Wideband Array Signal Processing Using Sequential Monte Carlo Methods

William Ng, James P. Reilly*, and Thia Kirubarajan
Department of Electrical and Computer Engineering,
McMaster University, 1280 Main St. W.,
Hamilton, Ontario,
Canada L8S 4K1

Abstract

This paper proposes a novel online tracking, waveform recovery, and detection algorithm for wideband array signal processing. The proposed method extends the wideband data model developed in [1] to the online case through use of particle filters. This proposed method finds application in various areas of telecommunications, including radar, sonar, and other wireless communication problems, where locating moving targets and extracting the waveforms for classification are needed.

The proposed method first detects the number of unknown wideband sources using a hypothesis testing procedure, followed by the sequential estimation of the DOAs of these identified sources, and the extraction of the waveform of these sources, using a particle filter. Computer simulations demonstrate the ability of the proposed method to detect the time-varying number of sources in different scenarios as time progresses. The results also show that the DOA trajectories of the identified sources are well tracked and their signal waveforms are also well restored. Finally, the posterior Cramér-Rao bound (PCRB) [2] is also presented to demonstrate that the estimates by the proposed method lie consistently within the bound.

I. Introduction

Array signal processing, a mature and specialized branch of signal processing, has found use in radar, sonar, communications, geophysical exploration, astrophysical exploration, biomedical signal processing, and acoustics [3], [4], [5]. It essentially deals with the processing of signals carried by propagating wave phenomena. The received signal is obtained by means of an array of sensors located at different points in space in a field of interest.

Array signal processing has to do with 1) detecting the number of incident sources (model order), 2) estimating parameters, like direction-of-arrival (DOA) or *inter-sensor delays* (ISD) of the sources impinging onto the array, and 3) recovering the incident source waveforms. For wideband scenarios, there exist many algorithms to solve these problems partially or completely. Common approaches like [6], [7], [8], [9], [10], sample the spectrum of the incoming signals at

Permission to publish abstract separately is granted.
J. Reilly, corresponding author: ph: 905 525 9140 x22895, fax: 905 521 2922, email: reillyj@mcmaster.ca

each sensor to form an array of narrowband signals. Then traditional signal processing algorithms developed for the narrowband case can be applied to the transformed signals to solve the problems. An alternative approach is proposed in [1] that adopts a novel data model in the time-domain, and incorporates the Markov chain Monte Carlo methods [11], [12], [13], [14] to solve the problems. However, all these approaches assume that the parameters of interest are stationary within an observation window such that batch or *offline* processing must be used. Unfortunately, this stationarity assumption is easily violated in practice, leading to suboptimal results. Since the problem of moving source localization is critical to several important applications in array signal processing, *online* approaches that can recursively estimate the parameters of interest are needed.

In this paper, we propose a new integrated procedure for joint recursive *online* detection of model order, estimation and tracking of DOAs, and recovery of the source waveforms for the *wideband* scenario. Previous approaches for array signal processing in the wideband case have addressed only one, or possibly only a few of these problems. Specifically, there exist many previous algorithms for DOA tracking, e.g., [15], [16], [17], [18], but these only apply to the narrowband scenario. The wideband tracking problem using arrays of sensors has not received much attention to date.

The array processing problem for the wideband scenario is of considerable interest in many communication systems, like radar, sonar, and the wireless 911-problem, etc. In radar applications [19], the instantaneous detection and location of targets of interest can be used in improving the accuracy of navigation aids and/or surveillance systems. In addition, the waveforms extracted from the multiple targets can be used to aid in target classification. Further, wideband target tracking followed by waveform recovery of sources is an essential component of wideband communications with highly maneuvering platforms in hostile mobile environments. In sonar applications, the important goals are the detection and tracking of submarines [5]. For applications in the wireless 911-problem, the need to locate wireless callers has recently gained attention. As a matter of fact, it is now a requirement for Commercial Mobile Radio Service providers in the US to provide ubiquitous location coverage for all wireless 911 callers within 125 meters *rms* by October 2001 [20], [21]. In each of these applications for the wideband scenario, the DOAs of the target sources must first be estimated and then the angle-bearing information be tracked to determine the positions and other parameters like velocities and accelerations of the targets [22], [23], given the knowledge of the platform. However, most of the existing methods treat the DOA estimation and the DOA tracking processes independently, rather than jointly, as is proposed in this paper.

The proposed method uses the sequential Monte Carlo (SMC) methods in conjunction with the Markov Chain Monte Carlo (MCMC) methods [14], [24], that have emerged as useful methods in the signal processing arena, aided by the persistent increase in available computer power. They are Bayesian methods based on the idea of numerically sampling posterior distributions of

interest that are difficult or impossible to handle analytically. Using the histogram so obtained from these samples, statistical inferences on parameters of interest can be readily made.

Simulation results show that the proposed particle filtering approach can track the DOAs of the moving sources and recover the source waveforms. In this paper, we also present posterior Cramér-Rao bound (PCRB) [2] to demonstrate that the estimates made by the proposed method are consistently close to the bound.

The outline is as follows. We start with the development of a novel data model for wideband signals in Section II, followed by the development of the posterior distribution of interest in Section III. In Section IV, a second-order Bayesian importance sampling function is presented, which is the core of the development of the recursive update equation for DOA tracking. In Section V, a rational statistical method for model order detection is given, followed by computer simulations in Section VI. Discussion is given in Section VII, and finally conclusions are drawn in Section VIII.

Notation: Bold upper case symbols denote matrices, bold lower case symbols denote vectors. The superscript ^{T} denotes the transpose operation, and the symbol "$\sim$" means "distributed as." The quantity $p(\cdot)$ denotes a prior probability distribution and $\pi(\cdot)$ denotes a posterior distribution.

II. Data Model

In this section, based on [1], we present a novel data model. Consider $k(t)$ wideband signals, $s_k(t)$, for $k = 0, ..., k(t) - 1$, where $k(t)$ is assumed known, impinging at distinct angles $\phi_k(t)$ onto a uniform linear array of M identical sensors. The frequency response of each wideband signal is band limited to

$$|f| \in \left[f_k^l, f_k^u\right], \quad f_k^u = f_k^l + \Delta f_k, \tag{1}$$

where f_k^l and f_k^u are the lower and upper frequencies, and Δf_k is the bandwidth of the kth signal that is received by the linear array. Denote the inter-sensor delay (ISD) of the kth source by $\tau_k(t)$ at time t as follows [3]

$$\tau_k(t) \triangleq \frac{\Delta}{C} \sin \phi_k(t), \tag{2}$$

where Δ is the interspacing of the sensors and C is the speed of propagation. Let T_{max} be the maximum time delay experienced by an incident source wave as it transverses between two adjacent sensors (see Fig. 1). It can be shown that [1] T_{max} can be expressed as

$$T_{max} = \min_{k=0,...,k(t)-1} \left\{ \frac{1}{2f_k^u} \right\}, \tag{3}$$

such that

$$|\tau_k(t)| \leq T_{max},\ k = 0, ..., k(t) - 1. \tag{4}$$

As shown Fig. 1, the signal undergoes a delay as it progresses along the array. In contrast to the narrowband case, the delayed signal $s_k(t-\tau_k(t))$, due to the kth incident source, cannot be determined by a corresponding phase shift. Moreover, since the time delay is seldom an integer multiple of the sampling frequency, interpolation is required to estimate the time-shifted version $s_k(t-\tau_k(t))$ from samples of the received signal. We now present a novel data model that uses signal interpolation to estimate the time-shifted versions of the wideband signals as they transverse a linear array.

Denote a snapshot vector $\boldsymbol{y}(t) \in \mathcal{R}^{M\times 1}$ at $t = nT_s$, where T_s is the sampling period. For notational convenience, we choose $T_s = 1$ from this point onwards. This snapshot vector represents the data received by a linear array of sensors whose mth element can be expressed as follows

$$y_m(n) = \sum_{k=0}^{k(n)-1} s_k(n - m\tau_k(n)) + \sigma_w w_m(n),\ m = 0, ..., M-1, \tag{5}$$

where $w_m(n)$ is the mth element of $\boldsymbol{w}(n) \in \mathcal{R}^{M\times 1}$, which is an *iid* Gaussian variable defined as,

$$\boldsymbol{w}(n) \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{I}_M\right), \tag{6}$$

and σ_w^2 is the observation noise variance.

Two situations, as shown in Figs. 1a and 1b where the incident angle $\phi_k(n)$ onto the array is respectively less than and greater than 90 degrees with respect to axis of the array, are considered. Note that the delay between the mth and $(m+1)$th sensor in the case of Fig. 1a is $\tau_k(n)$ given by (2); for the case in Fig. 1b, the corresponding delay is $-\tau_k(n)$.

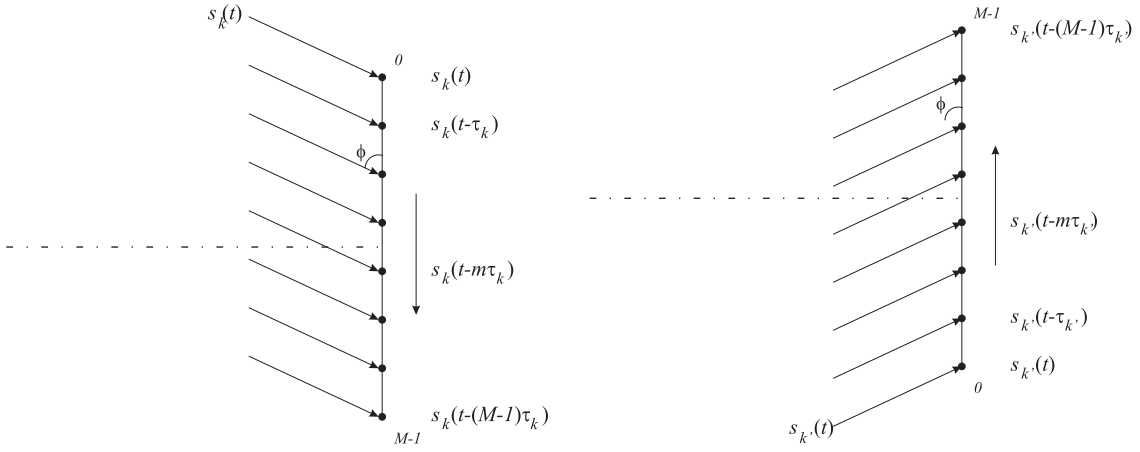

Fig. 1. Illustration of two different situations in a linear array when: (a) $\phi_k(t)_{|t=nT_s} \leq \pi/2$ and (b) $\phi_{k'}(t)_{|t=nT_s} > \pi/2$.

A. $\phi_k(n) \leq \pi/2$

Denote a signal vector, $\boldsymbol{s}_k(n) \in \mathcal{R}^{(2L-1)\times 1}$, for the kth source as follows

$$\boldsymbol{s}_k(n) = [s_k(n+L-1), s_k(n+L-2), ..., s_k(n), s_k(n-1), ..., s_k(n-L+1)]^T, \tag{7}$$

where $2L-1$ is the number of taps used in the interpolation. The vector $\boldsymbol{s}_k(n)$ is essentially a $(2L-1)$-dimensional discrete sample vector of the kth source. When $\phi_k(n) \leq \pi/2$, as shown in Fig. 1(a), the signal $s_k(t)$ first arrives the 0th sensor and then transverses *downward* along the array. Also, the corresponding delay becomes $\tau_k(t) \geq 0$, according to (2). Let $h_l(m\tau_k(t))$ be an interpolation function for the kth source at the mth sensor, and $\boldsymbol{h}(m\tau_k(t)) \in \mathcal{R}^{(2L-1)\times 1}$ be a truncated interpolating sequence of length $2L-1$, defined as

$$\boldsymbol{h}(m\tau_k(t)) = [h_{-L+1}(m\tau_k(t)), ..., h_0(m\tau_k(t)), h_1(m\tau_k(t)), ..., h_{L-1}(m\tau_k(t))]^T, \tag{8}$$

such that we can approximate the snapshot $s_k(n - m\tau_k(n))$ as follows

$$s_k(n - m\tau_k(n)) \approx \boldsymbol{h}^T(m\tau_k(n))\boldsymbol{s}_k(n), \tag{9}$$

$$= \sum_{l=-L+1}^{L-1} h_l(m\tau_k(n))s_k(n-l). \tag{10}$$

An example of a suitable interpolation sequence is given by

$$h_l(\tau_k(n)) = \psi_l(\tau_k(n)) \cdot w_l(\tau_k(n)), l = -L+1, ..., L-1, \tag{11}$$

where $\psi_l(\tau_k(n))$ is a $\text{sinc}(\cdot)$ function and $w_l(\tau_k(n))$ is a suitable window function. The interpolation process of (10) is non-causal. This implies that the processing must undergo a delay of $L-1$ samples for a causal implementation. In some cases, this delay is undesirable. We can eliminate this delay at the expense of some additional error in the interpolation proces, by replacing (10) with

$$s_k(n - m\tau_k(n)) \approx \begin{cases} \sum_{l=0}^{L-1} h_l(n\tau_k(n))s_k(n-l), & \tau \geq 0 \\ \sum_{l=-L+1}^{0} h_l(n\tau_k(n))s_k(n+l), & \tau < 0 \end{cases}. \tag{12}$$

The additional error introduced by the one-side interpolation process in (12) in most cases is small. This follows because the interpolation weights $h_l(\cdot)$ excluded from (12) are outside the main lobe of the interpolation function and are therefore small, especially when τ is small or close to T_s. For ease of notation, we replace the approximation in (12) with an equality in the sequel.

In [25], a comparison is performed for an 10-th order AR process, interpolated by (10) and (12), for different values of τ. The mean-squared errors between the two approaches are listed in Table 2.1.

From Table 2.1, it is seen that the errors are indeed small except possibly for the case $\tau = 0.5T_s$. Thus, the one-sided interpolation process offered by (12) is preferred in this paper, due to the

τ/T_s	MSE
0	5.6×10^{-8}
0.2	6.1×10^{-7}
0.5	3.1×10^{-4}
0.8	1.9×10^{-6}
1	1.6×10^{-9}

TABLE 2.1

MEAN-SQUARED ERROR BETWEEN THE HALF-SINC AND THE FULL-SINC APPROACHES.

simplicity and advantage of zero-delay processing. In some exceptional circumstances, especially when τ approaches $0.5T_s$, it may be necessary to delay the interpolation function several samples, and thus introduce processing delay, in order to maintain a sufficient level of accuracy.

On the basis of (12), we construct an interpolation matrix $\boldsymbol{H}(\tau_k(n)) \in \mathcal{R}^{M\times L}$ for the kth source as follows

$$\boldsymbol{H}(\tau_k(n)) = \begin{bmatrix} \boldsymbol{h}^T(0) \\ \boldsymbol{h}^T(\tau_k(n)) \\ \vdots \\ \boldsymbol{h}^T((M-1)\tau_k(n)) \end{bmatrix}. \tag{13}$$

We then have the following expression to approximate the snapshot vector of the kth source for all M sensors at time $t = n$

$$\begin{bmatrix} s_k(n) \\ s_k(n-\tau_k(n)) \\ \vdots \\ s_k(n-(M-1)\tau_k(n)) \end{bmatrix} \approx \boldsymbol{H}(\tau_k(n))\boldsymbol{s}_k(n). \tag{14}$$

B. $\phi_{k'}(n) > \pi/2$

When $\phi_{k'}(t) < \pi/2$, as shown in Fig. 1(b), the signal $s_{k'}(t)$ first arrives at the $M-1$th sensor and then transverses *upward* along the array. In this case, the corresponding delay becomes $\tau_{k'}(t) < 0$, according to (2). An interpolation sequence for the k'th source at the pth sensor where $p = M-1-m$ is formed as

$$\boldsymbol{h}^*(p\tau_{k'}(t)) = [h_0(p\tau_{k'}(t)), h_1(p\tau_{k'}(t)), ..., h_{L-1}(p\tau_{k'}(t))]^T. \tag{15}$$

The corresponding interpolation matrix $\boldsymbol{H}^*(\tau_{k'}(t)) \in \mathcal{R}^{M\times L}$ is then constructed as follows

$$\boldsymbol{H}^*(\tau_{k'}(t)) = \begin{bmatrix} \boldsymbol{h}^{*T}((M-1)\tau_{k'}(t)) \\ \boldsymbol{h}^{*T}((M-2)\tau_{k'}(t)) \\ \vdots \\ \boldsymbol{h}^{*T}(0) \end{bmatrix}, \tag{16}$$

$$= \boldsymbol{E}_M \boldsymbol{H}(\tau_{k'}(t)), \tag{17}$$

where $\boldsymbol{E}_M \in \mathcal{R}^{M\times M}$ is a *exchange matrix* defined as follows [26]

$$\boldsymbol{E}_M = \begin{bmatrix} 0 & 0 & \dots & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{18}$$

Note that the row vectors in $\boldsymbol{H}^*(\tau_{k'}(t))$ are in the *reverse order* of those defined in (13). We then have the following expression to approximate the snapshot vector of the k'th source all M sensors at time $t = n$

$$\begin{bmatrix} s_k(n-(M-1)\tau'_k(n)) \\ s_k(n-(M-2)\tau'_k(n)) \\ \vdots \\ s'_k(n) \end{bmatrix} \approx \boldsymbol{H}^*(\tau_{k'}(n))\boldsymbol{s}_{k'}(n). \tag{19}$$

As a result, the snapshot vector at the sensor output $\boldsymbol{y}(n)$ whose mth element is described in (5) can be expressed as follows

$$\boldsymbol{y}(n) \approx \sum_{k=0}^{k(n)-1} \tilde{\boldsymbol{H}}(\tau_k(n))\boldsymbol{s}_k(n) + \sigma_w \boldsymbol{w}(n), \tag{20}$$

where

$$\tilde{\boldsymbol{H}}(\tau_k(n)) = \begin{cases} \boldsymbol{H}(\tau_k(n)), & \text{if } \tau_k(n) \geq 0 \\ \boldsymbol{E}_M \boldsymbol{H}(\tau_k(n)), & \text{if } \tau_k(n) < 0 \end{cases}. \tag{21}$$

For notational convenience, from this point onwards we replace the approximation in (20) with an equality. Next we develop the desired data model by rearranging the columns in the interpolation matrices in (20).

In this vein, we define a delay vector $\boldsymbol{\tau}(n) \in \mathcal{R}^{k(n)\times 1}$ and a signal vector $\boldsymbol{a}(n) \in \mathcal{R}^{k(n)\times 1}$ at time n, respectively, as

$$\boldsymbol{\tau}(n) \triangleq [\tau_0(n), \tau_1(n), \dots, \tau_{k(n)-1}(n)]^T, \tag{22}$$

$$\boldsymbol{a}(n) \triangleq [s_0(n), s_1(n), \dots, s_{k(n)-1}(n)]^T. \tag{23}$$

We now define a new set of interpolation matrices $\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}(n)) \in \mathcal{R}^{M\times k(n)}$, for $l = 0, ..., L-1$, which is a collection of all the lth columns of the individual interpolation matrices of $\tilde{\boldsymbol{H}}(\tau_k(n))$ from (20) for $k = 0, ..., k(n)-1$, i.e.,

$$\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}(n)) = \left[\tilde{\mathbf{h}}_l(\tau_0(n)), \tilde{\mathbf{h}}_l(\tau_1(n)), ..., \tilde{\mathbf{h}}_l(\tau_{k(n)-1}(n))\right], \tag{24}$$

where $\tilde{\mathbf{h}}_l(\tau_k(n)) \in \mathcal{R}^{M\times 1}$ is the lth column of $\tilde{\boldsymbol{H}}(\tau_k(n))$. As a result, we can rearrange (20) using (23) and (24) as follows

$$\boldsymbol{y}(n) = \sum_{k=0}^{k(n)-1} \tilde{\boldsymbol{H}}(\tau_k(n))\boldsymbol{s}_k(n) + \sigma_w \boldsymbol{w}(n), \tag{25}$$

$$= \sum_{k=0}^{k(n)-1} \left\{\sum_{l=0}^{L-1} \tilde{\mathbf{h}}_l(\tau_k(n))\boldsymbol{s}_k(n-l)\right\} + \sigma_w \boldsymbol{w}(n), \tag{26}$$

$$= \sum_{l=0}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}(n))\boldsymbol{a}(n-l) + \sigma_w \boldsymbol{w}(n). \tag{27}$$

Accordingly, we define a vector $\boldsymbol{z}(n)$ as follows

$$\boldsymbol{z}(n) \triangleq \boldsymbol{y}(n) - \sum_{l=1}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}(n))\boldsymbol{a}(n-l), \tag{28}$$

which can be interpreted as the error between the observation $\boldsymbol{y}(n)$ and its approximation, based on the past, estimated values of the signals from $\boldsymbol{a}(n-1)$ to $\boldsymbol{a}(n-L+1)$ and the associated columns in the interpolation matrix $\boldsymbol{H}_l(\boldsymbol{\tau}(n))$ for $l = 1, ..., L-1$. Accordingly, we may rewrite (27) as follows

$$\boldsymbol{z}(n) = \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}(n))\boldsymbol{a}(n) + \sigma_w \boldsymbol{w}(n), \tag{29}$$

which represents the desired form of the model.

There are several interesting features regarding the model in (29). It is similar to the familiar narrowband model [3], except that the data is modified and the steering matrix takes the form of an interpolation matrix. Further, the same model in (29) can accomodate either narrowband or wideband sources, without change of structure or parameters. Also, all quantities in (29), including the data, are pure real, unlike previous models [6], [7], [8], [9] which require complex quantities. This latter point leads to significant savings in hardware, since quadrature mixing to IF frequencies is no longer required, and computational requirements are reduced.

III. The State-Space model

We assume the states $[\boldsymbol{\tau}(n), \boldsymbol{a}(n)]$ evolve according to

$$\boldsymbol{\tau}(n) = \boldsymbol{\tau}(n-1) + \sigma_v \boldsymbol{v}(n), \tag{30}$$

$$\boldsymbol{a}(n) \sim \mathcal{N}\left(\mathbf{0}, \sigma_a^2 \boldsymbol{I}_{k(n)}\right), \tag{31}$$

and the transformed observation model $\boldsymbol{z}(n)$ given in (29), is reproduced here as follows

$$\boldsymbol{z}(n) = \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}(n))\boldsymbol{a}(n) + \sigma_w \boldsymbol{w}(n). \tag{32}$$

The noise vector $\boldsymbol{v}(n) \in \mathcal{R}^{k(n)}$ is an *iid* Gaussian variable, defined as

$$\boldsymbol{v}(n) \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{I}_{k(n)}\right), \tag{33}$$

where σ_v^2 is the state noise variance. Note that we also assume that the ISD vector $\boldsymbol{\tau}(n)$ does not change significantly. The dimension $k(n)$ of the model is described by the following stochastic relationship at time n,

$$k(n) = k(n-1) + \epsilon_k(n), \tag{34}$$

where $\epsilon_k(n) \in [-1, 0, 1]$ are discrete *iid* random variables such that

$$\begin{aligned}
P(\epsilon_k(n) = -1) &= \begin{cases} 0, & k = 0 \\ h/2, & \text{otherwise,} \end{cases} \\
P(\epsilon_k(n) = 0) &= \begin{cases} 1 - h/2, & k = 0, k_{max} \\ 1 - h, & \text{otherwise,} \end{cases} \\
P(\epsilon_k(n) = 1) &= \begin{cases} 0, & k = k_{max} \\ h/2, & \text{otherwise,} \end{cases}
\end{aligned} \tag{35}$$

where h is the probability of a change in model order when $k \neq 0, k_{max}$, and k_{max} is the maximum allowable number of sources. In (35), it is tacitly assumed that the model order changes by no more than one in each sample period.

In the proposed system of equations, the noise variances σ_v^2 and σ_w^2 are assumed unknown and constant over time. The unknown vectors of amplitudes $\boldsymbol{a}(n)$ are assumed *iid* between snapshots.

We define the vector of all parameters $\boldsymbol{\theta}$ describing the received signal model as

$$\boldsymbol{\theta}_{1:n} \triangleq \left(\{\boldsymbol{\tau}\}_{1:n}, \{\boldsymbol{a}\}_{1:n}, \sigma_v^2, \sigma_w^2\right), \tag{36}$$

where the notation $(.)_{1:n}$ indicates all the elements from time 1 to time n.

Hence, the joint posterior distribution of all the parameters is $\pi\left(\boldsymbol{\theta}_{1:n}\right) \triangleq p\left(\boldsymbol{\theta}_{1:n} | \boldsymbol{z}_{1:n}\right)$, which can then be expanded using appropriately selected prior distributions of the parameters, according to Bayes' theorem as

$$\begin{aligned}
\pi\left(\boldsymbol{\theta}_{1:n}\right) \propto & p\left(\boldsymbol{z}_{1:n} | \boldsymbol{\tau}_{1:n}, \boldsymbol{a}_{1:n}, k_{1:n}, \sigma_v^2, \sigma_w^2\right) p\left(\boldsymbol{\tau}_{1:n} | k_{1:n}, \sigma_v^2\right) \times \\
& p\left(\boldsymbol{a}_{1:n} | \boldsymbol{\tau}_{1:n}, k_{1:n}, \sigma_w^2\right) p(k_{1:n}) p(\sigma_w^2) p(\sigma_v^2),
\end{aligned} \tag{37}$$

where $p(z_{1:n}|\cdot)$ is the likelihood term, and the remaining distributions constitute the joint prior distribution for the parameters $\boldsymbol{\theta}$. The individual terms in (37) are given as

$$p\left(\boldsymbol{z}_{1:n} | \boldsymbol{\tau}_{1:n}, \boldsymbol{a}_{1:n}, k_{1:n}, \sigma_w^2\right) = \prod_{j=1}^{n} \mathcal{N}\left(\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_j)\boldsymbol{a}_j, \sigma_w^2 \boldsymbol{I}_M\right), \tag{38}$$

$$p\left(\boldsymbol{\tau}_{1:n}|k_{1:n},\sigma_v^2\right)=\prod_{j=1}^{n}\mathcal{N}\left(\boldsymbol{\tau}_{j-1},\sigma_v^2\boldsymbol{I}_{k_j}\right), \tag{39}$$

$$p\left(\boldsymbol{a}_{1:n}|\boldsymbol{\tau}_{1:n},k_{1:n},\sigma_w^2\right)=\prod_{j=1}^{n}\mathcal{N}\left(\mathbf{0},\delta^2\sigma_w^2\left(\tilde{\boldsymbol{H}}_0^T\left(\boldsymbol{\tau}_j\right)\tilde{\boldsymbol{H}}_0\left(\boldsymbol{\tau}_j\right)\right)^{-1}\right), \tag{40}$$

$$p\left(k_{1:n}\right)=\prod_{j=1}^{n}p\left(k_j|k_{j-1}\right)=\prod_{j=1}^{n}\epsilon_k(j). \tag{41}$$

The distribution of (38) follows from (29) and the Normality assumption on the noise $\boldsymbol{w}(n)$. The prior distribution of (39) follows from (30) and the Normality assumption on $\boldsymbol{v}(n)$.

The prior distribution defined in (40) is the maximum entropy prior with the parameter δ^2 set to an estimate of the SNR. The prior distribution on the variances σ_v^2 and σ_w^2 are both assumed to follow the inverse Gamma distribution, defined as follows [27]

$$p\left(\sigma_v^2\right)\sim\mathcal{IG}\left(\frac{\nu_0}{2},\frac{\gamma_0}{2}\right), \tag{42}$$

$$p\left(\sigma_w^2\right)\sim\mathcal{IG}\left(\frac{\nu_1}{2},\frac{\gamma_1}{2}\right). \tag{43}$$

The inverse Gamma distribution is noninformative for $\nu,\gamma=0$.

The parameters of interest are primarily the ISDs $\boldsymbol{\tau}_{1:n}$ and the model order $k_{1:n}$. The amplitudes $\boldsymbol{a}_{1:n}$ along with the noise variances σ_v^2 and σ_w^2 can be estimated separately (as discussed later), and may be considered nuisance parameters.

Using the *iid* Normal distribution of the noise variables, and the model structure given by (29) to (31), and (38) to (43), the posterior distribution $\pi\left(\boldsymbol{\theta}_{1:n}\right)$ in (37) can be written as

$$\begin{aligned}\pi(\boldsymbol{\theta}_{1:n})\propto&\prod_{j=1}^{n}\frac{1}{(2\pi\sigma_w^2)^{M/2}}\exp\left\{-\frac{1}{2\sigma_w^2}\left(\boldsymbol{z}_j-\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_j)\boldsymbol{a}_j\right)^T\left(\boldsymbol{z}_j-\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_j)\boldsymbol{a}_j\right)\right\}\times\\&\prod_{j=1}^{n}\frac{1}{(2\pi\sigma_v^2)^{k_j/2}}\exp\left\{-\frac{1}{2\sigma_v^2}\left(\boldsymbol{\tau}_j-\boldsymbol{\tau}_{j-1}\right)^T\left(\boldsymbol{\tau}_j-\boldsymbol{\tau}_{j-1}\right)\right\}\times\\&\prod_{j=1}^{n}\frac{|\tilde{\boldsymbol{H}}_0^T\left(\boldsymbol{\tau}_j\right)\tilde{\boldsymbol{H}}_0\left(\boldsymbol{\tau}_j\right)|^{1/2}}{(2\pi\delta^2\sigma_w^2)^{k_j/2}}\exp\left\{-\frac{1}{2\delta^2\sigma_w^2}\boldsymbol{a}_j^T\tilde{\boldsymbol{H}}_0^T\left(\boldsymbol{\tau}_j\right)\tilde{\boldsymbol{H}}_0\left(\boldsymbol{\tau}_j\right)\boldsymbol{a}_j\right\}\times\\&\prod_{j=1}^{n}p\left(k_j|k_{j-1}\right)\times\left(\sigma_v^2\right)^{\frac{-\nu_0}{2}-1}\exp\left\{\frac{-\gamma_0}{2\sigma_v^2}\right\}\times\left(\sigma_w^2\right)^{\frac{-\nu_1}{2}-1}\exp\left\{\frac{-\gamma_1}{2\sigma_w^2}\right\}.\end{aligned} \tag{44}$$

Similar to the procedures in [11], the terms relating to the amplitudes can be collected together

to give the following expression

$$\begin{aligned}\pi(\boldsymbol{\theta}_{1:n}) \propto & \prod_{j=1}^{n} \frac{1}{(2\pi\sigma_w^2)^{M/2}} \exp\left\{\frac{-1}{2\sigma_w^2}\boldsymbol{z}_j^T \boldsymbol{P}_0^{\perp}(\boldsymbol{\tau_j})\boldsymbol{z}_j\right\} \times \\ & \prod_{j=1}^{n} \frac{1}{(2\pi\sigma_v^2)^{k_j/2}} \exp\left\{\frac{-1}{2\sigma_v^2}(\boldsymbol{\tau}_j - \boldsymbol{\tau}_{j-1})^T(\boldsymbol{\tau}_j - \boldsymbol{\tau}_{j-1})\right\} \times \\ & \prod_{j=1}^{n} \frac{|\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\tau}_j)|^{1/2}}{(2\pi\delta^2\sigma_w^2)^{k_j/2}} \exp\left\{\frac{-1}{2\delta^2\sigma_w^2}(\boldsymbol{a}_j - \boldsymbol{m}_j)^T \boldsymbol{\Sigma}_0(\boldsymbol{\tau}_j)(\boldsymbol{a}_j - \boldsymbol{m}_j)\right\} \times \\ & \prod_{j=1}^{n}(1+\delta^{-2})^{-k_j/2} \times \prod_{j=1}^{n} p(k_j|k_{j-1}) \times (\sigma_v^2)^{\frac{-\nu_0}{2}-1} \exp\left\{\frac{-\gamma_0}{2\sigma_v^2}\right\} \times (\sigma_w^2)^{\frac{-\nu_1}{2}-1} \exp\left\{\frac{-\gamma_1}{2\sigma_w^2}\right\},\end{aligned} \tag{45}$$

where

$$\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\tau}_j) = \left(1+\delta^{-2}\right)\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau}_j)\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_j), \tag{46}$$

$$\begin{aligned}\boldsymbol{m}_j &= \boldsymbol{\Sigma}_0(\boldsymbol{\tau}_j)\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau}_j)\boldsymbol{z}_j, \\ &= \boldsymbol{\Sigma}_0(\boldsymbol{\tau}_j)\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau}_j)\left(\boldsymbol{y}(j) - \sum_{l=1}^{L-1}\tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}_j)\boldsymbol{a}(t-l)\right),\end{aligned} \tag{47}$$

and

$$\boldsymbol{P}_0^{\perp}(\boldsymbol{\tau}_j) = \boldsymbol{I} - \frac{\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_j)\left(\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau}_j)\tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_j)\right)^{-1}\tilde{\boldsymbol{H}}_0^T(\boldsymbol{\tau}_j)}{(1+\delta^{-2})}. \tag{48}$$

From (45) and (47), a *maximum a posteriori* estimate of the amplitudes, knowing the other parameters is readily available as

$$\hat{\boldsymbol{a}}_{MAP}(n) = \boldsymbol{m}_n. \tag{49}$$

Thus the amplitude parameters need not be included in the particle filter. Instead, they can be estimated at each iteration, after sampling the other parameters.

Recognizing that the third line of (45) is a Normal distribution in $\boldsymbol{a}_t$, with mean $\boldsymbol{m}_t$ and covariance matrix $\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\tau}_t)$, it is then straightforward to integrate out the amplitudes in (45) to yield

$$\begin{aligned}\pi\left(\boldsymbol{\tau}_{1:n}, k_{1:n}, \sigma_u^2, \sigma_v^2\right) \propto & \int_{-\infty}^{\infty} \pi(\boldsymbol{\theta}_{1:n})\, d\boldsymbol{a}_1 d\boldsymbol{a}_2 ... d\boldsymbol{a}_t, \\ = & \prod_{j=1}^{n} \frac{1}{(2\pi\sigma_u^2)^{M/2}(1+\delta^2)^{k_j/2}} \exp\left\{\frac{-1}{2\sigma_w^2}\boldsymbol{z}_j^T \boldsymbol{P}_0^{\perp}(\boldsymbol{\tau_j})\boldsymbol{z}_j\right\} \times \\ & \prod_{j=1}^{n} \frac{1}{(2\pi\sigma_v^2)^{k_j/2}} \exp\left\{\frac{-1}{2\sigma_v^2}(\boldsymbol{\tau}_j - \boldsymbol{\tau}_{j-1})^T(\boldsymbol{\tau}_j - \boldsymbol{\tau}_{j-1})\right\} \times \\ & \prod_{j=1}^{n} p(k_j|k_{j-1}) \times (\sigma_v^2)^{\frac{-\nu_0}{2}-1} \exp\left\{\frac{-\gamma_0}{2\sigma_v^2}\right\} \times (\sigma_w^2)^{\frac{-\nu_1}{2}-1} \exp\left\{\frac{-\gamma_1}{2\sigma_w^2}\right\}.\end{aligned} \tag{50}$$

The MAP estimators of the variances can be readily obtained by comparing (50) with a product of Inverted Gamma distributions. Using the fact that the mode of the Inverted Gamma distribution is $\frac{\gamma}{\nu+1}$, it follows that

$$\sigma_{vMAP}{}^2(n) = \frac{\frac{\gamma_0}{2} + \frac{1}{2}\sum_{j=1}^{n}\left(\boldsymbol{\tau}_j - \boldsymbol{\tau}_{j-1}\right)^T\left(\boldsymbol{\tau}_j - \boldsymbol{\tau}_{j-1}\right)}{\frac{\nu_0}{2} + \frac{1}{2}\sum_{j=1}^{n} k_j + 1}, \tag{51}$$

$$\sigma_{wMAP}{}^2(n) = \frac{\frac{\gamma_1}{2} + \sum_{j=1}^{n}\boldsymbol{z}_j^T\boldsymbol{P}_0^{\perp}\left(\boldsymbol{\tau}_j\right)\boldsymbol{z}_j}{\frac{\nu_1}{2} + Mt + 1}. \tag{52}$$

We choose however to keep these parameters in the expression of the posteriori distribution in (50). Since the nuisance parameters can be estimated, we now define a new vector $\boldsymbol{\alpha}$ of parameters to sample with the particle filter, as

$$\boldsymbol{\alpha}_{1:n} \triangleq \left(\boldsymbol{\tau}_{1:n}, k_{1:n}\right). \tag{53}$$

Note that the probabilities in the term $\prod_{j=1}^{n} p(k_j|k_{j-1}) \equiv p(\epsilon_k(j))$ in (50) are independent of $\boldsymbol{\tau}$ and only weakly dependent on k. Hence, this term does not significantly affect the estimation of $\boldsymbol{\alpha}$ and is therefore ignored in subsequent analysis. Thus, our estimation problem is now independent of the parameters h and $\epsilon_k(j)$ in (35).

IV. Sequential MC (SMC)

In this section, we briefly describe the sequential Monte Carlo (SMC) procedure. For further details, interested readers should refer to [28], [29], [30].

The SMC procedure [29] is an extension of the Bayesian importance sampling scheme [14]. The SMC procedure allows a recursive approximation of the target distribution of a dynamic system with a set of time-varying importance weights, which are updated as new data arrive.

Since it may be difficult or impossible to draw samples directly from the desired distribution $\pi(\boldsymbol{\alpha}_n)$, an importance function $q(\boldsymbol{\alpha}_n)$ that is "easy-to-sample" is used instead to generate $N \gg 1$ samples. The distribution $q(\boldsymbol{\alpha}_n)$ must include the support of $\pi(\boldsymbol{\alpha}_n)$. The so-generated samples, also known as *particles*, are used to compute a set of importance weights $w(\boldsymbol{\alpha}_n^{(i)})$, $i = 1, ..., N$, that can numerically estimate the target distribution $\pi(\boldsymbol{\alpha}_n)$, i.e.,

$$\widehat{\pi}(d\boldsymbol{\alpha}_n) = \frac{\sum_{i=1}^{N} w(\boldsymbol{\alpha}_n^{(i)})\delta_{w(\boldsymbol{\alpha}_n^{(i)})}(d\boldsymbol{\alpha}_n)}{\sum_{i=1}^{N} w(\boldsymbol{\alpha}_n^{(i)})}, \tag{54}$$

where $(d\boldsymbol{\alpha}_n)$ is a small, finite region surrounding an $\boldsymbol{\alpha}_n$ of interest and $\delta_{w(\boldsymbol{\alpha}_n^{(i)})}$ is an indicator function defined as

$$\delta_{w(\boldsymbol{\alpha}_n^{(i)})}(d\boldsymbol{\alpha}_n) = \begin{cases} 1, & \text{if } \boldsymbol{\alpha}_n^{(i)} \in (d\boldsymbol{\alpha}_n), \\ 0, & \text{otherwise.} \end{cases} \tag{55}$$

For notational convenience, we adopt $w^{(i)}(n) = w(\boldsymbol{\alpha}_n^{(i)})$, $i = 1, ..., N$, from this point onwards. Denote a set of N normalized importance weights by $\tilde{w}^{(i)}(n)$, $i = 1, ..., N$, defined as follows

$$\tilde{w}^{(i)}(n) = \frac{w^{(i)}(n)}{\sum_{j=1}^{N} w^{(j)}(n)}. \tag{56}$$

The objective of the SMC procedure, otherwise known as *particle filtering* (PF), is to update the unnormalized weights $w^{(i)}(n)$ of the approximate posterior distribution in (56) recursively and sequentially, based on the arrival of new observations. It can be shown [18], [30], [31] that this set of weights can be recursively updated by the following expression

$$w^{(i)}(n) \quad = \quad \tilde{w}^{(i)}(n-1) \times \frac{p(\boldsymbol{z}_n|\boldsymbol{\alpha}_n^{(i)}) p(\boldsymbol{\alpha}_n^{(i)}|\boldsymbol{\alpha}_{n-1}^{(i)})}{q(\boldsymbol{\alpha}_n^{(i)}|\boldsymbol{\alpha}_{1:n-1}^{(i)}, \boldsymbol{z}_{1:n})}, \quad i = 1, ..., N, \tag{57}$$

where the terms in the numerator are the likelihood function and the prior distribution functions defined in (38) and (39), respectively, whereas the term in the denominator refers to the importance function.

To obtain an optimal importance function [29], [30], [31] that minimizes the overall variance of the weights, that includes the support of the target distribution $\pi(\boldsymbol{\alpha}_n)$, and allows the recursivity defined in (57), one can select it to be proportional to the distribution functions [27] in the numerator in (57), that is

$$q(\boldsymbol{\alpha}_n^{(i)}|\boldsymbol{\alpha}_{n-1}^{(i)}, \boldsymbol{z}_n) \sim p(\boldsymbol{z}_n|\boldsymbol{\alpha}_n^{(i)}) p(\boldsymbol{\alpha}_n^{(i)}|\boldsymbol{\alpha}_{n-1}^{(i)}). \tag{58}$$

To determine such an optimal importance function, we let $\mathcal{L}(\boldsymbol{\tau}_n)$ be the logarithm of the optimal importance function, i.e.,

$$\begin{aligned} \mathcal{L}\left(\boldsymbol{\tau}_n\right) &\triangleq \mathcal{L}_z(\boldsymbol{\tau}_n) + \mathcal{L}_\tau(\boldsymbol{\tau}_n), \\ &= \log p\left(\boldsymbol{z}_n|\boldsymbol{\tau}_{1:n}, \boldsymbol{z}_{1:n-1}\right) + \log p\left(\boldsymbol{\tau}_n|\boldsymbol{\tau}_{1:n-1}, \boldsymbol{z}_{1:n}\right). \end{aligned} \tag{59}$$

We use a second-order Taylor expansion on $\mathcal{L}\left(\boldsymbol{\tau}_n\right)$ about the sensibly chosen point $\boldsymbol{\tau}_{n-1}$ to give

$$\begin{aligned} \mathcal{L}\left(\boldsymbol{\tau}_n\right) \approx & \mathcal{L}_z\left(\boldsymbol{\tau}_{n-1}\right) + \mathcal{L}_\tau\left(\boldsymbol{\tau}_{n-1}\right) + \left(\nabla\mathcal{L}_z\left(\boldsymbol{\tau}_n\right) + \nabla\mathcal{L}_\tau\left(\boldsymbol{\tau}_n\right)\right)^T \left(\boldsymbol{\tau}_n - \boldsymbol{\tau}_{n-1}\right) + \\ & \frac{1}{2}\left(\boldsymbol{\tau}_n - \boldsymbol{\tau}_{n-1}\right)^T \left(\nabla^2\mathcal{L}_z\left(\boldsymbol{\tau}_n\right) + \nabla^2\mathcal{L}_\tau\left(\boldsymbol{\tau}_n\right)\right)\left(\boldsymbol{\tau}_n - \boldsymbol{\tau}_{n-1}\right), \end{aligned} \tag{60}$$

where $\nabla\mathcal{L}_x(\boldsymbol{\tau}_n) \in \mathcal{R}^{k\times 1}$ and $\nabla^2\mathcal{L}_x(\boldsymbol{\tau}_n) \in \mathcal{R}^{k\times k}$ are the gradient vectors and the Hessian matrix of $\mathcal{L}_x(\boldsymbol{\tau}_n)$, respectively, defined as follows

$$\nabla\mathcal{L}(\boldsymbol{\tau}_n) = \frac{\partial}{\partial\boldsymbol{\tau}_n}\left(\mathcal{L}_z(\boldsymbol{\tau}_n) + \mathcal{L}_\tau(\boldsymbol{\tau}_n)\right)\Big|_{\substack{\boldsymbol{\tau}_n=\boldsymbol{\tau}_{n-1} \\ \boldsymbol{a}_n=\boldsymbol{a}_{n-1}}}, \tag{61}$$

$$\nabla^2\mathcal{L}(\boldsymbol{\tau}_n) = \frac{\partial^2}{\partial\boldsymbol{\tau}_n^2}\left(\mathcal{L}_z(\boldsymbol{\tau}_n) + \mathcal{L}_\tau(\boldsymbol{\tau}_n)\right)\Big|_{\substack{\boldsymbol{\tau}_n=\boldsymbol{\tau}_{n-1} \\ \boldsymbol{a}_n=\boldsymbol{a}_{n-1}}}. \tag{62}$$

Analytical expressions for (61) and (62) can be found in [32]. Using (60), it can be shown [27] that an approximation to the optimal importance sampling function for $\boldsymbol{\tau}_n$ can be expressed as a Gaussian distribution with the following form

$$q(\boldsymbol{\tau}_n^{(i)}|\boldsymbol{\tau}_{n-1}^{(i)}, \boldsymbol{z}_n) \sim \mathcal{N}\left(\boldsymbol{m}_n^{(i)}, \boldsymbol{\Sigma}_n^{(i)}\right), \tag{63}$$

where

$$\boldsymbol{\Sigma}_n^{(i)} = -\left(\nabla^2\mathcal{L}_z(\boldsymbol{\tau}_n) + \nabla^2\mathcal{L}_\tau(\boldsymbol{\tau}_n)\right)^{-1}, \tag{64}$$

$$\boldsymbol{m}_n^{(i)} = \boldsymbol{\tau}_{n-1}^{(i)} + \boldsymbol{\Sigma}_n^{(i)}\left(\nabla\mathcal{L}_z(\boldsymbol{\tau}_n) + \nabla\mathcal{L}_\tau(\boldsymbol{\tau}_n)\right). \tag{65}$$

Unfortunately, the recursion in (57) quickly degenerates, so that after a few time steps, only a handful of particles have weights significantly different from zero. There are two procedures which can be used to reduce this degeneracy. One is the Sampling Importance Resampling (SIR), which resamples the particles according to their respective importance weights. It can be shown [31] that the resampling can be done very efficiently with order $\mathcal{O}(N)$ operations. Unfortunately, the trajectories with high importance weights are statisically selected many times, limiting the true statistical diversity amongst the particles. Alternatively, another procedure, which employs an MCMC step for each particle at time n, systematically re-introduces the statisical diversity amongst the resampled particles [18], [24], [28].

In summary, the proposed SMC approach is basically a combination of a sequential Bayesian importance sampling, the SIR procedure, and an MCMC step. Note that the only parameter in $\boldsymbol{\alpha}$ which is subjected to the sampling procedure is the ISD vector $\boldsymbol{\tau}$. The model order $k(n)$ is determined using a statistical testing procedure to be described in Section V. The desired amplitudes (which give the source waveforms), are estimated from (47), and the variances are estimated according to (51) and (52). We summarize these steps in the following table

---

Sequential Importance Sampling Algorithm

Initialization

For time $n = 1$,

- sample N particles $\boldsymbol{\tau}^{(i)}, i = 1, \ldots, N$ from $q(\cdot|\cdot)$.
- initialize the weights $w^{(i)}$, $i = 1, ..., N$ to $\frac{\pi(\boldsymbol{\tau}^{(i)})}{q(\boldsymbol{\tau}^{(i)})}$.
- normalize the weights to $\tilde{w}^{(i)}(n) = \frac{w^{(i)}(n)}{\sum_{j=1}^{N} w^{(j)}(n)}$.

Then for $n = 2, 3, ...$

1. *Sequential Importance Sampling Step*

(a) Sample N particles of $\boldsymbol{\tau}_n^{(i)}$ for $i = 1, 2, \ldots, N$ from the approximately optimal importance function given by (63) as follows

$$\boldsymbol{\tau}_n^{(i)} \sim q(\boldsymbol{\tau}_n^{(i)}|\boldsymbol{\tau}_{1:n-1}^{(i)}, \boldsymbol{z}_{1:n})$$

(b) Evaluate the importance weights for N particles as follows:

$$w^{(i)}(n) = \tilde{w}^{(i)}(n-1) \times \frac{p(\boldsymbol{z}_n|\boldsymbol{\tau}_n^{(i)})p(\boldsymbol{\tau}_n^{(i)}|\boldsymbol{\tau}_{n-1}^{(i)})}{q(\boldsymbol{\tau}_n^{(i)}|\boldsymbol{\tau}_{1:n-1}^{(i)}, \boldsymbol{z}_{1:n})},$$

and hence the normalized importance weights as follows:

$$\tilde{w}^{(i)}(n) = \frac{w^{(i)}(n)}{\sum_{j=1}^{N} w^{(j)}(n)}$$

2. *Sampling Importance Resampling Step*

Multiply/Supress the particles $\boldsymbol{\tau}^{(i)}(n)$ respectively with high/low importance weights $\tilde{w}^{(i)}(n)$ to obtain N random samples approximately distributed according to $\pi(\boldsymbol{\tau}_{1:n}^{(i)})$.

- Sample a vector of $\boldsymbol{l}$ distributed as:

$$P(l(j) = i) = w^{(i)}(n)$$

- Resample the particles with the index vector:

$$\boldsymbol{\tau}_n^{(i)} = \boldsymbol{\tau}_n^{(l(i))}$$

- Re-assign all the weights to $w^{(i)}(n) = \frac{1}{N}$

3. *MCMC Step*

Resample the particles using the MCMC procedure, with fixed model order $k = k(n)$, as follows. Here, we choose the Metropolis-Hastings (M-H) algorithm [14], [33], [34] for the sampling procedure. For each particle $\boldsymbol{\tau}_k^{(i)}(n)$ obtained from step 2, the algorithm samples a candidate $\boldsymbol{\tau}_k^{\star}$ according to the proposal distribution $q\left(\boldsymbol{\tau}_k^{\star}|\boldsymbol{\tau}_k(n)\right)$ which is given using (30) as

$$q\left(\boldsymbol{\tau}_k^{\star}|\boldsymbol{\tau}_k^{(i)}(n)\right) \sim \mathcal{N}\left(\boldsymbol{\tau}_k^{(i)}(n-1), \sigma_v^2 \boldsymbol{I}_k\right). \tag{66}$$

The candidate is accepted with probability

$$\alpha_{update}\left(\boldsymbol{\tau}_k^{(i)}(n), \boldsymbol{\tau}_k^{\star}\right) = \min\left\{1, r_{update}\right\}, \tag{67}$$

where r_{update} is given by [34]

$$r_{update} = \frac{\pi\left(\boldsymbol{\tau}_k^{\star}|\boldsymbol{z}^{\star}\right) q\left(\boldsymbol{\tau}_k^{(i)}(n)|\boldsymbol{\tau}_k^{\star}\right)}{\pi\left(\boldsymbol{\tau}_k^{(i)}(n)|\boldsymbol{z}_n\right) q\left(\boldsymbol{\tau}_k^{\star}|\boldsymbol{\tau}_k^{(i)}(n)\right)}. \tag{68}$$

Accordingly, substituting (50) into (68) yields

$$r_{update} = \frac{\exp\left\{-\frac{1}{2\sigma_w^2}\boldsymbol{z}_n^{\star T}\boldsymbol{P}_0^{\perp}(\boldsymbol{\tau}^{\star})\boldsymbol{z}_n^{\star}\right\}}{\exp\left\{-\frac{1}{2\sigma_w^2}\boldsymbol{z}_n^{T}\boldsymbol{P}_0^{\perp}(\boldsymbol{\tau}_n^{(i)})\boldsymbol{z}_n\right\}}. \tag{69}$$

If the candidate $\boldsymbol{\tau}_k^{\star}$ is accepted, then the current state becomes $\boldsymbol{\tau}_k^{(i)}(n) = \boldsymbol{\tau}_k^{\star}$. Otherwise, it remains at the current state, that is, $\boldsymbol{\tau}_k^{(i)}(n) = \boldsymbol{\tau}_k^{(i)}(n)$.

Note that the MCMC process normally requires a "burn–in" period for the chain to reach equilibrium. However in this case, this is not required, since the particles are already distributed according to the desired posterior distribution, which is the invariant distribution of the chain, before application of the MCMC procedure. Thus, it is necessary to apply only one MCMC step to each particle.

V. Model Order Detection

In this section, we present a rational technique that consists of two stages for determining which model, i.e. the number of sources, is most *consistent* with a given set of data. It is necessary that the number of sources be accurately determined before the particle filters can accurately estimate the DOA tracks. The proposed method assumes the observation noise variance σ_w^2 is known.

To facilitate the model order detection, we work on the data model that is expressed as a linear combination of individual sources, as in (20). For convenience, the model is reproduced here as follows

$$\boldsymbol{y}(n) = \sum_{k=0}^{k(n)-1} \tilde{\boldsymbol{H}}(\tau_k(n))\boldsymbol{s}_k(n) + \boldsymbol{w}(n), \tag{70}$$

$$= \hat{\boldsymbol{y}}(n) + \boldsymbol{w}(n), \tag{71}$$

where

$$\hat{\boldsymbol{y}}(n) = \sum_{k=0}^{k(n)-1} \tilde{\boldsymbol{H}}(\tau_k(n))\boldsymbol{s}_k(n). \tag{72}$$

Denote $\hat{\boldsymbol{y}}_{-k_0}(n)$ as follows

$$\hat{\boldsymbol{y}}_{-k_0}(n) = \sum_{k\neq k_0}^{k(n)-1} \tilde{\boldsymbol{H}}(\tau_k(n))\boldsymbol{s}_k(n), \tag{73}$$

which amounts to an estimate of the observation $\boldsymbol{y}(n)$ by excluding the contribution of the k_0th source.

In the model order detection, two hypotheses are considered at each time n. They are $\mathcal{H}_b$ and $\mathcal{H}_d$, which represent the birth of a new source and the death of an existing source, respectively. In the hypothesis $\mathcal{H}_b$, a new source in addition to the existing sources is proposed such that $k(n) \rightarrow k(n)+1$, whereas in the hypothesis $\mathcal{H}_d$, one of sources from the existing sources is proposed for removal such that $k(n) \rightarrow k(n)-1$. Deciding which hypothesis is taken amounts to testing whether the current model order fits the observation. Next, we present these procedures and the associated statistical testing proceduces in detail.

A. Death of an existing source

In this step, the existing $k(n)$ sources are sequentially removed, and the normalized squared residuals, $\epsilon_{-k_0}(n)$, for $k_0 = 0, ..., k(n)-1$, are each evaluated as follows

$$\epsilon_{-k_0}(n) = \frac{\boldsymbol{e}_y^T(n)\boldsymbol{e}_y(n)}{\sigma_w^2}, \tag{74}$$

where $\boldsymbol{e}_y(n)$ is the residual, defined as

$$\boldsymbol{e}_y(n) = \boldsymbol{y}(n) - \hat{\boldsymbol{y}}_{-k}(n). \tag{75}$$

Under the hypothesis that the k_0th source is zero, the quantity $\epsilon_{-k_0}(n)$ is χ^2-squared distributed with $M - k(n) + 1$ degrees of freedom, that is

$$\epsilon_{-k_0}(n) \sim \chi^2_{M-k(n)+1}. \tag{76}$$

Thus, it is possible to test whether a particular source contributes significantly to the observation $\boldsymbol{y}(n)$ at time n. The normalized error $\epsilon_{-k_0}(n)$ is compared to a threshold $c_d > 0$ to test whether the current model order $k(n)$ fits the data, that is

$$P(\epsilon_{-k_0}(n) > c_d) = \chi^2_{M-k(n)+1}(1-\alpha_d). \tag{77}$$

Here, the probability of the normalized error $\epsilon_{-k_0}(n)$ exceeding a threshold $c_d > 0$ is α_d. If $\epsilon_{-k_0}(n)$ is larger than c_d, it is likely that the model order $k(n)-1$ with the removal of the kth source is too low. In other words, the k_0th source should not be removed from the set of existing sources. On the other hand, if $\epsilon_{-k_0}(n)$ is smaller than c_d, it implies that the contribution of the k_0th source to the observation is probably insignificant and the removal of this source should be considered. However, in order to further justify the decision to remove the k_0th source, an additional statistical test is imposed. Since the signal amplitude is estimated according to (47), which is essentially a least-squared solution, if a source exists and contributes to the observation $\boldsymbol{y}(n)$ at time n, its corresponding amplitude $s_{k_0}(n)$, as well as the normalized signal power $\sigma^2_{s_{k_0}}(n)$, defined as

$$\sigma^2_{s_{k_0}}(n) = \frac{\boldsymbol{s}_{k_0}^T(n)\tilde{\boldsymbol{H}}^T(\tau_{k_0}(n))\tilde{\boldsymbol{H}}(\tau_{k_0}(n))\boldsymbol{s}_{k_0}(n)}{\sigma_w^2}, \tag{78}$$

should be persistently significant. On the other hand, if the source has indeed suddenly vanished, both the restored amplitude and the normalized signal power $\sigma^2_{s_{k_0}}(n)$ will become insignificant. To test the significance of $\sigma^2_{s_{k_0}}(n)$, which amounts to testing the presence of the source, an *F-test* [35] can be used.

An F-test, used to test if the estimated variances from two populations are equal, is given as follows

$$F = \frac{\sigma_1^2}{\sigma_2^2}, \tag{79}$$

where σ_1^2 and σ_2^2 are the estimates from the two populations. The more the ratio F deviates from 1, the stronger the evidence for unequal population variances. Given a significance level α_f, the hypothesis that the two variances are equal is rejected if

$$F > \mathcal{F}\left(\alpha_f/2, \beta_1 - 1, \beta_2 - 1\right), \tag{80}$$

where $\mathcal{F}\left(\alpha_f/2, \beta_1 - 1, \beta_2 - 1\right)$ is known as the critical value of the F distribution [35], where β_1 and β_2 are the degrees of freedom in the numerator and denominator in (78). The quantity α_f represents the significance level of the test. Therefore, to test whether $\sigma^2_{s_{k_0}}(n)$ is significant when compared with the noise variance, an F-test on $\sigma^2_{s_{k_0}}(n)$ is applied.

B. Birth of a new source

In the step, we test whether the current model order is too low to fit the observation $\boldsymbol{y}(n)$ at time n. Given the current model order $k(n)$, we will first compute the normalized residual as follows

$$\epsilon(n) = \frac{\boldsymbol{e}_y^T(n)\boldsymbol{e}_y(n)}{\sigma_w^2}, \tag{81}$$

where the residual $\boldsymbol{e}_y(n)$ is given by

$$\boldsymbol{e}_y(n) = \boldsymbol{y}(n) - \hat{\boldsymbol{y}}(n), \tag{82}$$

where $\hat{\boldsymbol{y}}(n)$ is given by (72). Under the hypothesis that the number of signals is correct, the quantity $\epsilon(n)$ is also χ^2-squared distributed with $M - k(n)$ degrees of freedom, that is

$$\epsilon(n) \sim \chi^2_{M-k(n)+1}. \tag{83}$$

To test whether the current model fits the current observation, the normalized error $\epsilon(n)$ is compared with a threshold $c_b > 0$, that is

$$P(\epsilon(n) > c_b) = \chi^2_{M-k(n)}(1 - \alpha_b). \tag{84}$$

The probability of the normalized error $\epsilon(n)$ exceeding a threshold $c_b > 0$ is α_b. If $\epsilon(n) < c_b$, it is likely that the current model is appropriate for the current observation, and there is no need to introduce a new source to the existing tracks. However, if $\epsilon(n) > c_b$, the current model order is likely too low to fit the observation, leading to larger estimation error. In this case, the current model order will be incremented by one to $k(n) = k(n) + 1$, and a new source will be generated according to importance sampling procedure described earlier.

Model Order Detection Using Hypothesis Testing

In summary, for $n = 1, 2, ...$, the MAP values of $\boldsymbol{\tau}(n)$ and $\boldsymbol{a}(n)$ obtained from the proposed Particle Filter method are used to detect the instantaneous model order, according to the following hypothesis testing procedure

Parameter	SNR (dB)	M	K	L	N	σ_w^2	σ_v^2	δ^{-2}	F_s (Hz)
Value	14	8	2	8	$1,000$	0.0169	$(3,000F_s)^{-2}$	25.12	1

TABLE 6.2

PARAMETERS FOR THE EXPERIMENT.

1. $\mathcal{H}_d$ – *Death of an existing source*

(a) evaluate $\epsilon_{-k}(n)$ in (74) for $k = 0, ..., k(n) - 1$, and compare it with the threshold c_d

- if $\epsilon_{-k}(n) < c_d$, the model order appears to fit the observation
- else, evaluate $\sigma^2_{s_k}(n)$ as in (78) and test it using a F-test

– if the F-test fails, the model order remains intact. Proceed with testing $\mathcal{H}_b$.

– else, remove the source k.

2. $\mathcal{H}_b$ – *Birth of a new source*

(a) evaluate $\epsilon(n)$ as in (81), and compare it with the threshold c_b

- if $\epsilon(n) < c_b$, the model order appears to fit the observation
- else, the current model order is probably too low,

– increment the model order by 1,

– re-run the sequential importance sampling steps for the new model order to estimate $\boldsymbol{\tau}(n)$ and $\boldsymbol{a}(n)$,

– once completed, by-pass the test for $\mathcal{H}_d$ and $\mathcal{H}_b$.

■

VI. Simulation Results

In the following simulations, a uniform linear array composed of $M = 8$ elements with a half-wavelength spacing of the elements at the highest frequency component is used in the simulations. The wideband signals in these experiments are bandlimited to the normalized frequency values as follows

$$f \in [0.1, 0.4]\,,$$

thus the normalized bandwidth of the signals is 0.3. The simulation environment is defined using the parameters listed in Table 6.2.

In these simulations, we demonstrate the ability of the proposed algorithm in jointly tracking the inter-sensor delays (ISDs) $\boldsymbol{\tau}(n)$ of the sources[1], detecting the model order $k(n)$, and recovering the source waveforms $\boldsymbol{a}(n)$, for the wideband scenario. The ISDs of the $K = 2$ sources are generated as autogressive (AR) processes whose coefficients represent a 10th–order low–pass

[1]Once the ISDs of the sources are available, the directions-of-arrival (DOAs) of the sources $\boldsymbol{\phi}(n)$ are determined through (2).

AR coefficients for ISD generation	AR coefficients for signal amplitude generation
1	1
-7.9923	-3.9877
28.9122	8.0944
-62.3154	-10.4736
88.5877	9.4233
-86.7671	-6.0842
59.2810	2.8353
-27.8903	-0.9364
8.6457	0.2089
-1.5942	-0.0283
0.1328	0.0018

TABLE 6.3

THE AR COEFFICIENTS FOR THE GENERATION OF THE ISDS AND SIGNAL AMPLITUDE PROCESSES.

Butterworth filter, with normalized cutoff frequency 0.1. The variance of the AR process is σ_v^2. The associated AR coefficients are specified in Table 6.3. Likewise, the source waveforms are also generated AR processes, corresponding to a 10th–order low–pass butterworth filter with normalized cutoff frequency 0.3 and variance $\delta^2\sigma_w^2$. The associated waveform AR coefficients are also listed in Table 6.3.

To demonstrate the algorithm's ability to detect the time-varying model order, the ISD trajectories $\boldsymbol{\tau}_t$ are constructed so that one of the sources suddenly vanishes or appears. The evolution of the model order $k(n)$ is first tracked by the proposed statistical testing procedure. Then the ISDs with the latest model order are tracked using the particle filter procedure. The source waveforms are then extracted using (47) and (49). Two different scenarios are presented in demonstrating the ability of the algorithm in handling the model order detection in different situations.

A. Scenario 1: A source suddenly appears

In this scenario, one source suddenly appears during the course of the ISD tracking. Fig. 2 shows the trajectories of the ISDs of the two sources. Before the algorithm starts, the parameters are appropriately initialized as described earlier in the SMC procedure with $k(0) = 1$. As the algorithm starts for $n \geq 1$, only one source is found, which is $\tau_0(0) = -0.9T_s$. According to Figs. 2 and 3, both the ISD (the lower trajectory in Fig. 2) and the signal amplitude of this source $s_0(n)$ are well estimated by the algorithm in the time region $n \in [1, 37]$. When $n > 37$, a new source $s_1(n)$ appears and the corresponding ISD is approximately $0.2T_s$. Based on the statistical testing procedure, the algorithm does not realize the presence of the new source until

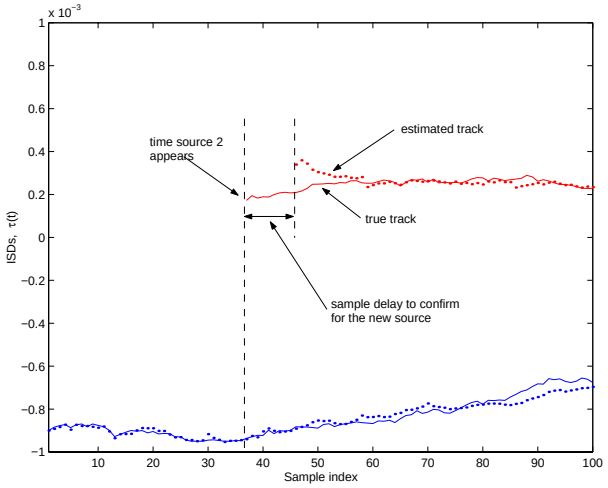

Fig. 2. The tracking of ISDs $\boldsymbol{\tau}(n)$ for scenario 1.

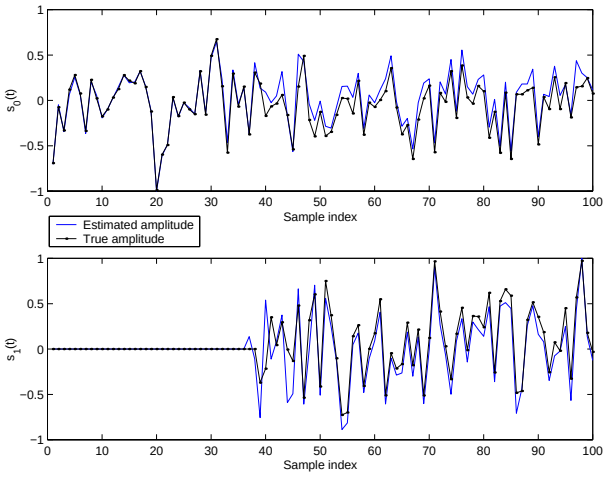

Fig. 3. The signal recovery of $\boldsymbol{a}(n)$ for scenario 1.

$n > 45$, as shown in Fig. 4. After the new source has been found, it takes approximately 15 additional time indicies for the algorithm to stabilize as shown in Fig. 2 before the new source is well tracked. Note that in the course of the process, the noise variances σ_v^2 and σ_w^2 can also be sequentially estimated using the MAP procedure in (51) and (52), respectively.

The mean-squared errors for the ISD estimation and the waveform recovery process of the sources are shown in Table 6.4, corresponding to 50 independent trials of the same simulation scenario. As one would expect, the errors from tracking and recovering the waveform of only one source are much smaller than in the case of multiple sources, especially when the SNR is about 14dB. Moreover, given the MAP procedure stated in (47), the performance of waveform recovery is heavily dependent on that of the ISD tracking, as evident from Table 6.4.

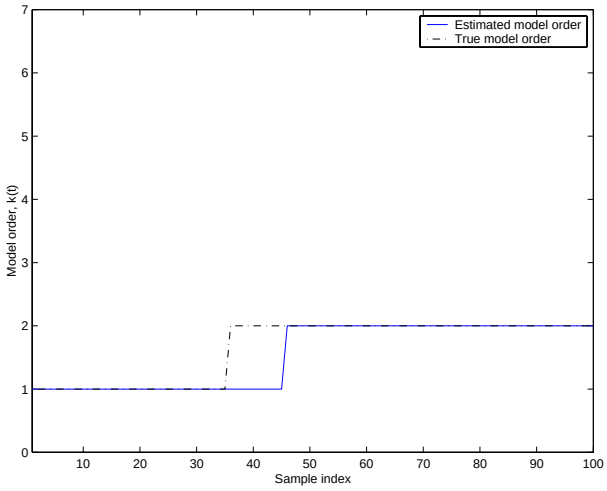

Fig. 4. The tracking of the model order $k(n)$ for scenario 1.

Time period	MSE (dB) ISD	MSE(dB) Amplitude
$n \in [1, 37]$	-27.4, –	-32.6, –
$n \in [38, 44]$	-26.3, –	-28.1, –
$n \in [45, 100]$	-33.0, -32.5	-16.5, -18.8

TABLE 6.4

THE MSE BETWEEN THE TRUE AND ESTIMATED ISDS OVER 50 INDEPENDENT TRIALS FOR THE PROPOSED METHOD FOR THE SCENARIO 1.

B. Scenario 2: A source suddenly vanishes

In this scenario, one source suddenly vanishes during the ISD tracking of all sources. The initial ISD values of the sources are $\boldsymbol{\tau}(0) = [0.05, -0.65] \times T_s$. As the algorithm starts when $n \geq 1$, these two sources are detected, and their ISD trajectories and waveform recovery are shown in Figs. 5 and 6, respectively. According to these figures, in the time period $n \in [1, 20]$, the algorithm tracks the ISDs and recovers their waveforms very well. At time $n > 20$, one of the sources, $s_1(n)$, the upper trajectory in Fig. 7, suddenly disappears. It takes about 11 time indicies for the algorithm to respond to this change and reduce the model order by 1. During the time period $n \in [21, 32]$, the ISD estimates and hence the restored waveform for the vanished source are essentially noise processes, whose variances are about σ_v^2 and σ_w^2, respectively. In the time period $n \in [33, 100]$, the algorithm tracks the ISD of the remaining source and recovers the waveform very well, as evident in Figs. 5 and 6.

We also run 50 independent trials for the same setup to evaluate the mean-squared errors for the ISD estimation and the waveform recovery process for the sources. Table 6.5 shows the results. Similar to the findings in the last scenario, the MSEs for both the ISD estimation and waveform recovery are larger when the algorithm tracks two sources than that when it tracks

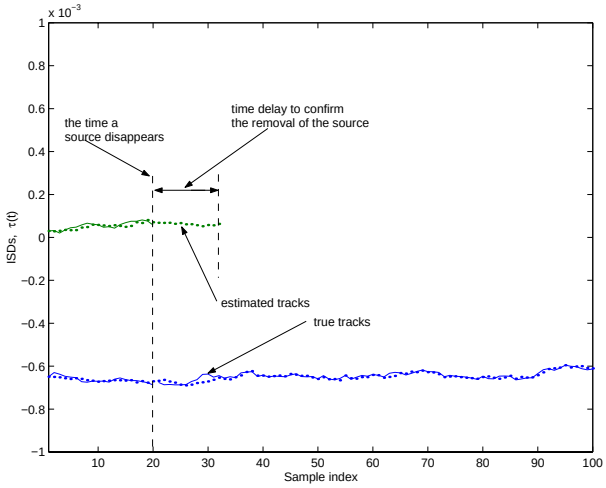

Fig. 5. The tracking of ISDs $\boldsymbol{\tau}(n)$ for scenario 2.

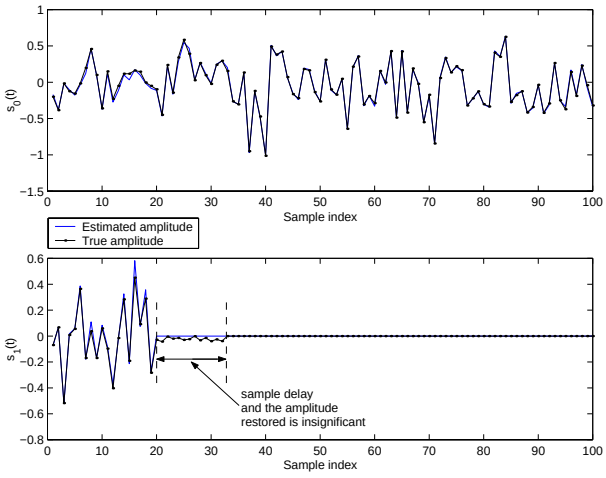

Fig. 6. The signal recovery of $\boldsymbol{a}(n)$ for scenario 2.

one only source.

VII. Performance Evaluation

Since there appears to be no previous method which offers the same functionality as the proposed approach, it is difficult to compare performance with previous approaches. Therefore, to investigate the performance of the proposed algorithm, we resort to simulations to compare proximity of the results to lower bounds corresponding to optimum performance. Since the problem in question is nonstationary, we resort to the posterior Cramér-Rao bound (PCRB) [2], derived for discrete-time nonlinear dynamical systems. The development of the recursive update equation for the variances of the parameters can be found in [32].

We now present the comparison between the performance of the proposed method in terms of the estimated variances of $\boldsymbol{\tau}(n)$ as a function of SNR and the PCRB. In this simulation, we use the parameter values listed in Table 6.2, except for SNR. We arbitrarily select a particular

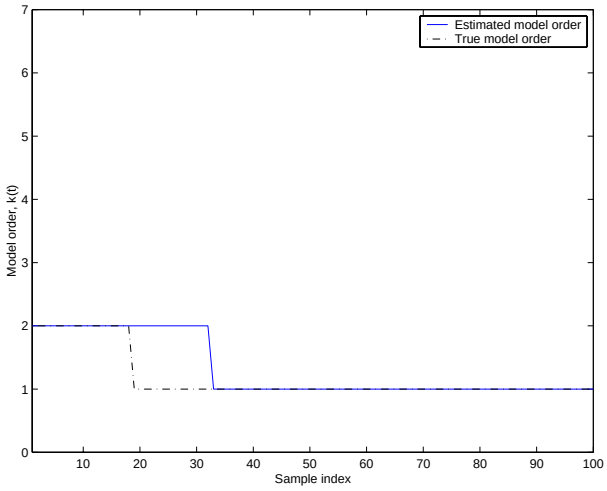

Fig. 7. The tracking of the model order $k(n)$ for scenario 2.

Time period	MSE (dB) ISD	MSE(dB) Amplitude
$n \in [1, 37]$	-29.4, -33.0	-22.6, -23.7
$n \in [38, 44]$	-33.1, –	-20.6, –
$n \in [45, 100]$	-34.7, –	-25.7, –

TABLE 6.5

THE MSE BETWEEN THE TRUE AND ESTIMATED ISDS OVER 50 INDEPENDENT TRIALS FOR THE PROPOSED METHOD FOR THE SCENARIO 2.

time sample at $n = 20$ for the evaluation of the variances of $\boldsymbol{\tau}(n)$. The algorithm is run for 100 independent trials over a range of SNR, from -5dB to 20 dB. Each trial uses 50 observations and the number of sources is assumed known. Fig. 8 shows the comparision between the estimated variances and the PCRB. As shown in Fig. 8, for SNR levels lower than -2dB, the algorithm starts to break down (i.e., departs rapidly from the PCRB). However, it is seen that the variances approach the PCRB closely, above this level. The reasons why the variances do not come closer to the theoretical PCRB are: 1) interpolation errors due to a non-ideal interpolation function being used and 2) the procedure for estimating the source amplitudes. The errors introduced in the signal recovery have an impact on the estimation accuracy of the ISD parameters, since the source amplitudes $\boldsymbol{a}(n)$ affect the transformed data $\boldsymbol{z}(n)$, as evident from (28). The quantities $\boldsymbol{z}(n)$ are then used in estimating the $\boldsymbol{\tau}(n)$ through the posterior distribution (50).

We also include a few more figures to compare the performance of the algorithm with the PCRB for $n \in [1, 50]$ for different SNR levels, as shown in Figs. 9, 10, and 11. The trajectories of the PCRB and the estimated variances of the error in the estimated $\boldsymbol{\tau}(n)$ are obtained by running 1,000 trials. We see that as SNR levels decrease, the gap between the estimated variances and the PCRB widens, which is consistent with the findings in Fig. 8.

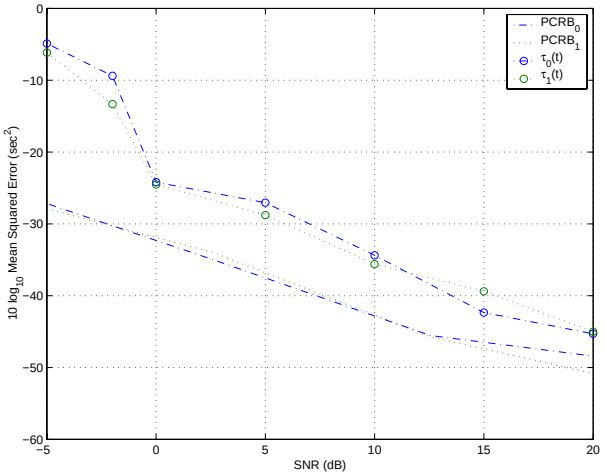

Fig. 8. The performance of the algorithm as a function of SNR for the sample at $n = 20$.

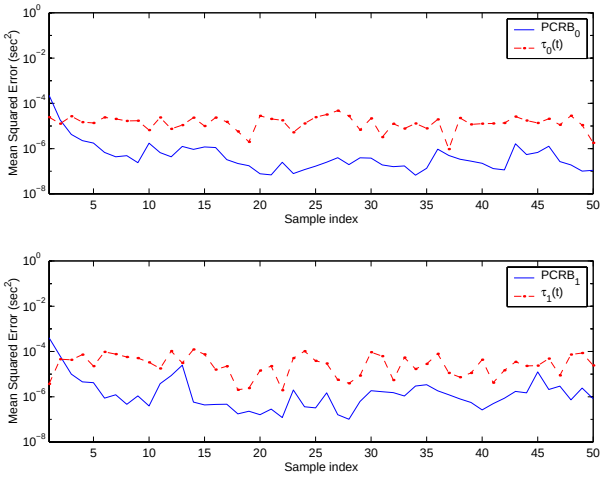

Fig. 9. A comparison between the estimate error variances of $\boldsymbol{\tau}(n)$, and the PCRB for SNR = 15 dB.

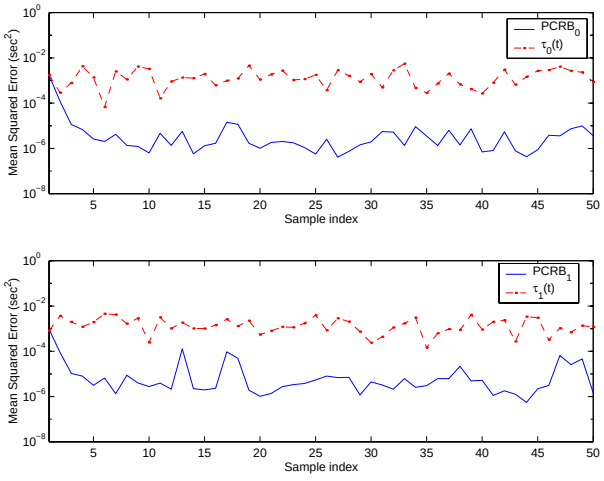

Fig. 10. A comparison between the estimate error variances of $\boldsymbol{\tau}(n)$, and the PCRB for SNR = 0 dB.

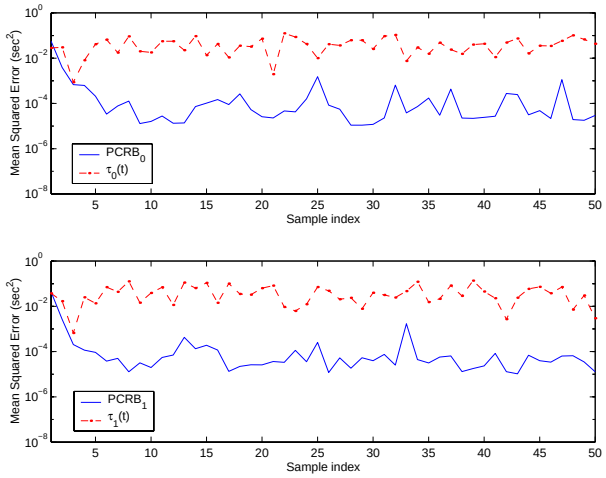

Fig. 11. A comparison between the estimate error variances of $\boldsymbol{\tau}(n)$, and the PCRB for SNR = -5 dB.

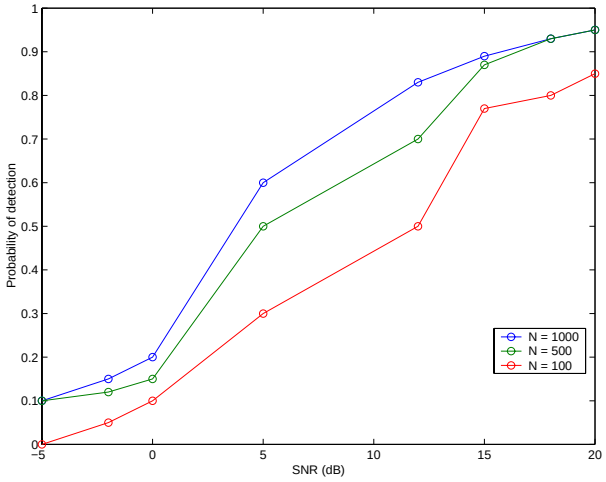

Fig. 12. The probability of detection of the number of unknown sources versus SNR levels for different numbers of particles.

We also investigate the detection performance of the algorithm as a function of SNR level. Fig. 12 clearly reveals that the larger the SNR level, the higher is the probability of correct detection of the number of sources. In addition, given an SNR level, the performance is dependent on the number of particles used in the particle filter. The larger the number of particles, the better is the performance in model order detection. In other words, in order to have a consistent tracking peformance that heavily relies on the consistent model order detection, one needs to use a large number of particles in the algorithm.

VIII. Conclusion

A new approach to wideband array signal processing is proposed that is based on an interpolation process to approximate wideband signals. A Bayesian approach is used for estimation of the parameters of interest, where a marginalized posterior density function is formulated.

The time-varying ISDs are tracked by particle filters, for which a novel second-order importance sampling function is proposed. The model order is determined by a hypothesis testing procedure. The source amplitudes are estimated using a MAP estimate. Simulation results support the effectiveness of the method, and demonstrate reliable detection of the number of sources and estimation of their ISDs in white noise environments with a single linear array. As a result, the source signals are reliably restored. Furthermore, a theoretical lower bound on error variances (the posterior Cramér-Rao bound, PCRB) is developed to evaluate the performance of the proposed algorithm. It has been shown that the performance of the proposed method is comparable to that of the PCRB.

References

[1] W. Ng, J. P. Reilly, T. Kirubarajan, and J.-R. Larocque, "Wideband array signal processing using MCMC methods," 2002. Submitted to *IEEE Transactions on Signal Processing*, Jan. 2003, also available at http://www.ece.mcmaster.ca/∼reilly.

[2] P. Tichavsky, H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1386–1396, May 1998.

[3] D. Johnson, "The application of spectral estimation to bearing estimation problem," *Proceedings of the IEEE*, vol. 70, pp. 1018–1028, Sept. 1982.

[4] B. D. V. Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," Apr. 1988.

[5] H. L. Van Trees, *Optimum Array Processing*. New York: John Wiley & Sons, 2002.

[6] S. Valaee and P. Kabal, "A unitary transformation algorithm for wideband array processing," in *Proceedings of Sixth IEEE SP Workshop on Statistical Signal & Array Processing*, (Victoria, BC), pp. 300–303, Oct. 1992.

[7] S. Valaee and P. Kabal, "Wideband array processing using a two-sided correlation transformation," *IEEE Transactions on Signal Processing*, vol. 44, pp. 160–172, Jan. 1995.

[8] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, pp. 823–831, Aug. 1985.

[9] H. Hung and M. Kaveh, "Focusing matrices for coherent signal-subspace processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1272–1281, Aug. 1988.

[10] S. Haykin, *Advances in Spectrum Analysis and Array Processing: Volume II*. Englewood Cliffs, NJ: Prentice Hall, 1991.

[11] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Transactions on Signal Processing*, vol. 47, pp. 2667–2676, Oct. 1999.

[12] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer Verlag, 1999.

[13] C. Andrieu, P. Djuric, and A. Doucet, "Model selection by MCMC computation," *Signal Processing*, vol. 81, pp. 19–37, Jan. 2001.

[14] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. New York: Chapman and Hall, 1998.

[15] A. Gershman, "Robust adaptive beamforming in sensor arrays," in *Int. Journal of Electronics and Communications*, vol. 53, pp. 305–314, 1999.

[16] V. Katkovnik and A. B. Gershman, "A local polynomial approximation based beamforming for source localization and tracking in nonstationary environments," vol. 7, pp. 3–5, Jan. 2000.

[17] T. Wigren and A. Eriksson, "Accuracy aspects of DOA and angular velocity estimation in sensor array processing," vol. 38, pp. 60–62, Apr. 1995.

[18] J.-R. Larocque, J. P. Reilly, and W. Ng, "Particle filter for tracking an unknown number of sources," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2926–2937, Dec. 2002.

[19] S. Haykin, *Array Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1985.

[20] S. Swales, J. Maloney, and J. Stevenson, "Locating mobile phones and the US wireless E-911 mandate," in *IEE Colloquium on Novel Methods of Location and Tracking of Cellular Mobiles and Their System Applications*, pp. 2/1 – 2/6, May 1999.

[21] J. Reed, K. Krizman, B. Woerner, and T. Rappaport, "An overview of the challenges and progress in meeting the E-911 requirement for location service," vol. 36, pp. 30–37, Apr. 1988.

[22] X.-R. Li, Y. Bar-Shalom, and T. Kirubarajan, *Estimation, Tracking and Navigation: Theory, Algorithms and Software*. New York: John Wiley & Sons, June, 2001.

[23] T. Kirubarajan and Y. Bar-Shalom, "Low observable target motion analysis using amplitude information," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, pp. 1367–1384, July 1996.

[24] C. Andrieu, N. D. Freitas, and A. Doucet, "Sequential MCMC for Bayesian model selection," in *Proceedings of the Internationl Workshop on Higher Order Statistics*, (Ceasarea, Israel), 1999.

[25] W. Ng, *Advances in Wideband Array Signal Processing Using Numerical Bayesian Methods*. PhD thesis, McMaster University, Hamilton, Ontario, Canada, 2003.

[26] G. H. Golub and C. F. V. Loan, *MATRIX Computations*. Baltimore, MA: The Johns Hopkins University Press, 2nd ed., 1993.

[27] M. Orton and W. Fitzgerald, "A Bayesian apporach to tracking multiple targets using sensor arrays and particle filters," *IEEE Transactions on Signal Processing*, vol. 50, pp. 216–223, Feb. 2002.

[28] A. Doucet, N. de Freitas, and N. Gordon, eds., *Sequential Monte Carlo in Practice*. New York: Springer-Verlag, 2001.

[29] N. Gordon, D. Salmond, and A. Smith, "Novel approach to non-linear/non-Gaussian Bayesian state estimation," *IEE Proceedings-F*, vol. 140, no. 2, pp. 107–113, 1993.

[30] A. Doucet, "On sequential simulation-based methods for Bayesian filtering.," Tech. Rep. TR.310, University of Cambridge, Department of Engineering, Signal Processing Group, England, 1998.

[31] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[32] W. Ng, J. P. Reilly, and T. Kirubarajan, "The derivation of the second order importance sampling function," May 2003. Please download the document at http://www.ece.mcmaster.ca/∼reilly.

[33] J. O. Ruanaidh and W. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing.* New York: Springer-Verlag, 1996.

[34] W. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[35] G. W. Snedecor and W. G. Cochran, *Statistical Methods*. Ames, Iowa: Iowa State University Press, 7th ed., 1980.