using the IS approach.

D. The IS ML TDE

Similar to the previous algorithm, we start from the expression of the LLF in (9). As mentioned, a direct maximization of this function imposes joint maximization over $\boldsymbol{\tau}$ and $\boldsymbol{R}_J$. Therefore, it will be of interest to formulate an objective function that depends on the time delays only. To that end, we first maximize the *actual* likelihood function in (9) with respect to the nuisance parameter matrix $\boldsymbol{R}_J$. For this purpose, it can be shown that the value of $\boldsymbol{R}_J$ that maximizes $L(\boldsymbol{\tau}, \boldsymbol{R}_J)$ for a fixed vector $\boldsymbol{\tau}$ is:

$$\widehat{\boldsymbol{R}}_J^{\text{ML}} = \left[\boldsymbol{\mathcal{D}}(\tau)^H \boldsymbol{\mathcal{D}}(\tau)\right]^{-1} \boldsymbol{\mathcal{D}}(\tau)^H \widehat{\boldsymbol{R}}_Z \boldsymbol{\mathcal{D}}(\tau) \left[\boldsymbol{\mathcal{D}}(\tau)^H \boldsymbol{\mathcal{D}}(\tau)\right]^{-1}$$
$$- \sigma^2 \left[\boldsymbol{\mathcal{D}}(\tau)^H \boldsymbol{\mathcal{D}}(\tau)\right]^{-1}. \tag{28}$$

Then, using $\widehat{\boldsymbol{R}}_J^{\text{ML}}$ back into (9) yields the CLF of the system:

$$L_c(\boldsymbol{\tau}) = \frac{1}{\sigma^2}\text{trace}\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_Z\right\} - \ln\left(\det\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_Z\boldsymbol{\Pi} + \sigma^2\boldsymbol{\Pi}^\perp\right\}\right), \tag{29}$$

where $\boldsymbol{\Pi}$ is an orthogonal projection matrix defined as $\boldsymbol{\Pi} = \boldsymbol{\mathcal{D}}(\tau)\left[\boldsymbol{\mathcal{D}}(\tau)^H \boldsymbol{\mathcal{D}}(\tau)\right]^{-1}\boldsymbol{\mathcal{D}}(\tau)^H$. Now, the ML estimates of the time delays are obtained by maximizing the obtained CLF $L_c(\boldsymbol{\tau})$ with respect to $\boldsymbol{\tau}$. In this section, as an alternative to the iterative method already presented in section II-C, we introduce a non-iterative implementation of the ML criterion. We resort to the global maximization theorem of Pincus [16] in order to find the global maximum of the multi-dimensional function at hand. In fact, according to [16], the global maximum,[5] $\widehat{\boldsymbol{\tau}} \triangleq [\widehat{\tau}_1, \ldots, \widehat{\tau}_P]^T$, of $L_c(\boldsymbol{\tau})$ with respect to $\boldsymbol{\tau}$ is given by:

$$\widehat{\tau}_p = \lim_{\rho \to \infty} \frac{\int_J \cdots \int_J \tau_p \exp\{\rho L_c(\boldsymbol{\tau})\}\, d\boldsymbol{\tau}}{\int_J \cdots \int_J \exp\{\rho L_c(\boldsymbol{\tau})\}\, d\boldsymbol{\tau}}, \quad p = 1, \ldots, P, \tag{30}$$

with $J = [0, T]$ being the interval in which the delays are supposed to be confined. Clearly, as ρ tends to infinity, the fraction $\frac{\int_J \cdots \int_J \tau_p \exp\{\rho L_c(\boldsymbol{\tau})\} d\boldsymbol{\tau}}{\int_J \cdots \int_J \exp\{\rho L_c(\boldsymbol{\tau})\} d\boldsymbol{\tau}}$ becomes a multidimensional Dirac function, centered at the global maximum of $L_c(.)$. Therefore, if we define the pseudo-PDF $L'_{c,\rho}(.)$ as:

$$L'_{c,\rho}(\boldsymbol{\tau}) = \frac{\exp\{\rho L_c(\boldsymbol{\tau})\}}{\int_J \cdots \int_J \exp\{\rho L_c(\boldsymbol{\tau})\}\, d\boldsymbol{\tau}}, \tag{31}$$

the ML estimates of $\{\tau_p\}_{p=1}^P$, given in (30), can be reformulated as:

$$\widehat{\tau}_p = \int_J \cdots \int_J \tau_p L'_{c,\rho_0}(\boldsymbol{\tau}) d\boldsymbol{\tau}, \quad p = 1, 2, \ldots, P, \tag{32}$$

where ρ_0 is a sufficiently large number (whose optimal value will be discussed later). The function $L'_{c,\rho_0}(.)$ is called a pseudo-PDF since it has all the properties of an PDF, although $\boldsymbol{\tau}$ is not truly random. As seen from (32), the ML estimate requires the evaluation of multi-dimensional integrals, which is usually a difficult task if not prohibitive as the number of delays gets larger. However, exploiting the fact that $L'_{c,\rho_0}(.)$ is a pseudo-PDF, the involved integral can be simply interpreted as the mean value of the pth element, τ_p, of the vector $\boldsymbol{\tau}$ when the latter is distributed according to $L'_{c,\rho_0}(.)$. Therefore, one can easily evaluate this mean — and hence the integrals in (32) — in order to obtain $\widehat{\boldsymbol{\tau}}$ using Monte Carlo techniques [21], [20] as follows:

$$\widehat{\boldsymbol{\tau}} = \frac{1}{R}\sum_{r=1}^R \boldsymbol{\tau}_r, \tag{33}$$

where $\{\boldsymbol{\tau}_r\}_{r=1}^R$ are R realizations of $\boldsymbol{\tau}$, with $\boldsymbol{\tau}$ being distributed according to $L'_{c,\rho_0}(.)$. Yet, another challenging problem arises here as how to jointly generate $\{\tau_p\}_{p=1}^P$ for a multidimensional random variable. Actually, $L'_{c,\rho_0}(.)$ is constructed using the actual CLF in (29), which is a multi-dimensional function;

---

[5]Cf. proof of convergence to the global maximum by Pincus in [16]. The only requirement to prove it is that the CLF be continuous and attain its global maximum at exactly one point of the compact $\boldsymbol{R}^P$. From estimation theory, $L_c(.)$ is continuous and the CLF has one global maximum.

making the generation of the vector $\boldsymbol{\tau}$ a very difficult task if not impossible. Therefore, it is of interest to find another pseudo-PDF to generate the realizations instead of directly using $\mathcal{L}'_{c,\rho_0}(.)$. To do so, we resort to the IS concept as detailed subsequently.

To begin with, we mention that IS is a powerful Monte-Carlo technique that allows generating realizations using another distribution which is simpler than the actual one. Through the IS technique, the generated samples are weighted and averaged in a judicious manner to obtain the desired ML estimates. It is this efficient weighting operation that improves considerably the performance achieved by the IS method compared to other Monte-Carlo techniques. The IS approach is based on the following simple observation:

$$\int_J \cdots \int_J f(\boldsymbol{\tau})\mathcal{L}'_{c,\rho_0}(\boldsymbol{\tau})d\boldsymbol{\tau} = \int_J \cdots \int_J f(\boldsymbol{\tau})\frac{\mathcal{L}'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})}g'(\boldsymbol{\tau})d\boldsymbol{\tau},$$
(34)

where $g'(.)$ is another pseudo-PDF called normalized *importance function* (IF), whose choice is discussed later and $f(.)$ is any given parameter transformation. Owing to (34), the problem can be recast into the computation of the expectation of $f(\boldsymbol{\tau})\frac{\mathcal{L}'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})}$ with respect to the distribution $g'(.)$, which is again simply performed via Monte-Carlo methods as follows:

$$\int_J \cdots \int_J f(\boldsymbol{\tau})\frac{\mathcal{L}'_{c,\rho_0}(\boldsymbol{\tau})}{g'(\boldsymbol{\tau})}g'(\boldsymbol{\tau})d\boldsymbol{\tau} \approx \frac{1}{R}\sum_{r=1}^R f(\boldsymbol{\tau}_r)\frac{\mathcal{L}'_{c,\rho_0}(\boldsymbol{\tau}_r)}{g'(\boldsymbol{\tau}_r)},$$
(35)

in which the realizations $\{\boldsymbol{\tau}_r\}_{r=1}^R$ are now generated according to $g'(.)$ and not $\mathcal{L}'_{c,\rho_0}(.)$. Yet, careful attention should be paid to the choice of $g'(.)$. In fact, the accuracy of the IS approach depends on the similarity between the shapes of $\mathcal{L}'_{c,\rho_0}(.)$ and $g'(.)$. In the best cases, the global maxima of $\mathcal{L}'_{c,\rho_0}(.)$ and $g'(.)$ are the same. Still, $\mathcal{L}'_{c,\rho_0}(.)$ is a complicated function of $\boldsymbol{\tau}$ and $g'(.)$ must be as simple as possible to easily generate the required realizations. Moreover, an appropriate choice of $g'(.)$ reduces the number R of realizations since the generated values will appear as if they were generated according to the original pseudo-PDF $\mathcal{L}'_{c,\rho_0}(.)$ when $g'(.)$ is closer to $\mathcal{L}'_{c,\rho_0}(.)$. Therefore, some trade-offs must be made in the construction of the *importance function* which is now discussed in some depth.

First, as an alternative to the actual multidimensional CLF, the *importance function* should be a separable function in the different delays, $\{\tau_p\}_{p=1}^P$, to reduce the generation of a P-dimensional random vector to the generation of P scalar random variables. We therefore simplify the expression of the CLF $\mathcal{L}_c(.)$ to find $g'(.)$. Indeed, it is seen from (29) that $\mathcal{L}_c(.)$ involves the sum of two independent terms. These two terms can be expressed in terms of the eigenvalues of $\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\boldsymbol{\Pi}$ as follows:

$$\ln\left(\det\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\boldsymbol{\Pi} + \sigma^2(\boldsymbol{I}_L - \boldsymbol{\Pi})\right\}\right)$$
$$= \ln\left((\sigma^2)^{L-P}\prod_{p=1}^P \lambda_p\right)$$
$$= \sum_{p=1}^P \ln\left(\frac{\lambda_p}{\sigma^2}\right) + L\ln\sigma^2,$$
(36)

$$\frac{1}{\sigma^2}\text{trace}\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\right\}$$
$$= \frac{1}{\sigma^2}\text{trace}\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\boldsymbol{\Pi}\right\} = \sum_{p=1}^P \frac{\lambda_p}{\sigma^2},$$
(37)

where $\lambda_1, \ldots, \lambda_P$ are the eigenvalues of $\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\boldsymbol{\Pi}$. We drop the term $L\ln\sigma^2$ in (36) since it does not depend on the delays and neglect the term $\sum_{p=1}^P \ln\left(\frac{\lambda_p}{\sigma^2}\right)$ in front of $\sum_{p=1}^P \frac{\lambda_p}{\sigma^2}$. Moreover, one can approximate the matrix $\boldsymbol{D}(\boldsymbol{\tau})^H\boldsymbol{D}(\boldsymbol{\tau})$ by the diagonal matrix $\left(\sum_{l=0}^{L-1}|c_l|^2\right)\boldsymbol{I}_P$ to avoid the computation of the inverse involved in $\boldsymbol{\Pi}$. This approximation is well justified since the off-diagonal terms of the matrix $\boldsymbol{D}(\boldsymbol{\tau})^H\boldsymbol{D}(\boldsymbol{\tau})$ are negligible compared to its diagonal elements (see [21] for further details). Using this assumption, the term trace $\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\right\}$ is approximated by:

$$\text{trace}\left\{\boldsymbol{\Pi}\widehat{\boldsymbol{R}}_z\right\} \approx \frac{1}{\left(\sum_{l=0}^{L-1}|c_l|^2\right)}\text{trace}\left\{\boldsymbol{D}(\boldsymbol{\tau})^H\widehat{\boldsymbol{R}}_z\boldsymbol{D}(\boldsymbol{\tau})\right\}$$
$$\approx \text{trace}\left\{\boldsymbol{D}(\boldsymbol{\tau})^H\widehat{\boldsymbol{R}}_z\boldsymbol{D}(\boldsymbol{\tau})\right\},$$
(38)

where the sum $\sum_{l=0}^{L-1}|c_l|^2$ is approximated by 1 when the shaping pulses are normalized. Lastly, considering all these observations, an approximate expression, $\overline{\mathcal{L}}(.)$, for the original CLF, $\mathcal{L}_c(.)$, with the unnecessary terms discarded, is given by:

$$\overline{\mathcal{L}}(\boldsymbol{\tau})$$
$$= \frac{1}{\sigma^2}\text{trace}\left\{\boldsymbol{D}(\boldsymbol{\tau})\boldsymbol{D}(\boldsymbol{\tau})^H\widehat{\boldsymbol{R}}_z\right\}$$
$$= \frac{1}{MN\sigma^2}\sum_{k=1}^{MN}\sum_{p=1}^P\left|\sum_{l=1}^L c_{l-1}\exp\left\{-\frac{j2\pi\ (l-1)\tau_p}{L}\right\}[\boldsymbol{Z}]_{k,l}\right|$$
$$= \sum_{p=1}^P I(\tau_p),$$
(39)

where

$$I(\tau) = \frac{1}{MN\sigma^2}\sum_{k=1}^{MN}\left|\sum_{l=1}^L c_{l-1}\exp\left\{-\frac{j2\pi\ (l-1)\tau}{L}\right\}[\boldsymbol{Z}]_{k,l}\right|,$$
(40)

can be evaluated using the FFT of $[\boldsymbol{Z}]_{k,l}$. Hence, from (39), the normalized IF is selected as:

$$g'_{\rho_1}(\boldsymbol{\tau}) = \frac{\prod_{p=1}^P \exp\left\{\rho_1 I(\tau_p)\right\}}{\left(\int_J \exp\left\{\rho_1 I(\tau)\right\}d\tau\right)^P},$$
(41)

which is the product of P elementary functions, each one depends on the single delay of a given single path. Here, we succeeded in finding a normalized IF for which the different delays are separable and identically distributed according to the same pseudo-PDF $p(.)$ given by:

$$p(\tau) = \frac{\exp\left\{\rho_1 I(\tau)\right\}}{\int_J \exp\left\{\rho_1 I(\tau)\right\}d\tau}.$$
(42)

Note here that the joint PDF of the delays in $g'_{\rho_1}(.)$ is split into the product of P individual PDFs. This transforms the problem of generating a P-dimensional random variable into
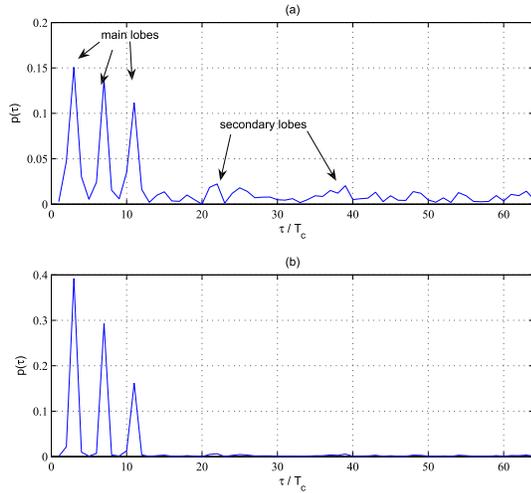
Fig. 1. Plot of $p(.)$ at SNR = 10 for (a) $\rho_1' = 5$ and (b) $\rho_1' = 10$.

the generation of P one-dimensional random variables according to a simpler common distribution. Moreover, the constant term ρ_1 in (41) and (42) is different from ρ_0 since it is more advantageous to use two different values as will be explained shortly. Yet, the appropriate selection of both parameters is critical to the performance behaviour of of the proposed IS-based estimator. In fact, the pseudo-PDF $p(.)$ in (42) exhibits P main lobes centered at the location of the true time delays. But the additive noise $\mathcal{N}$ (in (5)) makes other undesired lobes appear which in turn biases the generated values making them not faithful to the true delays. For this reason, ρ_1 must be increased to rid the pseudo-PDF in (42) of all the undesired lobes. However, we have noticed that very large values of ρ_1 may also destroy some useful lobes and hence their corresponding delays will not be generated. Therefore, the optimal value of ρ_1 is the highest one for which the PDF $p(.)$ still exhibits at least P main lobes. This behaviour is illustrated in Fig. 1 where we plot the pseudo-PDF $p(.)$ for two values of ρ_1. Moreover, one should keep in mind that the normalized IF in (41) is built upon an approximation of the original CLF, which we intend to maximize. Consequently, a bias will always appear in the mean of the values generated according to the normalized IF. Fortunately, this bias is alleviated by the weighting factor $\mathcal{L}_{c,\rho_0}'(.)/g_{\rho_1}'(.)$ introduced by the IS concept. Therefore, we maximize the contribution of the actual CLF in the weighting factor rather than its approximation by choosing ρ_0 higher than ρ_1. In conclusion, an appropriate choice of these two parameters reduces the number, R, of required realizations and ultimately the computational complexity. In this work we assume that the number of paths is exactly known. The problem of finding the number of paths is beyond the scope of this work.

To summarize, our new IS-based ML TDE is given by:

$$\widehat{\tau}_p = \frac{1}{R} \sum_{r=1}^{R} \boldsymbol{\tau}_r(p) \frac{\mathcal{L}_{c,\rho_0}'(\boldsymbol{\tau}_r)}{g_{\rho_1}'(\boldsymbol{\tau}_r)}, \qquad (43)$$

where $\boldsymbol{\tau}_r(p)$ is the pth element of the vector $\boldsymbol{\tau}_r$. We also mention that, in practice, the delays are often confined to the

interval $[0, LT_c]$ [13]. Since the delays are bounded, then there is room to further use the *circular mean* instead of the *linear mean* given by (43), as detailed below.

To define the concept of the *circular mean*, consider a random variable X taking values in the interval $[0, 1]$ according to a given distribution $f_X(.)$. The circular mean of X is [21]:

$$\mathbb{E}_c\{X\} = \frac{1}{2\pi} \angle \int_0^1 e^{j2\pi x} f_X(x) dx, \qquad (44)$$

where $\angle(.)$ returns the argument of a complex number. Then, if we have a set of R realizations, $x_1, \ldots, x_R$, drawn according to the PDF $f_X(.)$, the circular mean in (44) is computed as:

$$\mathbb{E}_c\{X\} = \frac{1}{2\pi} \angle \frac{1}{R} \sum_{r=1}^{R} e^{j2\pi x_r}. \qquad (45)$$

Thus, an alternative formulation of the estimate in (43) is given by:

$$\widehat{\tau}_p = \frac{LT_c}{2\pi} \angle \frac{1}{R} \sum_{r=1}^{R} F(\boldsymbol{\tau}_r) \exp\left\{ \frac{j2\pi \boldsymbol{\tau}_r(p)}{LT_c} \right\}, \qquad (46)$$

where the delays are transposed to the interval $[0, 1]$ after being normalized by LT_c and $F(.)$ is the weighting factor defined as:

$$F(\boldsymbol{\tau}_r) = \frac{\mathcal{L}_{c,\rho_0}'(\boldsymbol{\tau}_r)}{g_{\rho_1}'(\boldsymbol{\tau}_r)}. \qquad (47)$$

Note that the estimator in (46) relies on finding the angles of a complex number. Therefore, we no longer need to compute the two positive real-valued normalization factors $\int_J \ldots \int_J \exp\{\rho \mathcal{L}_c(\boldsymbol{\tau})\} d\boldsymbol{\tau}$ and $\left(\int_J \exp\{\rho_1 I(\tau)\} d\tau \right)^P$ since they can be dropped without affecting the final result. Also, when computing the weighting factor $F(\boldsymbol{\tau}_k)$, the exponential terms in his numerator and denominator may result in an overflow. To avoid this overflow, we substitute $F(.)$ by $F'(.)$ as:

$$F'(\boldsymbol{\tau}_k) = \exp\left\{ \rho_0 \mathcal{L}_c(\boldsymbol{\tau}_k) - \rho_1 \sum_{p=1}^{P} I(\boldsymbol{\tau}_k(p)) \right.$$
$$\left. - \max_{1 \le l \le R} \left(\rho_0 \mathcal{L}_c(\boldsymbol{\tau}_l) - \rho_1 \sum_{p=1}^{P} I(\boldsymbol{\tau}_l(p)) \right) \right\}, \quad (48)$$

by multiplying $F(.)$ by a positive number. In such a way, the exponential argument in (48) no longer exceeds zero, alleviating thereby any computation overflow.

The entire steps of the new IS-based ML time delay estimator are summarized in Table II.

III. EXTENSIONS TO THE MC AND MIMO NDA CASES

Without lack of generality, in a SIMO MC-DS-CDMA transmitter, the original data are spread over different subcarriers. Therefore, it is possible to transmit several DS-CDMA waveforms in parallel. At time index n, the input information is first converted into $N_c = 2K + 1$ parallel sequences and modulated at a rate $1/T_{MC}$, where $T_{MC} = N_c T_c$ is the symbol duration after serial to parallel conversion. Each of the parallel streams are then spread with a spreading code

TABLE II
SUMMARY OF STEPS OF THE IS-BASED ESTIMATOR

1) From the samples matrix, $\mathbf{Z}$, compute the periodogram $I(.)$ in (40) at discrete points $\{\tau_i\}_{i=1}^{K}$ of the interval $[0, LT_c]$. Then evaluate the elementary PDF as follows: $p(\tau_i) = \frac{\exp(\rho_1 \ I(\tau_i))}{\sum_{k=1}^{K} \exp(\rho_1 \ I(\tau_k))}, \ i = 1, \ 2, \dots, \ K,$ where we substitute the integral in the denominator by a summation over all the discrete points in the integration interval.
2) Generate one realization of the vector $\boldsymbol{\tau}$ according to the pseudo-PDF $g'_{\rho_1}(.)$. To simplify, we exploit the fact that the delays are separable in $g'_{\rho_1}(.)$ and we generate P realizations $\{\boldsymbol{\tau}_r(p)\}_{p=1}^{P}$ according to $p(.)$ using the inverse probability integration [26]. It is important to make sure that the P generated entries $\boldsymbol{\tau}_r(1), \dots, \boldsymbol{\tau}_r(P)$ are different. This condition is necessary since the delays are different in practice ensuring thereby that the matrix inverse $[\boldsymbol{\mathcal{D}}(\boldsymbol{\tau})^H \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})]^{-1}$ in $\mathcal{L}_c(.)$ always exists.
3) Repeat *step 2)* $R-1$ times then evaluate the weighting factors $F'(\boldsymbol{\tau}_r)$ for $r = 1, \dots, \ R$.
4) Find the maximum likelihood estimate of the delays using the *circular mean* in (46).

at a rate $1/T_c$ and modulated by the inverse discrete Fourier transform (IDFT).

At the receiver side, the post-correlation model for MC-DS-CDMA of the spatio-temporal observation at the kth subcarrier and the nth observation interval is given by [22]:

$$\boldsymbol{Z}_{k,n} = s_{k,n} \boldsymbol{J}_{k,n} \boldsymbol{D}_k^T(\boldsymbol{\tau}) + \boldsymbol{N}_{k,n}, \qquad (49)$$

where $s_{k,n}$ and $\boldsymbol{J}_{k,n}$ are the signal component and the spatial response matrix on the kth subcarrier. The column of the time response matrix $\boldsymbol{D}_k(\boldsymbol{\tau}) = [\mathbf{d}_{k,1}, \ \mathbf{d}_{k,2}, \dots, \ \mathbf{d}_{k,P}]$ are given by:

$$\begin{aligned}
\mathbf{d}_{k,p} = \ &e^{-j2\pi k\lambda \frac{\tau_p}{T_{MC}}} \\
&\times \Big[\rho_c(-\tau_p), \ \rho_c(T_c/k_s - \tau_p)e^{j2\pi k\frac{\lambda}{Lk_s}}, \dots, \\
&\qquad \rho_c((Lk_s - 1)T_c/k_s - \tau_p)e^{j2\pi k\frac{\lambda(Lk_s-1)}{Lk_s}}\Big]^T, \quad (50)
\end{aligned}$$

where λ determines the frequency spacing between two adjacent subcarriers[6] ($f_k = \lambda k/T_{MC}$) and k_s is the oversampling ratio [22]. Note here that the propagation time-delays are supposed to be the same for all subcarriers [22].

A. Extension to the MC NDA Case

Unlike the single-carrier case, the received samples $\boldsymbol{Z}_{k,n}$ cannot be directly used as an input to the algorithms because of the presence of the exponential terms $e^{j2\pi k\frac{\lambda l}{Lk_s}}$ in the elements of the vector $\mathbf{d}_{k,p}$. Therefore, we introduce the intermediate transformation of the samples matrix, denoted $\dot{\boldsymbol{Z}}_{k,n}$, given by:

$$\dot{\boldsymbol{Z}}_{k,n} = \boldsymbol{Z}_{k,n} \odot \big(\boldsymbol{a} \ \mathbf{1}_M^T\big) = s_{k,n} \boldsymbol{J}_{k,n} \dot{\boldsymbol{D}}_k^T(\boldsymbol{\tau}) + \dot{\boldsymbol{N}}_{k,n}, \quad (51)$$

where $\boldsymbol{a} \triangleq \big[1, \ e^{-j2\pi \frac{\lambda}{Lk_s}}, \dots, \ e^{-j2\pi \frac{\lambda \ (Lk_s-1)}{Lk_s}}\big]^T$ and $\mathbf{1}_M \triangleq [1, \dots, 1]^T$. The pth column of $\dot{\boldsymbol{D}}_k(\boldsymbol{\tau})$ is:

$$\begin{aligned}
\dot{\mathbf{d}}_{k,p} = \ &\mathbf{d}_{k,p} \odot \boldsymbol{a} \\
= \ &e^{-j2\pi k\lambda \frac{\tau_p}{T_{MC}}} \big[\rho_c(-\tau_p), \ \rho_c(T_c/k_s - \tau_p), \dots, \\
&\qquad \rho_c((Lk_s - 1)T_c/k_s - \tau_p)\big]^T. \quad (52)
\end{aligned}$$

[6] The transceiver belongs to the class of MT-CDMA or MC-DS-CDMA when λ is set to 1 or L, respectively.

Hence, in the spectral domain, we eliminate in $\dot{\boldsymbol{D}}_k(\boldsymbol{\tau})$ the row-wise dependence of the phase slope of each column vector on k. While the formulation in (51) seems to be adapted to our estimation process, the phase shift $e^{-j2\pi k\lambda \frac{\tau_p}{T_{MC}}}$ in each column of $\dot{\boldsymbol{D}}_k(\boldsymbol{\tau})$ prevents us from using directly the signal structure in (51). To overcome this problem, we note that the matrix $\dot{\boldsymbol{D}}_k(\boldsymbol{\tau})$ can be written as $\dot{\boldsymbol{D}}_k(\boldsymbol{\tau}) = \boldsymbol{D}(\boldsymbol{\tau})A_k$ where A_k is a $P \times P$ diagonal matrix whose diagonal elements are $\left\{e^{-j2\pi k\lambda \frac{\tau_p}{T_{MC}}}\right\}_{p=1}^{P}$. Hence, we transfer this matrix into the spatial-response matrix $\dot{\boldsymbol{J}}_{k,n}^T = A_k \boldsymbol{J}_{k,n}^T$ that preserves the statistical properties of $\boldsymbol{J}_{k,n}$ to obtain a structure similar to the one in (3). Finally, to optimally exploit gains from the frequency dimension, we gather all the transformed observations over the different subcarriers into the following compact representation:

$$\dot{\boldsymbol{Z}}_n \triangleq \Big[\dot{\boldsymbol{Z}}_{1,n}^T, \ \dot{\boldsymbol{Z}}_{2,n}^T, \dots, \ \dot{\boldsymbol{Z}}_{N_c,n}^T\Big] = \boldsymbol{D}(\boldsymbol{\tau})\dot{\boldsymbol{J}}_n^T + \dot{\boldsymbol{N}}_n^T, \quad (53)$$

where $\dot{\boldsymbol{J}}_n^T \triangleq \Big[A_1 \boldsymbol{J}_{1,n}^T, \ A_2 \boldsymbol{J}_{2,n}^T, \dots, \ A_{N_c} \boldsymbol{J}_{N_c,n}^T\Big]$ and $\dot{\boldsymbol{N}}_n^T \triangleq \Big[\dot{\boldsymbol{N}}_{1,n}^T, \ \dot{\boldsymbol{N}}_{2,n}^T, \dots, \ \dot{\boldsymbol{N}}_{N_c,n}^T\Big]$. The interesting feature with the formulation in (53) is that it increases the number of observations proportionally to the number of subcarriers. Considering N received symbols, we obtain a compact representation similar to (4) by concatenating the N observed symbols as:

$$\dot{\boldsymbol{Z}} \triangleq \big[\dot{\boldsymbol{Z}}_1, \ \dot{\boldsymbol{Z}}_2, \dots, \ \dot{\boldsymbol{Z}}_N\big] = \boldsymbol{D}(\boldsymbol{\tau})\dot{\boldsymbol{J}}^T + \dot{\boldsymbol{N}}^T, \quad (54)$$

where $\dot{\boldsymbol{J}}^T \triangleq \Big[\dot{\boldsymbol{J}}_1^T, \dots, \ \dot{\boldsymbol{J}}_N^T\Big]$ and $\dot{\boldsymbol{N}}^T \triangleq \Big[\dot{\boldsymbol{N}}_1^T, \dots, \ \dot{\boldsymbol{N}}_N^T\Big]$. Similar to the SC case, we perform a column-by-column FFT of $\dot{\boldsymbol{Z}}$ to obtain:

$$\dot{\boldsymbol{z}} = \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\dot{\boldsymbol{J}}^T + \dot{\boldsymbol{\mathcal{N}}}^T. \quad (55)$$

Then, using (55), the extension of the two proposed SC algorithms to MC-DS-CDMA is straightforward by substituting $\boldsymbol{Z}$ in (5) by $\dot{\boldsymbol{z}}$ in (55). Compared to a scheme that estimates the delays over each subcarrier separately, the proposed model offers better performance since it exploits the information carried by all the subcarriers jointly and not independently.

B. Extension to the MIMO NDA Case

Although we have derived our TDEs for the SIMO case, they can be extended to MIMO systems. Indeed, if we assume M_{Tx} co-located transmit antennas characterized each by its own spreading code[7] and M_{Tx} sources s_k for $k = 1, \ 2, \dots, M_{\text{Tx}}$ that stem from any given mixture of other M_{Tx} sources s'_k for $k = 1, 2, \dots, M_{\text{Tx}}$ to translate pure transmit diversity (i.e., $s_k = s'_1$), pure transmit multiplexing (i.e., $s_k = s'_k$), or any combination thereof (e.g., unitary precoding with $[s_1, \ s_2, \dots, s_{M_{\text{Tx}}}]^T = \boldsymbol{U}^H[s'_1, s'_2, \dots, s'_{M_{\text{Tx}}}]^T$ where $\boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I}_{M_{\text{Tx}}}$), then (55) also holds in the MIMO case:

$$\boldsymbol{z}_k = \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})\boldsymbol{J}_k^T + \boldsymbol{\mathcal{N}}_k^T, \quad (56)$$

[7] The quasi-orthogonality between codes allow us to extract each of the source signals from any given transmit antenna, thereby, allowing us to treat them separately.

where k now stands for the transmit antenna index ranging from 1 to M_{Tx}. And if we stack all the M_{Tx} matrices in (56) into one matrix $\boldsymbol{Z}_{\text{MIMO}} \triangleq [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{M_{\text{Tx}}}]$, then, we can substitute $\boldsymbol{Z}$ in (5) by $\boldsymbol{Z}_{\text{MIMO}}$ and thereby extend the two proposed TDEs to the MIMO configuration in both the SC-DS-CDMA and MC-DS-CDMA cases.

C. The CRLB

We are able to derive the corresponding CRLB following the same steps presented in Section II. This reveals that the CRLB has a similar expression as in (11) scaled by a factor of $1/N_c$. The resulting CRLB in the MIMO MC DSSS NDA case is given by:

$$
\begin{aligned}
\text{CRLB} &= \frac{\sigma^2}{2MNN_c} \\
&\times \left[\Re\left\{ \left(\boldsymbol{U}^H \boldsymbol{\Pi}^\perp \boldsymbol{U} \right) \odot \left(\boldsymbol{R_J} \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})^H \boldsymbol{R}_{\dot{\boldsymbol{Z}}}^{-1} \boldsymbol{\mathcal{D}}(\boldsymbol{\tau}) \boldsymbol{R_J} \right)^T \right\} \right]^{-1} \\
&= \frac{1}{M_{\text{Tx}} M N N_c} \text{CRLB}_0,
\end{aligned} \tag{57}
$$

where $\boldsymbol{R_J}$ and $\boldsymbol{R}_{\dot{\boldsymbol{Z}}}$ are the covariance matrices of $\dot{\boldsymbol{J}}$ and $\dot{\boldsymbol{Z}}$, respectively, and CRLB_0 is the CRLB when $M_{\text{Tx}} = M = N = N_c = 1$. As a consequence, the TDE in MC-DS-CDMA systems merges space, time and frequency dimensions whereby spatial (both transmit or receive), temporal and frequency samples exactly the very same impact on estimation performance regardless of any channel correlation type and amount in each dimension (due to the observation's randomization with the symbols' quasi-independence across all dimensions). It can be verified from the expression of the CRLB in (57) that the overall achievable performance is inversely proportional to the product of the numbers of both receive and transmit antennas, received symbols, and subcarriers.

IV. EXTENSION TO THE DA CASE

In this section, we extend the two proposed ML TDEs and the CRLB established in all previous scenarios to the DA case. In the previous NDA case, the columns of the observation matrix are uncorrelated due to the presence of uncorrelated transmitted symbols. In the DA case, however, these columns become correlated due to space, time, and/or frequency channel correlations whose proper incorporation in the estimation process and in the CRLB derivation is extremely challenging in practice. In the following, we are able to properly cope with these correlations by following a new approach that is completely different from previous works. Next, we focus on MC DSSS RITs. The special case of SC DSSS can be obtained by setting N_c equal to 1. And extension to MIMO with any diversity-multiplexing configuration follows easily the very same lines as above.

A. The ML TDEs

Considering again the formulation in (49), $\boldsymbol{Z}_{k,n}$ can be written as follows:

$$
\boldsymbol{Z}_{k,n} = \boldsymbol{H}_{k,n} s_{k,n} + \boldsymbol{N}_{k,n}, \tag{58}
$$

in which $\boldsymbol{H}_{k,n} = \boldsymbol{J}_{k,n} \boldsymbol{D}^T(\boldsymbol{\tau})$ denotes the overall spatio-temporal propagation matrix. Interestingly, the model in (58) can be used to estimate the channel response, $\boldsymbol{H}_{k,n}$, in an efficient way. One can use any blind channel estimator to obtain an estimate, $\widehat{\boldsymbol{H}}_{k,n}$, of $\boldsymbol{H}_{k,n}$ (more details can be found in [22]). Then, taking into account the estimation error, $\widehat{\boldsymbol{H}}_{k,n}$ is written as:

$$
\widehat{\boldsymbol{H}}_{k,n}^T = \boldsymbol{D}(\boldsymbol{\tau}) \boldsymbol{J}_{k,n}^T + \boldsymbol{E}_{k,n}^T, \tag{59}
$$

where $\boldsymbol{E}_{k,n}^T$ is the channel estimation error matrix. The variance of the entries of $\boldsymbol{E}_{k,n}^T$ depends of course on the noise variance of the received signal [24] where the power of $\boldsymbol{E}_{k,n}$ is lower than the power of $\boldsymbol{N}_{k,n}$ due to the SNR gain stemming from channel identification. Note here that the matrix $\boldsymbol{J}_{k,n}$ no longer contains the transmitted symbols $s_{k,n}$ thereby resulting in a DA scenario.[8] Clearly, the time delays can be estimated from the column-by-column FFT of $\widehat{\boldsymbol{H}}_{k,n}^T$ in (59) as well as from (54), with the only difference that the noise power is reduced in (59).

In the NDA case, the columns of the observation matrix are uncorrelated due to the presence of uncorrelated transmitted symbols. However, when the channel-coefficients matrix estimate $\widehat{\boldsymbol{H}} \triangleq \left[\widehat{\boldsymbol{H}}_{1,1}^T, \ldots, \widehat{\boldsymbol{H}}_{N_c,1}^T, \ldots, \widehat{\boldsymbol{H}}_{1,N}^T, \ldots, \widehat{\boldsymbol{H}}_{N_c,N}^T \right]$ is directly used in the estimation process, these symbols are no longer present and we are actually estimating the delays from the inherently correlated columns of $\widehat{\boldsymbol{H}}$. Therefore, we slightly modify the signal model to keep the two algorithms valid. First, we perform a column-by-column FFT of $\left\{ \widehat{\boldsymbol{H}}_{k,n}^T \right\}$ to obtain:

$$
\begin{aligned}
\widehat{\boldsymbol{\mathcal{H}}}_{k,n} &= \boldsymbol{\mathcal{D}}(\boldsymbol{\tau}) \dot{\boldsymbol{J}}_{k,n}^T + \boldsymbol{\mathcal{E}}_{k,n} \\
&= \boldsymbol{\mathcal{D}}(\boldsymbol{\tau}) \left[\boldsymbol{g}_{k,n}(1), \boldsymbol{g}_{k,n}(2), \ldots, \boldsymbol{g}_{k,n}(M) \right] + \boldsymbol{\mathcal{E}}_{k,n}, \quad (60)
\end{aligned}
$$

where $\boldsymbol{g}_{k,n}(m)$ is the mth column of $\dot{\boldsymbol{J}}_{k,n}^T$. Then we stack the channel coefficients in the matrix $\widehat{\boldsymbol{\mathcal{H}}}$ given by:

$$
\widehat{\boldsymbol{\mathcal{H}}} = \overline{\boldsymbol{\mathcal{D}}}(\boldsymbol{\tau}) \left[\boldsymbol{g}(1), \boldsymbol{g}(2), \ldots, \boldsymbol{g}(M) \right] + \boldsymbol{\mathcal{E}}, \tag{61}
$$

with $\overline{\boldsymbol{\mathcal{D}}}(\boldsymbol{\tau}) = \boldsymbol{I}_N \otimes \boldsymbol{I}_{N_c} \otimes \boldsymbol{\mathcal{D}}(\boldsymbol{\tau})$ where the operator $\otimes$ stands for the Kronecker product and $\boldsymbol{g}(m) \triangleq \left[\boldsymbol{g}_{1,1}^T(m), \ldots, \boldsymbol{g}_{N_c,1}^T(m), \boldsymbol{g}_{1,2}^T(m), \ldots, \boldsymbol{g}_{N_c,2}^T(m), \ldots, \boldsymbol{g}_{1,N}^T(m), \ldots, \boldsymbol{g}_{N_c,N}^T(m) \right]^T$. Here, the two algorithms can be applied to the observation matrix $\widehat{\boldsymbol{\mathcal{H}}}$ in (61) instead[9] of $\dot{\boldsymbol{Z}}$ in (55).

B. The CRLB

Denote the autocorrelation of the channel transfer function as $\phi(\Delta f, \Delta t)$. Here, we consider uncorrelated scattering where this autocorrelation across subcarriers is a function of the frequency difference, Δf, only [30]. The covariance matrix of $\boldsymbol{g}(m)$ is hence $\boldsymbol{R_g} = \boldsymbol{\Phi} \otimes \boldsymbol{R_J}$ where the elements of $\boldsymbol{\Phi}$

---

[8]Alternatively, we can address the DA TDE case without intermediate channel identification directly from the observation $\boldsymbol{Z}_{k,n} = \boldsymbol{H}_{k,n} + \boldsymbol{N}_{k,n}$ with $s_{k,n}$ set to 1 (i.e., reference or pilot signal).

[9]Extension to the MIMO case is straightforward as we only need to substitute $\overline{\boldsymbol{\mathcal{D}}}$ by $\boldsymbol{I}_{M_{\text{Tx}}} \otimes \boldsymbol{I}_N \otimes \boldsymbol{I}_{N_c} \otimes \boldsymbol{\mathcal{D}}$ and increase $\boldsymbol{g}(m)$ to include all the channel coefficients resulting from the M_{Tx} transmit antennas.

are function of $\phi(\Delta f, \Delta t)$. Injecting $\boldsymbol{R}_g$ and $\overline{\mathcal{D}}(\boldsymbol{\tau})$ in (11), the CRLB can be written in the following alternative form:

$$
\begin{aligned}
&\left[\mathrm{CRLB}^{-1}(\boldsymbol{\tau})\right]_{i,j} \\
&= \frac{2M}{\sigma^2} \Re\left\{\mathrm{trace}\left\{\overline{\boldsymbol{D}}_j^H(\boldsymbol{\tau})(\boldsymbol{I}_{LNN_c} - \boldsymbol{I}_N \otimes \boldsymbol{I}_{N_c} \otimes \boldsymbol{\Pi})\right.\right. \\
&\quad \left.\left. \times \overline{\boldsymbol{D}}_i(\boldsymbol{\tau})\left(\boldsymbol{R}_g\overline{\boldsymbol{D}}^H(\boldsymbol{\tau})\boldsymbol{R}_{\mathcal{H}}^{-1}\overline{\boldsymbol{D}}(\boldsymbol{\tau})\boldsymbol{R}_g\right)^T\right\}\right\} \\
&= \frac{2M}{\sigma^2}\Re\left\{\mathrm{trace}\left\{\boldsymbol{I}_N \otimes \boldsymbol{I}_{N_c} \otimes \left(\boldsymbol{D}_j^H(\boldsymbol{\tau})\boldsymbol{\Pi}^\perp\boldsymbol{D}_i(\boldsymbol{\tau})\right)\right.\right. \\
&\quad \left.\left. \times \left(\boldsymbol{R}_g\overline{\boldsymbol{D}}^H(\boldsymbol{\tau})\boldsymbol{R}_{\mathcal{H}}^{-1}\overline{\boldsymbol{D}}(\boldsymbol{\tau})\boldsymbol{R}_g\right)^T\right\}\right\},
\end{aligned}
\tag{62}
$$

where $\boldsymbol{R}_{\mathcal{H}}$ is the covariance of $\mathcal{H}$ and $\overline{\boldsymbol{D}}_i(\boldsymbol{\tau})$ and $\boldsymbol{D}_i(\boldsymbol{\tau})$ are the derivatives of $\overline{\boldsymbol{D}}(\boldsymbol{\tau})$ and $\boldsymbol{D}(\boldsymbol{\tau})$ with respect to $\boldsymbol{\tau}_i$, respectively. If we denote by $\boldsymbol{B}_k$ the kth ($P \times P$) diagonal block of $\left(\boldsymbol{R}_g\overline{\boldsymbol{D}}^H\boldsymbol{R}_{\mathcal{H}}^{-1}\overline{\boldsymbol{D}}\boldsymbol{R}_g\right)^T$, we obtain:

$$
\left[\mathrm{CRLB}^{-1}(\boldsymbol{\tau})\right]_{i,j} = \frac{2M}{\sigma^2}\Re\left\{\sum_{k=1}^{NN_c}\left(\boldsymbol{u}_j^H\boldsymbol{\Pi}^\perp\boldsymbol{u}_i\right)[\boldsymbol{B}_k]_{i,j}\right\}. \tag{63}
$$

And accounting for the fact[10] $\boldsymbol{B}_k = \boldsymbol{B}$ we obtain the following expression of the CRLB:

$$
\mathrm{CRLB}(\boldsymbol{\tau}) = \frac{1}{MNN_c}\mathrm{CRLB}_1(\boldsymbol{\tau}), \tag{64}
$$

where $\boldsymbol{U}$ and $\boldsymbol{u}_i$ are defined in (12) and $\mathrm{CRLB}_1(\boldsymbol{\tau}) = \Re\left\{(\boldsymbol{U}^H\boldsymbol{\Pi}^\perp\boldsymbol{U}) \odot \boldsymbol{B}\right\}$. Under certain conditions, $\boldsymbol{\Phi}$ can be easily expressed. In fact, $\phi(\Delta f, \Delta t)$ is given in [30] by:

$$
\phi(\Delta f, \Delta t) = \int_{-\infty}^{+\infty}\phi_c(\tau, \Delta t)e^{-j2\pi\Delta f\tau}d\tau, \tag{65}
$$

where $\phi_c(\tau, \Delta t)$ is the autocorrelation function of the channel impulse response. If we consider that the signal is transmitted through a Rayleigh channel, with the uniform Jakes spectrum, and if we denote the maximum Doppler frequency by f_D, $\phi(\Delta f, \Delta t)$ can be written as:

$$
\begin{aligned}
\phi(\Delta f, \Delta t) &= \int_{-\infty}^{+\infty}P(\tau)J_0(2\pi f_D\Delta t)e^{-j2\pi\Delta f\tau}d\tau \\
&= J_0(2\pi f_D\Delta t)F_f(\Delta f), \tag{66}
\end{aligned}
$$

where $J_0(.)$ is a zeroth-order Bessel function of the first kind and $F_f(\Delta f) = \int_{-\infty}^{+\infty}g(\tau)e^{-j2\pi\Delta f\tau}d\tau$ with $g(\tau)$ is the delay profile. Then $\boldsymbol{\Phi} = \boldsymbol{J}_0 \otimes \boldsymbol{F}_f$ where $[\boldsymbol{F}_f]_{i,j} = F_f((i-j)\lambda/T_{MC})$ and $[\boldsymbol{J}_0]_{i,j} = J_0(2\pi f_D(i-j)T)$.

V. SIMULATION RESULTS

In this section, we compare the performance of the two proposed ML estimators against the popular root-MUSIC algorithm and the newly derived CRLBs. Unless specified otherwise, we consider in all the simulations a multipath Rayleigh-fading channel with $f_DT = 0.01$ and 3 equal-power paths in a challenging scenario of closely-spaced delays set to 0.12 T, 0.15 T, and 0.18 T (i.e., $\boldsymbol{\tau} = [7.68\ T_c,\ 9.6\ T_c,\ 11.52\ T_c]$). The mean square error (MSE)
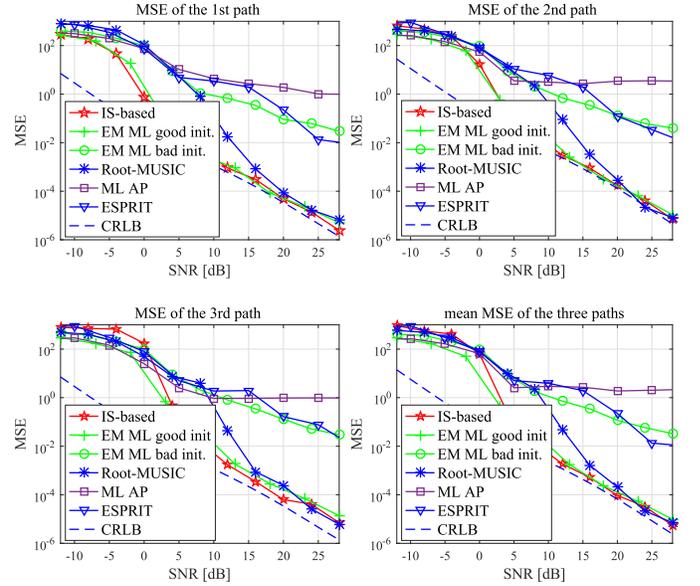

Fig. 2. MSE vs. SNR performance of the different algorithms with closely-spaced delays and $M = 4$ in the NDA case.

is adopted as our performance measure. First, recall that the EM algorithm is iterative in nature; hence its initialization is critical. Therefore, the initial values for this estimator are selected as random variables, centered at the true time delays with variance $(0.05\ T)^2$. The processing gain is fixed to $L = 64$ and the optimal values of the parameters ρ_0 and ρ_1 for the IS-based technique are set to 20 and 10, respectively. The number of realizations is $R = 100$ and the number of transmit antennas is fixed to $M_{\mathrm{Tx}} = 1$ (i.e., SIMO). The SNR is defined in the DA and NDA cases with respect to the powers of the channel estimation error and noise,[11] respectively.

First, we consider a SIMO SC transceiver with $M = 4$ receiving antenna branches and one received sample ($N = 1$) and compare in Fig. 2 the MSE of the two proposed ML algorithms to those of the root-MUSIC, ESPRIT and the ML alternating projection (ML AP) [32] algorithms in the NDA case. In the same figure, we plot the performance of the EM ML estimator when the initial values have a variance of $(0.18\ T)^2$ thereby accounting for less accurate initializations. Clearly, both the IS and the EM TDE, with good initialization, outperform root-MUSIC over a large range of SNR values. While the two ML algorithms exhibit almost the same performance, we advocate in practice the EM ML approach in this configuration since it entails less computational complexity than the IS TDE. Indeed, the EM ML estimator has the advantage of performing P parallel maximizations. Therefore, as the number of paths P increases, there is no additional noticeable increase in computation time. On the other hand, the IS TDE is a far more robust to initialization errors, in contrast to EM ML.

---

[10] $\boldsymbol{B}_k = \boldsymbol{B}$, for $k = 1, \ldots, NN_c$ is verified since the second-order statistics of the P paths are the same for all received signals.

[11] In a multi-user setup, signals received from other users very often contribute to the observation as additive noise after despreading. Yet if the need be, in near-far situations for instance, inter-user interference can be suppressed by one of the many multiuser detection techniques such as [31] which is well-adapted to the PCM model and, hence, can be easily operated jointly with our new ML TDEs.

TABLE III
COMPLEXITY ASSESSMENT OF THE ALGORITHMS

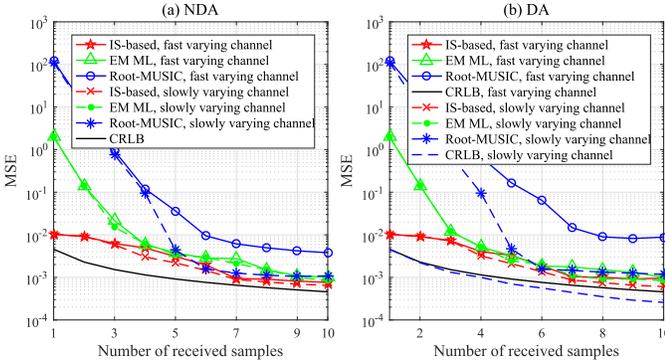| Algorithm	Complexity	Complexity ratio
Root-MUSIC	$O(L^3 + PL^2 + PL^2 M_{\text{Tx}} N M N_c K + L^2 K)$	0.73
EM ML	$O((L^3 + LP + L*log(N_s) + N_s + P*K*(L^2))*N_{iter})$	1
IS-based	$O(M_{\text{Tx}} M N N_c L K + R*(P^2 * L + P^3 + L^2))$	2.24
Grid-search	$O((L^3 + P^2 L + P^3)*K^P)$	10^4

Fig. 3. MSE vs. the number of symbols N with $M = 1$ antenna, $N_c = 1$ subcarrier, and SNR $= 10$ dB for fast- ($f_D T = 0.01$) and slowly-varying ($f_D T = 10^{-4}$) channels in the (a): NDA and (b): DA cases.

Various methods can be adopted to choose the initial guesses for the EM ML algorithm. In fact, one can simply use random initialization or apply a sub-optimal but low-cost algorithm such as the correlation-based algorithm that identifies the correlation maxima as initial estimates for the delays. On the other hand, the ESPRIT algorithm uses an estimate of the covariance matrix of the received signal from the columns of the matrix $\mathcal{Z}$, which needs a large number of received samples to obtain a good estimate of the covariance matrix.

In Fig. 2, we also compare the accuracy of the proposed estimators with the ML AP algorithm to show that the latter performs poorly compared to the proposed techniques, more so at higher SNR values. Therefore we keep root-MUSIC and the CRLB as benchmarks against which we gauge in the following the performance of our proposed ML-based estimators.

In Table III, we assess the complexity of the proposed IS-based and EM ML estimators and compare them to the other algorithms in terms of computational intensity. In this table, N_{iter} stands for the number of iterations of the EM ML and K stands for the number of discrete points in the interval $[0, LT_c]$ of the grid-search method. The third column of this table lists the achieved complexity ratios against EM ML taken as a reference using the specific parameter values $M_{\text{Tx}} = 1$; $M = 1$; $N = 1$; $N_c = 1$; $K = 64$; $L = 64$; $P = 3$; $R = 100$; $N_s = 64$; and $N_{iter} = 3$. The *grid-search* algorithm was included in the table for complexity comparisons purposes only, but will not be considered for assessment by simulations due to its prohibitive cost. Obviously, EM ML offers the best trade-off between complexity and accuracy if it is accurately initialized. Whereas the IS TDE is slightly more complex but always provides good performance regardless of initialization.
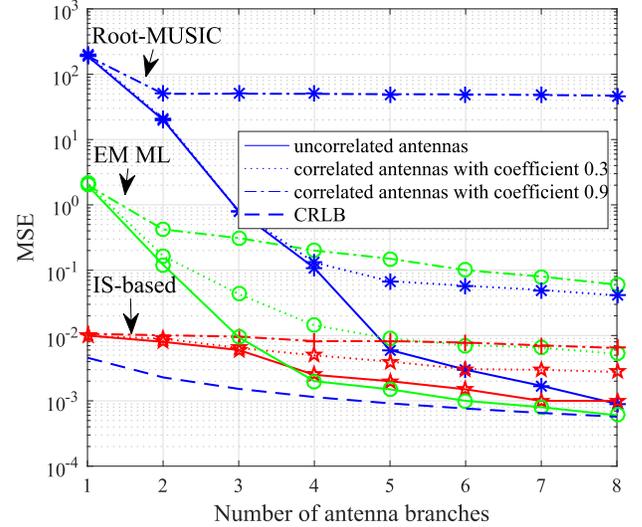

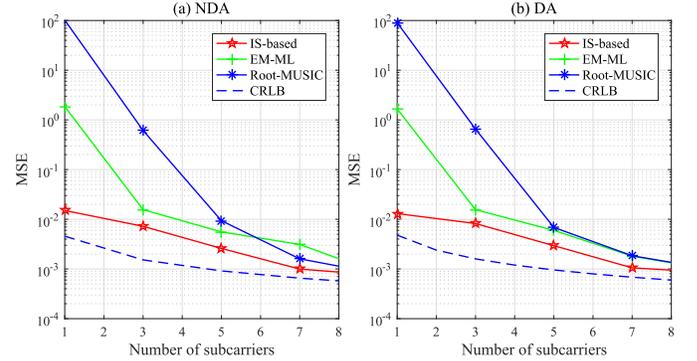Fig. 4. MSE vs. the number of antennas M with $N = 1$ symbol, $N_c = 1$ subcarrier, and SNR $= 10$ dB.

Fig. 5. MSE vs. the number of subcarriers N_c with $M = 1$ antenna, and $N = 1$ symbol, and SNR $= 10$ dB in the (a): NDA and (b): DA cases.

So far, we have seen that all the methods exhibit good performance, with a remarkable advantage for the two new ML estimators. However, a quick look at the EM and root-MUSIC algorithms reveals that they rely on an estimate of the covariance matrix of the received signal from the columns of the matrix $\mathcal{Z}$. The accuracy of this estimate depends on the number of spatial, temporal, and frequency snapshots. Therefore, we assess in Figs. 3 to 5 the performance of all these algorithms[12] where one of the three snapshot dimensions is changed while the others are kept constant and set to 1. We also fix the SNR to 10 dB. Compared to the previous case of Fig. 2, root-MUSIC is very sensitive to the number of snapshots and its performance degrades considerably over short data records. In fact, it fails completely in estimating the delays due to the very poor estimate of the covariance matrix $\boldsymbol{R}_{\mathcal{Z}}$. For the very same reason, the performance of the EM ML algorithm also deteriorates significantly. On the other hand, the new IS-based algorithm performs nearly the same in this challenging scenario and still provides good estimates

---

[12]In the following simulations, we illustrate only the MSE of the delay of the first path since the very same conclusions hold for the other paths.

with relatively few data snapshots. To assess the impact of channel time variations or correlation on performance, we consider now both the NDA and DA cases. We plot in Fig. 3 the MSE versus the number of received symbols N, considering only one receiving antenna branch. For the fast varying channel (i.e., short temporal correlation), the three estimators exhibit almost the same performance starting from $N = 5$. However, collecting more samples from a larger number of symbols (i.e., by increasing N) might not be useful since the values of the delays may change appreciably over large observation windows or even from one symbol to another. Usually, a tracking technique (as the one developed in [22]) is necessary to track the time variations of the delays. Hence, in the more practical region of small numbers of snapshots (i.e., $N < 5$), the new ML-based methods perform better than root-MUSIC and the advantage of the IS-based estimator becomes even more prominent as N decreases to cover the single-snapshot case. Therefore, we succeed here in developing an efficient technique, namely the IS-based that is able to reliably estimate the time delays from a single received symbol and thereby offer a very useful feature for real-time applications. In the NDA case, there is almost no-dependency on the channel time variations or correlation, as intuitively expected. In the DA case, when using the channel matrix estimate, the ML-based methods still perform well, whereas root-MUSIC saturates at $N = 8$ for slowly time varying channels.

To investigate the impact of spatial correlation, we assess in Fig. 4 the performance of the three estimators when the two adjacent antennas are correlated with correlation factors equal to 0.3 and 0.9. In this figure, we plot the MSE versus M from 1 to 8 with $N = 1$. As expected, increasing the number of antennas does not bring much improvement when the antennas are strongly correlated.

In order to evaluate the performance of the estimators in MC DSSS systems, we plot in Fig. 5 the MSE versus the number of subcarriers N_c at SNR = 10 dB in both the NDA and DA cases. We fix $M = 1$ and $N = 1$ to better isolate the sole impact of the number of subcarriers N_c on performance. As N_c increases, the estimation performance improves but then saturates at large values of N_c due to the increase of inter-carrier interference with N_c stemming from the loss of orthogonality between subcarriers in a multipath environment [22]. To better investigate the impact of frequency correlation between subcarriers, we plot in the same figure the MSE versus N_c using another delay profile, i.e., delays equal to 0.12 T, 0.13 T and 0.16 T (i.e., $\tau = [7.68\ T_c,\ 8.32\ T_c,\ 10.24\ T_c]$). Clearly, the estimation performance is almost the same in the DA and NDA cases, while the CRLB states that estimation error is a little bit lower in the former than in the latter. We also emphasize the similarity between Figs. 4, 3, and 5 which corroborates the fact that the three dimensions (time, space and frequency) have the same impact on the estimation performance of the algorithms. The IS-based estimator can reach an MSE accuracy as low as 10^{-2} with a single observation sample in time, space, and frequency and is best for data records of $MNN_c = 1$ to $MNN_c = 5$ combined samples in space, time and frequency.
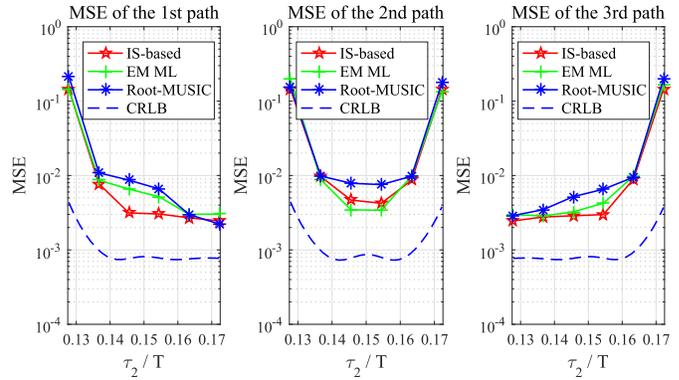

Fig. 6. MSE vs. τ_2/T with $M = 1$ antenna, $N = 1$, $N_c = 5$ subcarriers, and SNR = 10 dB in the DA case.

The EM ML algorithm is preferred in case of short data records MNN_c between 5 and 10 combined samples in space, time and frequency due its lower complexity. Root-MUSIC, however, requires data records MNN_c larger than 10 to present satisfactory.

To further investigate impact of frequency correlation across subcarriers (or equivalently the impact of the delay spacing), Fig. 6 depicts the estimation performance with $N_c = 5$ when the value of the delay of the second path (i.e., τ_2) varies between τ_1 and τ_3. As expected, the estimation performance becomes worse if any of the two delay spacing decreases.

VI. CONCLUSION

In this paper, we developed two new multi-path ML TDEs and derived their underlying estimation CRLBs in closed-form for both SC and MC DSSS RITs with either SISO, SIMO, or MIMO transceiver structures in both the NDA and DA cases. While the two proposed TDEs implement the ML criterion, each one has its own attractive features. The first TDE relies on the iterative EM concept with a substantially reduced computational cost compared to exhaustive grid-search techniques since it transforms the multidimensional grid-search problem into parallel easy searches over identical one-dimensional spaces. Compared to other eigen-based methods such as root-MUSIC, the EM approach exhibits better performance with a relatively good initialization, which also should be recognized as a limitation in EM TDE performance. The later also degrades considerably in the case of short data records (the product of the number of transmit or receive antennas, carriers, and symbols) where root-MUSIC fails completely to estimate the delays. The second ML TDE relies on the global maximization theorem and the IS concept to empirically find the global maximum of the CLF. Like EM, the IS TDE avoids any multidimensional grid-search by approximating the CLF and thereby splitting it into separable one-dimensional functions of the delays. In contrast to EM, however, without any initialization required, it always performs as well as the EM TDE — only when the latter is accurately initialized — at the expense of a moderate computational cost increase. Moreover, only the IS TDE is able to deliver estimates from very short data records. In the

NDA case, we revealed, both analytically and by simulations, that the three time, frequency and space (transmit and receive) dimensions interchangeably have exactly the very same impact on estimation performance regardless of any channel correlation type and amount in each one due to the observation's randomization with the symbols' quasi-independence across all of them. Besides, we properly cope with such channel correlations that do indeed arise in practice and, hence, become very challenging both in estimation and CRLB derivation in the DA case, but that have been so far overlooked in previous works.

REFERENCES

[1] A. Masmoudi, F. Bellili, and S. Affes, "Time delays estimation from DS-CDMA multipath transmissions using expectation maximization," in *Proc. IEEE VTC-Fall*, Quebec, QC, Canada, Sep. 2012, pp. 1–5.

[2] A. Masmoudi, F. Bellili, and S. Affes, "Maximum likelihood time delay estimation for direct-spread CDMA multipath transmissions using importance sampling," in *Proc. 45th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2011, pp. 618–623.

[3] W. Choi, J. G. Andrews, and R. W. Heath, Jr., "Multiuser antenna partitioning for cellular MIMO–CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2448–2456, Sep. 2007.

[4] D. N. Knisely, Q. Li, and N. S. Ramesh, "CDMA2000: A third-generation radio transmission technology," *Bell Labs Tech. J.*, vol. 3, no. 3, pp. 63–78, Jul. 1998.

[5] D. Benenati, P. M. Feder, N. Y. Lee, S. Martin-Leon, and R. Shapira, "A seamless mobile VPN data solution for CDMA2000,* UMTS, and WLAN users," *Bell Labs Tech. J.*, vol. 7, no. 2, pp. 143–165, 2002.

[6] Q. Sun, S. Wang, S. Han, and I. L. Chih, "Unified framework towards flexible multiple access schemes for 5G," *ZTE Commun.*, vol. 14, no. 4, pp. 26–34, Oct. 2016.

[7] K. Cheikhrouhou, S. Affes, and P. Mermelstein, "Impact of synchronization on performance of enhanced array-receivers in wideband CDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 12, pp. 2462–2476, Dec. 2001.

[8] S. Han, Y.-C. Liang, and B.-H. Soong, "Robust joint resource allocation for OFDMA-CDMA spectrum refarming system," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1291–1302, Mar. 2016.

[9] Y. Cai, R. C. de Lamare, and D. Le Ruyet, "Transmit processing techniques based on switched interleaving and limited feedback for interference mitigation in multiantenna MC-CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1559–1570, May 2011.

[10] J. He, M. O. Ahmad, and M. N. S. Swamy, "Joint space-time parameter estimation for multicarrier CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3306–3311, Sep. 2012.

[11] F. Zabini, B. M. Masini, A. Conti, and L. Hanzo, "Partial equalization for MC–CDMA systems in non-ideally estimated correlated fading," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3818–3830, Oct. 2010.

[12] W. Yang, J. Chen, Z. Tan, and S. Cheng, "Performance analysis of multicarrier DS-CDMA with antenna array in a fading channel," *Wireless Pers. Commun.*, vol. 52, pp. 273–287, Jan. 2010.

[13] S. Affes and P. Mermelstein, "A new receiver structure for asynchronous CDMA: STAR-the spatio-temporal array-receiver," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1411–1422, Oct. 1998.

[14] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[15] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[16] M. Pincus, "A closed form solution of certain programming problems," *Oper. Res.*, vol. 16, no. 3, pp. 690–694, 1962.

[17] S. Kay and S. Saha, "Mean likelihood frequency estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1937–1946, Jul. 2000.

[18] H. Wang and S. Kay, "Maximum likelihood angle-Doppler estimator using importance sampling," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 2, pp. 610–622, Apr. 2010.

[19] G. Wang and H. Chen, "An importance sampling method for TDOA-based source localization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1560–1568, May 2011.

[20] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A non-data-aided maximum likelihood time delay estimator using importance sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4505–4515, Oct. 2011.

[21] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A maximum likelihood time delay estimator in a multipath environment using importance sampling," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 182–193, Jan. 2013.

[22] B. Smida, S. Affes, J. Li, and P. Mermelstein, "A spectrum-efficient multicarrier CDMA array-receiver with diversity-based enhanced time and frequency synchronization," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2315–2327, Jun. 2007.

[23] S. Affes and P. Mermelstein, "A blind coherent spatiotemporal processor of orthogonal Walsh-modulated CDMA signals," *Wireless Commun. Mobile Comput.*, vol. 2, no. 7, pp. 763–784, 2002.

[24] S. Affes and P. Mermelstein, "Adaptive space-time processing for wireless CDMA," in *Adaptive Signal Processing: Applications to Real-World Problems*, J. Benesty and A. H. Huang, Eds. Berlin, Germany: Springer, Jan. 2003.

[25] M. I. Miller and D. R. Fuhrmann, "Maximum-likelihood narrow-band direction finding and the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 9, pp. 1560–1577, Sep. 1990.

[26] S. Kay, *Intuitive Probability and Random Processes Using MATLAB*. New York, NY, USA: Springer, 2005.

[27] M. I. Miller and D. L. Snyder, "The role of likelihood and entropy in incomplete-data problems: Applications to estimating point-process intensities and Toeplitz constrained covariances," *Proc. IEEE*, vol. 75, no. 7, pp. 892–907, Jul. 1987.

[28] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.

[29] D. Shutin and B. H. Fleury, "Sparse variational Bayesian SAGE algorithm with application to the estimation of multipath wireless channels," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3609–3623, Aug. 2011.

[30] J. G. Proakis, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2008.

[31] B. Smida, S. Affes, K. Jamaoui, and P. Mermelstein, "A multicarrier-CDMA space–time receiver with full-interference-suppression capabilities," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 363–379, Jan. 2008.

[32] J.-Y. Lee, R. E. Hudson, and K. Yao, "Acoustic DOA estimation: An approximate maximum likelihood approach," *IEEE Syst. J.*, vol. 8, no. 1, pp. 131–141, Mar. 2014.

Ahmed Masmoudi was born in Ariana, Tunisia. He received the Diplôme d'Ingénieur degree in communications from the Ecole Supérieure des Communications de Tunis-Sup'Com, Tunisia, in 2010, the M.Sc. degree from the Institut National de la Recherche Scientifique in 2012, and the Ph.D. degree from McGill University, Montréal, QC, Canada. He is currently a Post-Doctoral Fellow with McGill University. His research activities include statistical signal processing and wireless communications.

Faouzi Bellili received the B.Eng. degree (Hons.) in signals and systems from the Tunisia Polytechnic School in 2007, and the M.Sc. and Ph.D. degrees (Hons.) from the National Institute of Scientific Research, University of Quebec, Montréal, QC, Canada, in 2009 and 2014, respectively. From 2014 to 2016, he was a Research Associate with INRS-EMT, where he coordinated a major multi-institutional NSERC collaborative research and development project on fifth-generation wireless access virtualization enabling schemes. He is currently a Post-Doctoral Fellow with the University of Toronto, ON, Canada. He has authored/co-authored over 50 peer-reviewed papers in reputable IEEE journals and conferences. His research focuses on statistical and array signal processing for wireless communications and 5G-enabling technologies. He received the very prestigious NSERC PDF Grant for the period 2017–2018. He also received the prestigious PDF Scholarship offered over the same period (but declined) from the Fonds de Recherche du Québec Nature et Technologies, the INRS Innovation Award for the year 2014/2015, the very prestigious Academic Gold Medal of the Governor General of Canada for the year 2009–2010, the Excellence Grant of the Director General of INRS for the year 2009–2010, the Award of the Best M.Sc. Thesis in INRS-EMT for the year 2009–2010, and twice for both the M.Sc. and Ph.D. programs, the National Grant of Excellence from the Tunisian Government. In 2011, he received the Merit Scholarship for Foreign Students from the Ministère de l'Éducation, du Loisir et du Sport of Québec, Canada, for the period 2011–2013. He was also selected by the INRS as its candidate for the 2009–2010 competition of the very prestigious Vanier Canada Graduate Scholarships program and for the 2014–2015 competition of the very prestigious Doctoral Dissertation Award of the Northeastern Association of Graduate schools. In 2015, he was also selected by the University of Waterloo and the University of Toronto as one of their candidates to the very prestigious Banting Postdoctoral Fellowship competition for the year 2015/2016. He serves regularly as a TPC member for major IEEE conferences and acts as a reviewer for many international scientific journals and conferences.

Sofiène Affes (S'94–SM'04) received the Diplôme d'Ingénieur degree in telecommunications and the Ph.D. degree (Hons.) in signal processing from the École Nationale Supérieure des Télécommunications, Paris, France, in 1992 and 1995, respectively. He was a Research Associate with INRS, Montréal, QC, Canada, until 1997, an Assistant Professor until 2000, and an Associate Professor until 2009. From 2003 to 2013, he was a Canada Research Chair in wireless communications. He is currently a Full Professor at INRS and also the Director of PERSWADE, a unique CAD 4M research training program on wireless in Canada involving twenty-seven faculty members from eight universities and ten industrial partners. He was a recipient of the Discovery Accelerator Supplement Award twice from NSERC, from 2008 to 2011 and from 2013 to 2016. He was an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He served as a General Co-Chair of the IEEE VTC'2006-Fall and the IEEE ICUWB 2015, both held in Montréal. For his contributions to the success of both events, he received a Recognition Award from the IEEE Vehicular Technology Society in 2008 and a Certificate of Recognition from the IEEE Microwave Theory and Techniques Society in 2015. He is an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and the *Journal on Wireless Communications and Mobile Computing* (Wiley). He is currently serving as the General Chair of the 28th IEEE PIMRC to be held in Montréal in the fall 2017.

Ali Ghrayeb received the Ph.D. degree in electrical engineering from the University of Arizona, Tucson, AZ, USA, in 2000. He was with Concordia University, Montréal, Canada. He is currently a Professor with the Department of Electrical and Computer Engineering, Texas A&M University at Qatar. He has co-authored the book *Coding for MIMO Communication Systems* (Wiley, 2008). His research interests include wireless and mobile communications, physical layer security, massive MIMO, wireless cooperative networks, and ICT for health applications. He was a co-recipient of the IEEE Globecom 2010 Best Paper Award. He has served as an instructor or co-instructor in technical tutorials at several major IEEE conferences. He served as the Executive Chair of the 2016 IEEE WCNC Conference and the TPC Co-Chair of the Communications Theory Symposium at the 2011 IEEE Globecom. He serves as an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS. He has served on the editorial boards of several IEEE and non-IEEE journals.