

# A NEW IMPORTANCE-SAMPLING ML ESTIMATOR OF TIME DELAYS AND ANGLES OF ARRIVAL IN MULTIPATH ENVIRONMENTS

Faouzi Bellili, Souheib Ben Amor, Sofiène Affes and Abdelaziz Samet

INRS-EMT, 6900-800 de la Gauchetiere Ouest, Montreal (Quebec), H5A 1K6.

{bellili, souheib.ben.amor, affes, samet}@emt.inrs.ca

## ABSTRACT

In this paper, the importance sampling (IS) concept is exploited for the first time in the context of maximum likelihood (ML) estimation of both the time delays and angles of arrival (AoAs) in multipath propagation environments. The global maximum of the compressed likelihood function (CLF) is found empirically with a low computational cost. Simulations suggest that the new IS-based ML-type estimator outperforms, in terms of accuracy, the main state-of-the-art techniques published on the topic. It is also able to reach the Cramér-Rao-lower bound (CRLB) [13] with few received samples.

**Index Terms**— time delays, AoAs, multipath propagation.

## 1. INTRODUCTION

In many fields such as radar, sonar or wireless communications, it is crucial to have accurate estimates for the received signal's time delays and angles of arrival (AoAs). Different methods have been so far proposed on this topic and most of them consider the separate estimation of either the angles or the delays [1–8]. Yet, estimating the AoAs separately from the delays is not optimal when the multipath components overlap. Hence, joint estimation of these parameters can improve the performance significantly. Further, having the information about the angle and the delay of each reflected path can improve the performance of the equalization process [9].

In this context, there have been some works in the open literature that tackle the joint angles and delays estimation (JADE) problem. Roughly speaking, all the existing JADE methods can be categorized into two major categories: subspace-based algorithms and ML-based ones. The most common subspace-based JADE estimators rely on extensions of the well-known Multiple Signal Classification (MUSIC) [10–12] and Estimation of Signal Parameters by Rotational Invariance Techniques (ESPRIT) [13]. Under the ML class, we mention the notable solution of [9] which is iterative in nature and, therefore, its performance is closely tied to the initial guess it requires. Motivated by these facts, we propose in this paper a new ML JADE technique which is based on the IS concept [17], a well-known Monte-Carlo technique that is usually used to solve high-dimensional optimization problems. The main advantage of the new ML technique is that it always returns the global maximum of the CLF and does not require any initial guess.

The rest of this paper is structured as follows. In section 2, we introduce the system model. Section 3 is dedicated to the derivation of the CLF and the new ML solution. In section 4, we present the simulation results before drawing out some concluding remarks in section 5.

## 2. SYSTEM MODEL

Consider a uniform linear array (ULA) consisting of  $P$  antenna elements with half-wavelength spacing immersed in a homogeneous medium in the far field of one source that is transmitting a planar

wave. After traveling through a multi-path environment, the received analog signal on each  $\{p^{th}\}_{p=1}^P$  antenna element is modeled as a complex signal as follows:

$$x_p(t) = \sum_{q=1}^Q \gamma_q a_p(\theta_q) s(t - \tau_q) + w_p(t), \quad (1)$$

where  $Q$  is the number of reflections and  $s(t)$  is a known transmitted signal. The noise components  $w_p(t)$  are modeled by zero-mean complex Gaussian random processes with independent real and imaginary parts, each of variance  $\sigma^2/2$ . The complex channel coefficients  $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_Q]^T$  are assumed to be constant but unknown. Moreover,  $a_p(\theta_q)$  is the amplitude response of the  $p^{th}$  sensor to a wavefront impinging from location  $\theta_q$  with time delay  $\tau_q$ . We also define the parameter vectors  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_Q]^T$  and  $\boldsymbol{\tau} = [\tau_1, \tau_2, \dots, \tau_Q]^T$ . After sampling the received analog signal at time instants  $t_m = mT_s$ , the obtained samples are given by [9]:

$$x_p(t_m) = \sum_{q=1}^Q \gamma_q s(t_m - \tau_q) e^{j\pi(p-1)\cos(\theta_q)} + w_p(t_m), \quad (2)$$

where  $m = 1, 2, \dots, M$  with  $M$  being the total number of received samples that are gathered in a space-time matrix:

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S}(\boldsymbol{\tau}) + \boldsymbol{\Omega}, \quad (3)$$

whose entries are given by  $[\mathbf{X}]_{p,m} = x_p(t_m)$ . In (3),  $\mathbf{S}(\boldsymbol{\tau})$  is a  $(Q \times M)$  matrix whose  $q^{th}$  row is  $\mathbf{s}(\tau_q) = [s(t_1 - \tau_q), s(t_2 - \tau_q), \dots, s(t_M - \tau_q)]$ ,  $\boldsymbol{\Omega}$  is a  $(P \times M)$  noise matrix with entries  $[\boldsymbol{\Omega}]_{p,m} = w_p(t_m)$  and  $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_Q)]$  is a  $(P \times Q)$  matrix containing the  $Q$  steering vectors given by:

$$\mathbf{a}(\theta_q) = \left[ 1, e^{j\pi \cos(\theta_q)}, e^{j2\pi \cos(\theta_q)}, \dots, e^{j(P-1)\pi \cos(\theta_q)} \right]^T.$$

## 3. THE IS-BASED ML-TYPE ESTIMATOR

### 3.1. Derivation of The CLF

We begin by deriving the *actual* log-likelihood function from which the CLF will be found. In fact, from (2) it can be shown that:

$$\begin{aligned} \ln(p[\mathbf{x}(t_1), \dots, \mathbf{x}(t_M)]) &= -\pi M \ln(\sigma^2) \\ &- \frac{1}{\sigma^2} \sum_{m=1}^M \left\| \mathbf{x}(t_m) - \sum_{q=1}^Q \gamma_q \mathbf{a}(\theta_q) s(t_m - \tau_q) \right\|^2, \end{aligned} \quad (4)$$

in which  $\mathbf{x}(t_m) = [x_1(t_m), x_2(t_m), \dots, x_P(t_m)]^T$ . Now, maximizing the right-hand side of (4) with respect to  $\{\theta_q\}_{q=1}^Q$  and  $\{\tau_q\}_{q=1}^Q$  is equivalent to minimizing its second term. Hence, by applying the

discrete Fourier transform (DFT) and using the Parseval's equality, we obtain the following objective function:

$$L(\hat{\mathbf{x}}|\boldsymbol{\tau}, \boldsymbol{\theta}, \gamma) = \frac{1}{\sigma^2} \left\| \hat{\mathbf{x}} - \hat{\mathbf{D}}\boldsymbol{\gamma} \right\|^2, \quad (5)$$

where  $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^T, \hat{\mathbf{x}}_2^T, \dots, \hat{\mathbf{x}}_P^T]$  with  $\{\hat{\mathbf{x}}_p = \text{DFT}(\mathbf{x}_p)\}_{p=1}^P$  and  $\mathbf{x}_p = [x_p(t_1), x_p(t_2), \dots, x_p(t_M)]^T$ . We also introduce  $\hat{s}(\omega)$  as the Fourier transform of  $s(t)$  which is evaluated at the involved DFT frequency points  $\{\omega_m\}_{m=1}^M$  and the  $(PM \times Q)$  matrix  $\hat{\mathbf{D}} = [\hat{\mathbf{D}}_1, \hat{\mathbf{D}}_2, \dots, \hat{\mathbf{D}}_P]^T$  where  $\hat{\mathbf{D}}_p$  is a  $(Q \times M)$  matrix whose  $m^{\text{th}}$  column,  $\hat{\mathbf{D}}_p(m)$ , is given by:

$$\hat{\mathbf{D}}_p(m) = [a_p(\boldsymbol{\theta}_1)\hat{s}(\omega_m)e^{-j\omega_m\tau_1}, a_p(\boldsymbol{\theta}_2)\hat{s}(\omega_m)e^{-j\omega_m\tau_2}, \dots, a_p(\boldsymbol{\theta}_Q)\hat{s}(\omega_m)e^{-j\omega_m\tau_Q}]^T. \quad (6)$$

In order to decrease the number of unknown parameters upon which (5) depends, we first proceed to maximizing it with respect to  $\boldsymbol{\gamma}$ , which yields:

$$\hat{\boldsymbol{\gamma}} = (\hat{\mathbf{D}}^H \hat{\mathbf{D}})^{-1} \hat{\mathbf{D}}^H \hat{\mathbf{x}}. \quad (7)$$

Then, by using  $\hat{\boldsymbol{\gamma}}$  in (5) instead of  $\boldsymbol{\gamma}$ , we obtain the so-called CLF:

$$L_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) = \frac{1}{\sigma^2} \left( \hat{\mathbf{x}}^H \hat{\mathbf{D}} (\hat{\mathbf{D}}^H \hat{\mathbf{D}})^{-1} \hat{\mathbf{D}}^H \hat{\mathbf{x}} \right). \quad (8)$$

### 3.2. Global maximization of the CLF

By a closer look at the CLF derived in (8), it is seen that finding an analytical expression of its global maximum is trivially intractable. Therefore, it becomes clear that a numerical or an empirical approach must be envisaged. In this context, an iterative method was previously developed in [9]. Unfortunately, it requires a good initialization, which is not always available in practice. An alternative way that allows to avoid all the drawbacks of iterative implementations is to make use of the global maximization theorem proposed by Pincus in [14]. It simply states that for a given cost function  $g(\cdot)$  that depends on any unknown parameter vector  $\boldsymbol{\alpha}$ , its global maximum is reached at the vector  $\hat{\boldsymbol{\alpha}}$  whose  $q^{\text{th}}$  entry is given by:

$$\hat{\boldsymbol{\alpha}}_q = \lim_{\rho \rightarrow \infty} \frac{\int \dots \int \boldsymbol{\alpha}_q \exp \{ \rho g(\boldsymbol{\alpha}) \} d\boldsymbol{\alpha}}{\int \dots \int \exp \{ \rho g(\boldsymbol{\alpha}) \} d\boldsymbol{\alpha}}. \quad (9)$$

The limit involved in (9) is approximated, for some sufficiently high value  $\rho_0$  of  $\rho$ , as follows:

$$\hat{\boldsymbol{\alpha}}_q = \frac{\int \dots \int \boldsymbol{\alpha}_q \exp \{ \rho_0 g(\boldsymbol{\alpha}) \} d\boldsymbol{\alpha}}{\int \dots \int \exp \{ \rho_0 g(\boldsymbol{\alpha}) \} d\boldsymbol{\alpha}}. \quad (10)$$

Applying this theorem to our estimation problem yields the following ML estimates:

$$\hat{\tau}_q = \int \dots \int \tau_q \bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\tau}, \boldsymbol{\theta}) d\boldsymbol{\theta} d\boldsymbol{\tau}, \quad q = 1, 2, \dots, Q \quad (11)$$

$$\hat{\boldsymbol{\theta}}_q = \int \dots \int \boldsymbol{\theta}_q \bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\tau}, \boldsymbol{\theta}) d\boldsymbol{\theta} d\boldsymbol{\tau}, \quad q = 1, 2, \dots, Q \quad (12)$$

where  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\tau}, \boldsymbol{\theta})$  is the joint pseudo-probability density function (pseudo-pdf) defined as:

$$\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) = \frac{\exp \{ \rho_0 L_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) \}}{\int \dots \int \exp \{ \rho_0 L_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) \} d\boldsymbol{\theta} d\boldsymbol{\tau}}. \quad (13)$$

Here, we denote from now the  $q^{\text{th}}$  element of any vector  $\mathbf{z}$  as  $z_q$ . Therefore in (11) and (12), we have  $\hat{\tau}_q = \hat{\tau}_q$  and  $\hat{\boldsymbol{\theta}}_q = \hat{\boldsymbol{\theta}}_q$  for  $q = 1, 2, \dots, Q$ . Intuitively, when  $\rho_0$  tends to infinity,  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  in (13) becomes a Dirac-delta function centered at the location of the true maximum value. Thus, the ML estimates are easily obtained by evaluating the integrals in (11) and (12), yet with a careful choice of the design parameter  $\rho_0$ . Fortunately, these integrals can be interpreted as the expected values of multiple realizations which are generated using the pseudo-pdf  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$ , i.e.,  $\hat{\tau}_q = E_{\boldsymbol{\tau}}\{\tau_q\}$  and  $\hat{\boldsymbol{\theta}}_q = E_{\boldsymbol{\theta}}\{\boldsymbol{\theta}_q\}$ . Indeed, if we are able to generate  $R$  realizations of random vectors,  $\{\boldsymbol{\tau}^{(r)}\}_{r=1}^R$  and  $\{\boldsymbol{\theta}^{(r)}\}_{r=1}^R$ , which are distributed according to  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$ , it will be very accurate to approximate the integrals in (11) and (13) as follows:

$$\hat{\tau}_q = \frac{1}{R} \sum_{r=1}^R \tau_q^{(r)} \quad \text{and} \quad \hat{\boldsymbol{\theta}}_q = \frac{1}{R} \sum_{r=1}^R \boldsymbol{\theta}_q^{(r)}. \quad (14)$$

Clearly, as the number of generated values  $R$  increases, the variances of the above arithmetic means become smaller and the estimates approach the global maximum of the CLF.

### 3.3. The IS technique

A remaining practical issue is how to generate the required realizations according to  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$ . Unfortunately, the proposed pseudo-pdf in (13) is extremely non-linear. Our goal is then is to approximate it by a product of multiple two-dimensional pdfs and, therefore, transform the problem of generating two vectors of parameters, simultaneously, into generating multiple *couples* of realizations. This can be accomplished through the IS concept. Applied to our situation, it leads to the following result:

$$\hat{\boldsymbol{\alpha}}_q = \int \dots \int \boldsymbol{\alpha}_q \frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})}{\bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})} \bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) d\boldsymbol{\theta} d\boldsymbol{\tau}. \quad (15)$$

where  $\boldsymbol{\alpha}_q$  stands for  $\tau_q$  or  $\boldsymbol{\theta}_q$  depending on the context and  $\bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  is another pseudo-pdf to be designed as close as possible to  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  while allowing at the same time to separate the multidimensional integrals. After doing so, we can use the following Monte-Carlo approximation to easily evaluate the involved integrals:

$$\int \dots \int \boldsymbol{\alpha}_q \frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})}{\bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})} \bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) d\boldsymbol{\theta} d\boldsymbol{\tau} \approx \frac{1}{R} \sum_{r=1}^R \boldsymbol{\alpha}_q^{(r)} \frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})}{\bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})},$$

where  $\boldsymbol{\alpha}_q^{(r)}$  is the  $r^{\text{th}}$  generated realization according to the simpler pseudo-pdf  $\bar{g}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$ . In fact, by revisiting (8), one can notice that it is the presence of the inverse  $[\hat{\mathbf{D}}^H \hat{\mathbf{D}}]^{-1}$  that actually makes the generation of the required realizations with respect to  $\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  so difficult. Fortunately, we show in the sequel that the matrix  $\hat{\mathbf{D}}^H \hat{\mathbf{D}}$  can be approximated by a diagonal matrix. Indeed, we have:

$$\hat{\mathbf{D}}^H \hat{\mathbf{D}} = \sum_{m=1}^M |\hat{s}(\omega_m)|^2 \boldsymbol{\Phi}^H(\omega_m, \boldsymbol{\tau}) \mathbf{A}^H(\boldsymbol{\theta}) \mathbf{A}(\boldsymbol{\theta}) \boldsymbol{\Phi}(\omega_m, \boldsymbol{\tau}),$$

where:

$$\boldsymbol{\Phi}(\omega_m, \boldsymbol{\tau}) = \text{diag} \left( e^{-j\omega_m\tau_1}, e^{-j\omega_m\tau_2}, \dots, e^{-j\omega_m\tau_Q} \right).$$

The diagonal entries of  $\hat{\mathbf{D}}^H \hat{\mathbf{D}}$  are thus given by:

$$[\hat{\mathbf{D}}^H \hat{\mathbf{D}}]_{ll} = P \sum_{m=1}^M |\hat{s}(\omega_m)|^2, \quad (16)$$

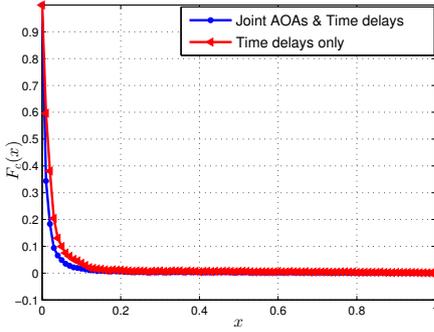
and the off-diagonal ones can be rewritten as follows:

$$[\widehat{\mathbf{D}}^H \widehat{\mathbf{D}}]_{lk} = \left( \sum_{m=1}^M |\hat{s}(\omega_m)|^2 e^{j\omega_m(\tau_l - \tau_k)} \right) \left( \sum_{p=1}^P e^{j\pi(p-1)(\cos\theta_k - \cos\theta_l)} \right).$$

Now, it becomes clear that it is indeed reasonable to neglect the off-diagonal elements in front of the diagonal ones, i.e.,  $[\widehat{\mathbf{D}}^H \widehat{\mathbf{D}}]_{ll} \gg [\widehat{\mathbf{D}}^H \widehat{\mathbf{D}}]_{lk}$ . To validate this assumption, we define  $\tau_{lk} = (\tau_l - \tau_k)$  and the ratio,  $\beta(\tau_{lk}, \theta_l, \theta_k) = [\widehat{\mathbf{D}}^H \widehat{\mathbf{D}}]_{lk} / [\widehat{\mathbf{D}}^H \widehat{\mathbf{D}}]_{ll}$ , which is given by:

$$\beta(\tau_{lk}, \theta_l, \theta_k) = \frac{1}{P \sum_{m=1}^M |\hat{s}(\omega_m)|^2} \times \left( \sum_{m=1}^M |\hat{s}(\omega_m)|^2 e^{j\omega_m(\tau_l - \tau_k)} \right) \times \left( \sum_{p=1}^P e^{j\pi(p-1)(\cos\theta_k - \cos\theta_l)} \right). \quad (17)$$

Then, we generate a very large number of realizations of  $\tau_{lk}$ , and  $(\theta_l, \theta_k)$  uniformly over  $[-10T_s, 10T_s]$  and  $[0, \pi]^2$ , respectively. By using these realizations in (17), we compute the complementary cumulative density function  $F_c(x) = Pr[\beta(\tau_{lk}, \theta_l, \theta_k) > x]$  and plot it in Fig. 1.



**Fig. 1.** Complementary cumulative density function of  $\beta(\tau_{lk}, \theta_l, \theta_k)$ .

This figure, suggests that the off-diagonal elements of  $\widehat{\mathbf{D}}^H \widehat{\mathbf{D}}$  can be neglected with very high probability in front of the diagonal ones. In fact, the ratio  $\beta(\tau_{lk}, \theta_l, \theta_k)$  has an extremely low probability to exceed 0.1. We see also that this approximation is more valid in joint delays and AoAs estimation as compared to the estimation of the delays only, for which the ratio is given by:

$$\beta(\tau_{lk}) = \frac{\sum_{m=1}^M |\hat{s}(\omega_m)|^2 e^{j\omega_m(\tau_l - \tau_k)}}{P \sum_{m=1}^M |\hat{s}(\omega_m)|^2}. \quad (\text{delays only})$$

Thus, we will henceforth use the following approximation:

$$\widehat{\mathbf{D}}^H \widehat{\mathbf{D}} \approx \left( P \sum_{m=1}^M |\hat{s}(\omega_m)|^2 \right) \mathbf{I}_{Q \times Q}, \quad (18)$$

with  $\mathbf{I}_{Q,Q}$  being the  $(Q \times Q)$  identity matrix. By doing so, we obtain:

$$L_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) \approx \frac{\rho_0}{\sigma^2 P \sum_{m=1}^M |\hat{s}(\omega_m)|^2} \hat{\mathbf{x}}^H \widehat{\mathbf{D}} \widehat{\mathbf{D}}^H \hat{\mathbf{x}}, \quad (19)$$

which can be rewritten as follows:

$$L_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) \approx \rho_1 \sum_{q=1}^Q I(\boldsymbol{\theta}_q, \boldsymbol{\tau}_q), \quad (20)$$

where  $\rho_1 = \rho_0 \left( \sigma^2 P \sum_{m=1}^M |\hat{s}(\omega_m)|^2 \right)^{-1}$  and  $I(\boldsymbol{\theta}, \boldsymbol{\tau})$  is the periodogram of the signal:

$$I(\boldsymbol{\theta}, \boldsymbol{\tau}) = \left\| \sum_{p=1}^P e^{j\pi(p-1)\cos(\theta)} \sum_{m=1}^M s(\omega_m) \hat{\mathbf{x}}_p^*(\omega_m) e^{-j2\pi\tau\omega_m} \right\|^2.$$

From (20) it becomes clear that the simpler pseudo-pdf that should be used in (15) is given by:

$$\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) = \frac{\exp\{g_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})\}}{\int \cdots \int \exp\{g_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})\} d\boldsymbol{\theta} d\boldsymbol{\tau}}, \quad (21)$$

where  $g_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  is given by:

$$g_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}) \triangleq \rho_1 \sum_{q=1}^Q I(\boldsymbol{\theta}_q, \boldsymbol{\tau}_q), \quad (22)$$

In the following, we provide the estimation of the time delays only since the very same steps hold for the AoAs as well. By injecting (21) in (15) and using the fact that  $\frac{L_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})}{\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})} \approx 1$ , it can be shown that the multiple integrals in (15) can be split into  $Q$  pairs of double integrals each of which corresponding to one of the couples  $(\boldsymbol{\tau}_q, \boldsymbol{\theta}_q)$ . Hence, we obtain the following  $q^{th}$  estimate:

$$\hat{\tau}_q = \frac{\iint \tau_q \exp\{\rho_1 I(\boldsymbol{\theta}_q, \boldsymbol{\tau}_q)\} d\boldsymbol{\theta}_q d\boldsymbol{\tau}_q}{\iint \exp\{\rho_1 I(\boldsymbol{\theta}_q, \boldsymbol{\tau}_q)\} d\boldsymbol{\theta}_q d\boldsymbol{\tau}_q}. \quad (23)$$

At this point, it becomes obvious that the pdf that can be used to generate the required realizations is only two-dimensional and is given by:

$$\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}_q, \boldsymbol{\tau}_q) = \frac{\exp\{\rho_1 I(\boldsymbol{\theta}_q, \boldsymbol{\tau}_q)\}}{\iint \exp\{\rho_1 I(\boldsymbol{\theta}_q, \boldsymbol{\tau}_q)\} d\boldsymbol{\theta}_q d\boldsymbol{\tau}_q}. \quad (24)$$

Using (24) in (23), it follows that the estimates  $\hat{\tau}_q$  can be rewritten as:

$$\hat{\tau}_q = \int \tau_q \left( \int \bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}_q, \boldsymbol{\tau}_q) d\boldsymbol{\theta}_q \right) d\boldsymbol{\tau}_q. \quad (25)$$

We further notice that the inner integral in (25) is nothing but the marginal pdf of  $\boldsymbol{\tau}_q$  denoted as:

$$\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\tau}_q) = \int \bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}_q, \boldsymbol{\tau}_q) d\boldsymbol{\theta}_q,$$

and, hence, the ML estimates of the time delays are expressed as follows:

$$\hat{\tau}_q = \int \tau_q \bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\tau}_q) d\boldsymbol{\tau}_q, \quad \text{for } q = 1, \dots, Q, \quad (26)$$

which is equivalent to the expected value of  $\boldsymbol{\tau}_q$  that is computed as done in (14) after generating  $R$  realizations  $\{\boldsymbol{\tau}_q^{(r)}\}_{r=1}^R$  according to  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\tau}_q)$ . At low SNR values, an estimation bias can be introduced as explained in [15]. To overcome this shortcoming, one can use the circular mean instead of the linear mean in (14), i.e.:

$$\hat{\tau}_q \approx \frac{1}{2\pi} \angle \sum_{r=1}^R \exp\{j2\pi\boldsymbol{\tau}_q^{(r)}\}, \quad (27)$$

where  $\angle(\cdot)$  returns the argument of any complex number. The estimates for the AoAs can be obtained by applying the same procedure.

Yet, it is seen that we have so far neglected the coefficient  $\frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})}{\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})}$  that appears in (15). This term can be actually used in (27) in order to improve the estimation performance. In this case, the estimates of  $\boldsymbol{\tau}$  and  $\boldsymbol{\theta}$  are simply given by:

$$\hat{\boldsymbol{\tau}} = \frac{1}{2\pi} \angle \sum_{r=1}^R \frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})}{\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})} \exp \{j2\pi\boldsymbol{\tau}^{(r)}\}, \quad (28)$$

$$\hat{\boldsymbol{\theta}} = \frac{1}{2\pi} \angle \sum_{r=1}^R \frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})}{\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})} \exp \{j2\pi\boldsymbol{\theta}^{(r)}\}. \quad (29)$$

The implementation of the IS algorithm can be summarized as follows:

---

#### Algorithm 1 Implementation of IS algorithm

---

Evaluate  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  at many grid points  $(\tau^i, \theta^i)$  and take the  $Q$  largest maxima  $\hat{\boldsymbol{\tau}}^{\max}$  and  $\hat{\boldsymbol{\theta}}^{\max}$  of  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\tau})$  and  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta})$ , respectively.

Reevaluate  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau})$  locally around  $\hat{\boldsymbol{\tau}}^{\max}$  and  $\hat{\boldsymbol{\theta}}^{\max}$ .

**for**  $r \leftarrow 1$  to  $R$  **do**

**for**  $q \leftarrow 1$  to  $Q$  **do**

    generate  $\boldsymbol{\tau}_q^{(r)}$  locally around  $\hat{\boldsymbol{\tau}}_q^{\max}$  using  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\tau})$ .

    generate locally  $\boldsymbol{\theta}_q^{(r)}$  around  $\hat{\boldsymbol{\theta}}_q^{\max}$  using  $\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}, \boldsymbol{\tau}_q^{(r)})$ .

**end for**

  Evaluate  $\frac{\bar{L}_c(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})}{\bar{g}_{\rho_1}(\hat{\mathbf{x}}|\boldsymbol{\theta}^{(r)}, \boldsymbol{\tau}^{(r)})}$

**end for**

Use (28) and (29) to obtain  $\hat{\boldsymbol{\theta}}$  and  $\hat{\boldsymbol{\tau}}$ .

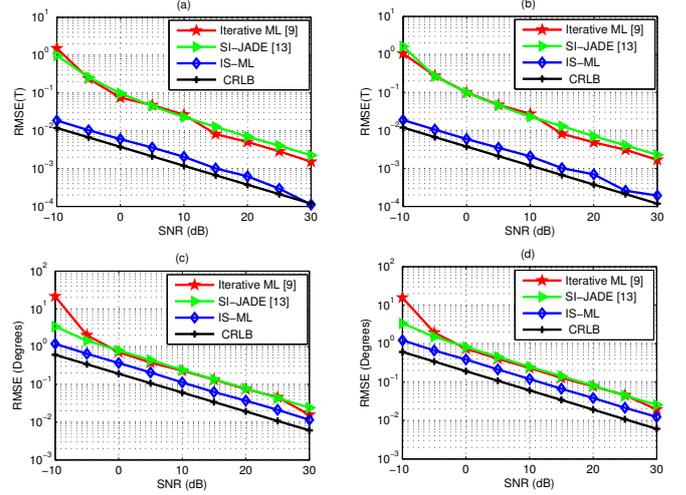
---

For more details about the generation of the required realizations, we refer the reader to [16].

## 4. SIMULATION RESULTS

In this section, we assess the performance of the new IS-based ML estimator using  $R = 2000$  and  $M_c = 1000$  Monte-Carlo runs in all simulations. We also consider the iterative ML estimator of [9], the shift invariance joint angles-of-arrival and time-delays estimation (SI-JADE) technique of [13] and the CRLB [13], as benchmarks. For the sake of brevity and without loss of generality, we consider the case of two equi-powered paths only. We will also consider the case of a five-element ULA, i.e.,  $P = 5$ . The sensors are separated by half the wavelength and a raised-cosine signal  $s(t)$ , with excess bandwidth  $\Delta f = 0.3$ , is considered as a known transmitted waveform. We also sample the received signal at twice the Nyquist rate.

Fig. 2 compares the performance of the proposed method against the benchmarks listed above in terms of RMSE (root mean square error) versus the SNR for both paths. The SNR is the signal-to-noise ratio given by  $\text{SNR} = |\gamma_q|^2 \sum_{m=1}^M |s(t_m)|^2 / (M \sigma^2)$  for each  $q^{\text{th}}$  path. The two paths are located at directions  $90^\circ$  and  $75^\circ$  with delays  $5 T_s$  and  $2.5 T_s$ , respectively. Fig. 2 shows that our IS-based ML-type solution (IS-ML) algorithm exhibits the best performance over the entire SNR range. In fact, the iterative ML estimator [9] relies on another initialization algorithm and the latter is not guaranteed to provide sufficiently accurate initial guesses about the delays and the AoAs. This makes the iterative ML technique converge to a local maximum affecting thereby its overall performance. The new IS-based solution, however, always returns the global maximum of the CLF without any need for an initial guess. The new IS-based algorithm outperforms, as well, SI-JADE which is a subspace-based technique and, therefore, it requires a large number of received samples to reach the CRLB.



**Fig. 2.** RMSE for (a) the first delay, (b) the second delay, (c) the first AoA, and (d) the second AoA with:  $M = 128$  samples,  $\rho_1 = 10$ , and  $\rho_0 = 4000$ .

Furthermore, we compare in Tab. 1 the complexity of the three estimators in terms of computational complexity. The iterative ML of [9] require the maximization of multiple cost functions with respect to  $\boldsymbol{\tau}$  and  $\boldsymbol{\theta}$ , separately. The cost functions are evaluated in  $L$  and  $K$  grid points for  $\boldsymbol{\tau}$  and  $\boldsymbol{\theta}$ , respectively. These two parameters are set to 3000 and 2800 for the iterative solution. Clearly, we use smaller number of grid points with IS ( $L_{\text{IS}} = 1000$  and  $K_{\text{IS}} = 1180$ ) since we collect many realizations locally around rough maxima and then refine the final result through the circular mean without the need for a dense grid discretization. We also set  $N_{\text{iter}} = 5$  as the number of iterations for the iterative ML of [9]. To evaluate the computational complexity of SI-JADE [13], we take  $m_1 = 1$  and  $m_2 = 5$  as stacking parameters. We also fix the number of samples to  $M = 128$  and the number of paths to  $Q = 2$ .

**Table 1.** Complexity assessment

Algorithm	Complexity	ratio
IS-based	$O((P+1)M \ln(M) + L_{\text{IS}} P K_{\text{IS}} + K_{\text{IS}} M^2 + P M L_{\text{IS}}) + R(Q^2 M P + 2(M P Q + Q^2) + Q^3 + 2Q)$	1
Iterative ML	$O((P+1)M \ln M \log(M) P) + ((Q(P+1) + P^2 + 1)M + M P^2 + 2P + P^2)K + ((2Q^2 P + Q^2 + P^2)M + M + 2ML) + (P Q M + Q^2 M Q^2 + P Q^2 + 2P + 1)K) N_{\text{iter}} Q$	9.5119
SI-JADE	$M \log(M) (P+1) + P M + m_1 (M - m_2 + 1)^3 + (m_2 m_1 (M - m_2 + 1))^2 + (m_2 - 1)(M - m_1 + 1)(m_1 * (M - m_2 + 1)^2 + 4 m_1 (M - m_2 + 1)) + m_1 (M - m_2 + 1)(2(m_2 - 1) * (M - m_1 + 1))^2 + 2 Q^2 (m_1 (M - m_2 + 1)(m_1 - 1) - (2/3) Q^3) 4 + 3 Q^3$	4.7597

It turns out that the iterative algorithm is nine times more complex than the new IS-based ML estimator. Accuracy of the former depends on the discretization step and intensive discretization is hence required to achieve higher performance. Clearly, the use of multiple SVD operations and the real processing as described in [13] makes SI-JADE more complex than the IS ML algorithm, as well.

## 5. CONCLUSION

In this paper, we proposed a new non-iterative implementation of the ML estimator that exploits the importance sampling concept for the joint AoAs and time delays ML estimation. One of the main advantages of the new estimator is that it always returns the global maximum of the likelihood function reaching thereby the CRLB. Moreover, it remarkably incurs less computational burden than the most accurate state-of-the-art techniques while achieving remarkably better performance.

## 6. REFERENCES

- [1] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques", *IEEE Trans. on Acous. Speech and Sig. Proc.*, vol. 37, no. 7, pp. 984-995, Aug. 1989.
- [2] C. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay", *IEEE Trans. on Acous. Speech and Sig. Proc.*, vol. 24, no. 4, pp. 320-327, Jul. 1976.
- [3] M. L. McCloud and L. L. Scharf, "A new subspace identification algorithm for high-resolution DOA estimation", *IEEE Trans. Sig. Process.*, vol. 50, no. 10, pp. 1382-1390, Oct. 2002.
- [4] A. Masmoudi, F. Bellili, and S. Affes, and A. Stéphenne, "A maximum likelihood time delay estimator using importance sampling", in *Proc. of IEEE GLOBECOM*, Houston, TX, USA, 2011, pp. 1-6.
- [5] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A maximum likelihood time delay estimator in a multipath environment using importance sampling", *IEEE Trans. Sig. Process.*, vol. 61, no. 1, pp. 182-193, Jan. 2013.
- [6] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A new importance-sampling-based non-data-aided maximum likelihood time delay estimator", in *IEEE Wireless Communications and Networking Conference*, Cancun, Quintana Roo, Mexico, 2011, pp. 1682-1687.
- [7] H. Wang, S. Kay, and S. Saha, "An importance sampling maximum likelihood direction of arrival estimator", *IEEE Trans. Sig. Process.*, vol. 56, no. 10, pp. 5082-5092, Oct. 2008.
- [8] A. Masmoudi, F. Bellili, S. Affes, and A. Stéphenne, "A non-data-aided maximum likelihood time delay estimator using importance sampling", *IEEE Trans. Sig. Process.*, vol. 59, no. 10, pp. 4505 - 4515, Oct. 2011.
- [9] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal", *IEEE Trans. Sig. Process.*, vol. 45, no. 10, pp. 2477-2477, Oct. 1997.
- [10] Y. Wang, J. Chen, and W. Fang, "TST-MUSIC for joint DOA-delay estimation", *IEEE Trans. Sig. Process.*, vol. 49, no. 4, pp. 721-729, Apr. 2001.
- [11] M. Vanderveen, A. Vanderveen, and A. Paulraj, "Estimation of multipath parameters in wireless communications", *IEEE Trans. Sig. Process.*, vol. 46, no. 3, pp. 682-690, Mar. 1998.
- [12] M. C. Vanderveen, C. B. Papadias, and A. Paulraj, "Joint angle and delay estimation (JADE) for multipath signals arriving at an antenna array", *IEEE Sig. Process. Lett.*, vol. 1, no. 1, pp. 12-14, Jan. 1997.
- [13] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance properties", *IEEE Sig. Process. Lett.*, vol. 46, no. 2, pp. 405-418, Feb. 1998.
- [14] M. Pincus, "A closed form solution for certain programming problems", *Oper. Res.*, pp. 690-694, 1962.
- [15] S. Kay, "Comments on frequency estimation by linear prediction", *IEEE Trans. on Acous. Speech and Sig. Process.*, vol. 27, no. 2, pp. 198-199, Apr. 1979.
- [16] S. Saha and S. M. Kay, "Maximum likelihood parameter estimation of superimposed chirps using monte carlo importance sampling", *IEEE Trans. Sig. Process.*, vol. 58, no. 2, pp. 224-230, Feb. 2002.
- [17] L. Stewart, "Bayesian analysis using Monte Carlo integration: A powerful methodology for handling some difficult problems", *Amer. Statist.*, pp. 195200, 1983.