A distributed particle filtering approach for multiple acoustic source tracking using an acoustic vector sensor network

Xionghu Zhong [a,*], Arash Mohammadi [b], A.B. Premkumar [c], Amir Asif [d]

[a] Centre for Multimedia and Network Technology, The School of Computer Engineering, College of Engineering, Nanyang Technological University, Singapore 639798, Singapore
[b] Department of Electrical and Computer Engineering, University of Toronto, Canada M5S 2J7
[c] Electrical Engineering Department, Faculty of Engineering University of Malaya, Kuala Lumpur, Malaysia
[d] Department of Computer Science and Engineering, York University, Canada M3J 1P3

ARTICLE INFO

Article history:
Received 2 September 2013
Received in revised form
21 September 2014
Accepted 24 September 2014
Available online 2 October 2014

Keywords:
Acoustic vector sensor
Sensor network
Particle filtering
Acoustic source tracking
Distributed posterior Cramér–Rao bound

ABSTRACT

Different centralized approaches such as least-squares (LS) and particle filtering (PF) algorithms have been developed to localize an acoustic source by using a distributed acoustic vector sensor (AVS) array. However, such algorithms are either not applicable for multiple sources or rely heavily on sensor-processor communication. In this paper, a distributed unscented PF (DUPF) approach is proposed for multiple acoustic source tracking. At each distributed AVS node, the first-order and the second-order statistics of the local state are estimated by using an unscented information filter (UIF) based PF. The UIF is employed to approximate the optimum importance function due to its simplicity, by which the matrix operation is the state information matrix rather than the covariance matrix of the measurement sequence. These local statistics are then fused between neighbor nodes and a consensus filter is applied to achieve a global estimation. In such an architecture, only the state statistics need to be transmitted among the neighbor nodes. Consequently, the communication cost can be reduced. The distributed posterior Cramér–Rao bound is also derived. Simulation results show that the performance of the DUPF tracking approach is similar to that of centralized PF algorithm and significantly better than that of LS algorithms.

1. Introduction

Traditionally, multiple acoustic sources are tracked using either a single or multiple array setup comprising of several pressure sensors. In the case of a single array, a large number of pressure sensors are needed to provide enough aperture for accurate estimation. Alternatively, a combination of multiple hybrid arrays of pressure sensors can be used. Utilizing the correlation information between pressure measurements recorded from different arrays in the processing algorithms adds to the complexity of the multiple array setup. Irrespective of the configuration used, the pressure measurements obtained from the arrays are processed using an estimation algorithm such as the Kalman filter (KF) and particle filter (PF) to track the acoustic sources. In recent years, acoustic vector sensors (AVS) [1,2] that have the capability of measuring both particle vibration velocity and acoustic pressure at the sensor's location have been widely employed for acoustic source detection and localization. Accordingly,

* Corresponding author.
E-mail addresses: xhzhong@ntu.edu.sg (X. Zhong), marash@cse.yorku.ca (A. Mohammadi), benjamin.premkumar@gmail.com (A.B. Premkumar), asif@cse.yorku.ca (A. Asif).

http://dx.doi.org/10.1016/j.sigpro.2014.09.031

several signal processing approaches [3–21] have been developed for source localization and tracking using AVSs. These approaches are, however, centralized and based on a fusion centre. In such a centralized configuration, all pressure measurements have to be communicated to the fusion centre adding unnecessary latency to the estimation mechanism. The computational complexity of the central approaches is also fairly high with all processing completed at the fusion node. The fusion node is, therefore, critical and its failure makes the network in-operational. Compared to the centralized estimation techniques, distributed estimation, where all nodes collaborate in the estimation process by locally processing the pressure measurements and combining the local results, offers several advantages, including scalability, higher immunity to network failure, and dynamic adaptability to changes in the network topology. These recent advances in the sensor network technology motivates the proposed deployment of distributed sensor network for acoustic source localization [22]. In this paper, we derive a novel multiple acoustic source tracking algorithm based on the unscented particle filter for the distributed AVS network and show that its performance is comparable with the centralized approach.

1.1. *Acoustic vector sensor*

The AVS employs a co-located sensor structure which consists of two or three orthogonally oriented velocity sensors and a pressure sensor [1,2]. Given an AVS with three velocity components that is located at the origin of the 3-D (x-, y- and z-coordinates) space, the sensor manifold $\boldsymbol{u} \in \mathbb{R}^{4\times 1}$ has the following form:

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ x(\phi,\psi) \\ y(\phi,\psi) \\ z(\psi) \end{bmatrix} \triangleq \begin{bmatrix} 1 \\ \cos\phi\cos\psi \\ \sin\phi\cos\psi \\ \sin\psi \end{bmatrix}, \tag{1}$$

where the first component represents the pressure measurement, and the other three components are the particle velocities. The 2-D direction of arrival (DOA) $\phi \in (-\pi, \pi]$ and $\psi \in [-\pi/2, \pi/2]$ are the azimuth and elevation angles, respectively. The manifold suggests that the AVS has following advantages over pressure sensors:

1. It produces both azimuth and elevation angle information and enables 2-D DOA estimation with a single AVS.
2. It allows elevation angle estimation unambiguously.
3. The manifold is independent of the frequency of the source signal, which makes AVS suitable for wideband source signal or scenarios where the frequency of the source signal is unknown *a priori*.

Due to a number of advantages mentioned above, both the theoretical aspects and the applications of AVS have been widely studied during the last decade [3–21]. An intensity based algorithm and a velocity covariance based algorithm were firstly presented in [3]. Maximum likelihood based DOA estimation algorithms are developed in [23,24]. The beamforming approaches [7] and the subspace based approaches such as MUSIC [10,15] and ESPRIT [5,8,10,13] have been developed for 2-D DOA estimation problem. Tracking the DOA of a single acoustic source has been recently studied in [25–27]. In the authors' previous work [28–31], different particle filtering (PF) approaches are developed for DOA tracking. Applications of the AVS in room acoustic signals and underwater acoustic environment are investigated in [32] and [33,34], respectively. The AVS signal based source localization in impulsive noise environment has been considered in [35]. The authors in [36] employ a towed AVS to track the angles and frequencies of sperm whales. However, such investigations focus only on the 2-D DOA estimation, rather than the 3-D position in space.

1.2. *Centralized approaches for 3-D position estimation*

Recently, an AVS array consisting of a number of spatially distributed AVSs has been deployed for 3-D position estimation problem and, accordingly, different centralized approaches have been developed [13,37,38]. In [13], least-squares (LS) approaches have been derived for wideband acoustic source localization based on the distributed AVS array. The 2-D DOAs of the source are firstly estimated at each distributed AVS by using the Capon beamforming method [7]. The weighted least-squares (WLS) and reweighted least squares (RWLS) algorithms are then developed to obtain the 3-D location parameters at the central processor (CP) by triangulating the DOAs. Such a centralized approach is illustrated in Fig. 1(a). The DOAs are estimated at each local node, and then transmitted to a CP. The triangulation approaches are implemented in the CP to estimate the 3-D position of the source. These algorithms can be categorized into *indirect* methods since the position is estimated from the DOA estimates rather than from the received signals directly. The advantage of indirect methods is that each AVS needs only to transmit the DOA estimates to the CP and thus the communication cost can be reduced significantly. However, such approaches cannot be applied for multiple source position estimation unless the DOA estimates can be correctly associated to each source. Also, the accuracy of the triangulation can seriously degrade due to the inaccurate DOA estimates at the local AVSs.

In [38,39], a PF approach is developed to track the 3-D position of multiple sources using the same array configuration. It can be referred to as a *direct* method since all received signals are directly transmitted to the CP to estimate the source position. There is no need to preprocess the data to extract the DOA estimates. The advantages of the proposed approach is that it is able to track dynamic sources and, in particular, it can be applied to track multiple wideband acoustic sources. Also, compared to the indirect approaches, it is able to provide better estimation accuracy. However, transmitting all received signals to the CP and processing them can be cumbersome since both the communication and the computation cost can be high. It is worth mentioning that in [37], a random finite set (RFS) approach is developed to jointly detect and track a time-varying number of multiple sources. Also, distributed deployment of the sensor components of an AVS has been studied by considering the acoustic pressure sensor component placed far away from the velocity-triad [40] and sparsely located linear AVS array [41].

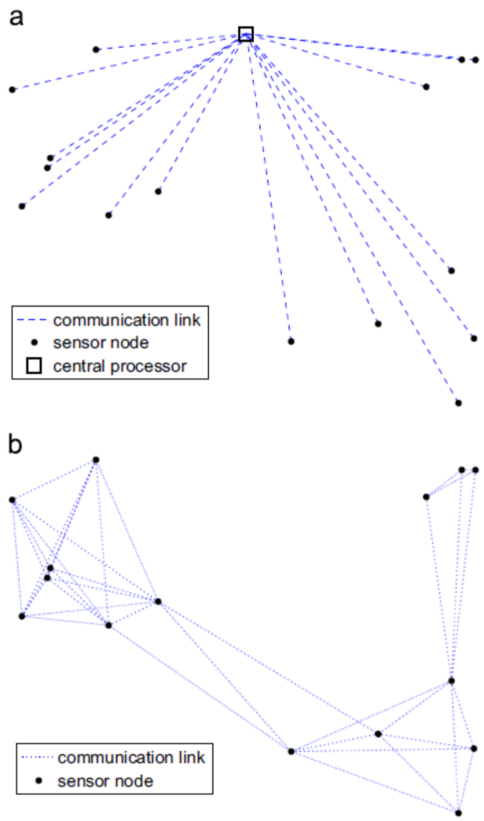

Fig. 1. Illustration of different AVS network configurations. (a) Centralized configuration and (b) decentralized configuration.

1.3. Proposed distributed tracking approach for AVS network

In this paper, a diffusive estimation scheme where each node diffuses its local first- and second-order statistics throughout the network by communicating with its neighbor nodes is developed. Such a diffusive estimation approach is able to fully exploit the merits of the distributed AVS array/network, e.g., low energy consumption and high robustness and scalability. Fig. 1(b) gives an illustration of the network structure. At each distributed AVS node, the statistics of the state are estimated by using an unscented particle filter (UPF). Considering that a number of measurements are used in acoustic applications, an unscented information filter (UIF) [42] is employed to approximate the optimum importance function due to its simplicity. The UIF needs only to perform the inversion of a small matrix, i.e., the covariance matrix of the state noise process, while the UKF requires the inverse operation of the covariance matrix of the measurement noise process that can be computationally very expensive. The estimated statistics are then employed for formulating an optimum importance function and a PF is applied to refine the local statistics estimation.

A consensus-based distributed estimation approach is then proposed for the AVS network to obtain global position estimates. Basically, consensus filtering is the process of establishing a consistent value for the statistics of the state across the network by exchanging relevant information between the connected neighbor nodes. In this work, we focus on a distributed implementation of the UIF in which the global estimates are obtained using a consensus filter. In such an architecture, the information that needs to be transmitted among the AVS nodes is only the first-order and the second-order statistics. Hence, the communication cost can be reduced. Also, the local states are estimated *directly* from the AVS signals.

For sensor selection decisions, the posterior Cramér–Rao bound (PCRB) has been utilized as an effective criterion [43–45] since it can be computed predictively and is independent of estimation mechanism. The need for computing the PCRB distributively occurs in the distributed/decentralized sensor networks, where sensor resource management (sensor selection) issues are often based on the online computation of the PCRB for the specific system configuration. Following our work in [46,47], the PCRB of the distributed estimation (referred to as the DPCRB) is derived via a distributed factorization of the posterior distribution. Through simulations, we show that the proposed bound is a little higher than its centralized counterpart in both single source and multiple source scenarios. The justification for this observation is provided in the experimental section.

The core contribution of this work is that a distributed tracking algorithm is developed to track multiple acoustic sources using an AVS network, and the corresponding DPCRB has also been derived. The rest of this paper is organized as follows. In Section 2, the AVS signal model and centralized PF (CPF) method are introduced. Section 3 presents the distributed algorithm development for multiple source tracking. Derivation of the DPCRB is presented in Section 4. Simulations are organized in Section 5. Conclusions and future directions of this work are discussed in Section 6.

2. Problem formulation

This section presents the AVS signal model. The distributed AVS network and the position estimation problem are then addressed. Note that general assumptions for the source signal and the noise process are the same as those in [13].

2.1. AVS signal model

At discrete time instance k, assume that N AVSs are deployed at arbitrary distinct 3-D locations $\mathbf{x}_0^n=[x_0^n, y_0^n, z_0^n]^T \in \mathbb{R}^{3\times1}$, for $n=1,\ldots,N$ to receive the signals emitted by M sources located at $\mathbf{x}_{m,k}=[x_{m,k}, y_{m,k}, z_{m,k}]^T \in \mathbb{R}^{3\times1}$, for $m=1,\ldots,M$. According to the sensor-source geometry, the 2-D DOAs of source signals $\boldsymbol{\theta}_{m,k}^n=[\phi_{m,k}^n, \psi_{m,k}^n]^T \in \mathbb{R}^{2\times1}$ are related to the source position by

$$\phi_{m,k}^n=\tan^{-1}\left(\frac{x_{m,k}-x_0^n}{y_{m,k}-y_0^n}\right);$$

$$\psi_{m,k}^n = \tan^{-1}\left(\frac{z_{m,k}-z_0^n}{\sqrt{(x_{m,k}-x_0^n)^2+(y_{m,k}-y_0^n)^2}}\right); \quad (2)$$

where $\phi_{m,k}^n \in (-\pi, \pi]$ and $\psi_{m,k}^n \in [-\pi/2, \pi/2]$ represent the azimuth and the elevation angles, respectively, and the superscript T denotes the transpose. The complex acoustic source signals $s_m(k) \in \mathbb{C}$ are assumed to be wideband and independent of each other. Let $\boldsymbol{u}_m^n(k)$ be the unit direction vector pointing from the nth sensor toward to the mth source and given as

$$\boldsymbol{u}_m^n(k) = \begin{bmatrix} \cos\psi_{m,k}^n \cos\phi_{m,k}^n \\ \cos\psi_{m,k}^n \sin\phi_{m,k}^n \\ \sin\psi_{m,k}^n \end{bmatrix}. \quad (3)$$

The received signal model for nth AVS can be written as [3]

$$\boldsymbol{y}_k^n = \sum_{m=1}^{M} \begin{bmatrix} 1 \\ \boldsymbol{u}_m^n(k) \end{bmatrix} s_m(k) + \boldsymbol{\epsilon}_n(k), \quad (4)$$

where $\boldsymbol{\epsilon}_n(k) \in \mathbb{C}^{4\times 1}$ represents the channel noise including the pressure and the velocity noise terms. Note that in (4), the pressure and the velocity measurements are assumed to be proportional to the signal amplitude, and we have normalized the particle velocity terms by multiplying by a constant term $-\rho_0 c_0$, where ρ_0 and c_0 represent the ambient density and the propagation speed of the acoustic wave in the medium, respectively. The noise process $\boldsymbol{\epsilon}_k^n$ is a sequence of complex-valued i.i.d. circular Gaussian random variables with zero mean and covariance matrix $\boldsymbol{\Gamma}$, given as $\boldsymbol{\epsilon}_k^n \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma})$, where $\mathcal{CN}(\cdot|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for the circular complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Also, the noise processes at different AVSs are assumed to be independent of each other.

At each time step, a frame of signal that consists of a number of snapshots are employed to estimate the position. For the AVS data structure (4), each snapshot contains 4×1 samples. Assume that T_0 snapshots are employed at each time step. When T_0 is small, the source can be assumed to be stationary during each measurement frame. Eq. (4) can thus be written in a matrix form as

$$\mathbf{Y}_k^n = \mathbf{A}^n(\mathbf{X}_k)\mathbf{S}_k + \boldsymbol{\epsilon}_k^n, \quad (5)$$

where $\mathbf{Y}_k^n \in \mathbb{C}^{4\times T_0}$ and $\boldsymbol{\epsilon}_k^n \in \mathbb{C}^{4\times T_0}$, and

$$\mathbf{A}^n(\mathbf{X}_k) = [\mathbf{a}^n(\mathbf{x}_{1,k}), \ldots, \mathbf{a}^n(\mathbf{x}_{M,k})] \in \mathbb{C}^{4\times M}, \quad (6)$$

$$\mathbf{S}_k = [\mathbf{s}_{1,k}, \ldots, \mathbf{s}_{M,k}]^T \in \mathbb{C}^{M\times T_0}, \quad (7)$$

with $\mathbf{a}^n(\mathbf{x}_{m,k}) = \left[1, (\boldsymbol{u}_m^n(k))^T\right]^T \in \mathbb{C}^{4\times 1}$. In [13], a two step method has been developed for 3-D localization. The AVS signal model (4) is employed to obtain DOA measurements by using the Capon beamforming method. The DOAs are then regarded as measurements and employed to triangulate the 3-D source position by using a WLS approach or an RWLS approach. However, these approaches are limited to the single source scenario. Although the DOAs of multiple sources can be found by obtaining multiple peaks in the Capon beamforming response, achieving a 3-D position estimation is not a trivial task since a sophisticated data association technique is required to assign the DOAs to each source.

2.2. Centralized particle filtering

Eq. (5) shows a direct relationship between the measurements and the 3-D position states. Hence, it is possible to estimate the 3-D position of the source directly given by the signals collected from all sensors, i.e., $\mathbf{Y}_k = [(\mathbf{Y}_k^1)^T, \ldots, (\mathbf{Y}_k^N)^T]^T$. Given a signal sequence $\mathbf{Y}_{1:k} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_k]$ obtained until time step k, the state to be estimated is $\mathbf{X}_k$ (e.g., $\mathbf{X}_k = [\mathbf{x}_{1,k}^T, \ldots, \mathbf{x}_{M,k}^T]^T$). The posterior distribution of the state $p(\mathbf{X}_k|\mathbf{Y}_{1:k})$ can be obtained via a Bayesian recursive estimation, given as [48]

- Predict:

$$p(\mathbf{X}_k|\mathbf{Y}_{1:k-1}) = \int p(\mathbf{X}_k|\mathbf{X}_{k-1})p(\mathbf{X}_{k-1}|\mathbf{Y}_{1:k-1})\, d\mathbf{X}_{k-1}; \quad (8)$$

- Update:

$$p(\mathbf{X}_k|\mathbf{Y}_{1:k}) = \frac{p(\mathbf{Y}_k|\mathbf{X}_k)p(\mathbf{X}_k|\mathbf{Y}_{1:k-1})}{p(\mathbf{Y}_k|\mathbf{Y}_{1:k-1})}. \quad (9)$$

In this recursion, $p(\mathbf{X}_{k-1}|\mathbf{Y}_{1:k-1})$ is the posterior distribution estimated at the last time step, and $p(\mathbf{X}_k|\mathbf{X}_{k-1}) = p(\mathbf{X}_k|\mathbf{X}_{k-1}, \mathbf{Y}_{1:k-1})$ is the state transition probability density function (PDF) for the current time step. The Bayesian recursion states that given the transition PDF and the likelihood (which can usually be obtained according to the system models), the current state PDF can be estimated recursively from the state PDF at the previous time step $k-1$.

Although the Kalman filter can be used to solve the Bayesian recursion in (8) and (9), its use is limited to the case of linear and Gaussian system models. Since the measurement function is nonlinear, the PF [49] that provides an excellent solution to the nonlinear problem is employed. The core idea of the PF is that it uses a set of particles and importance weights of these particles to approximate the posterior distribution. Assuming that L particles are used to approximate the above Bayesian recursion, the PDF $p(\mathbf{X}_k|\mathbf{Y}_{1:k})$ is represented by $\{\mathbf{X}_k^{(\ell)}, w_k^{(\ell)}\}_{\ell=1}^{L}$. The entire procedure of PF processing can be summarized as follows. At time step k, the particles are sampled according to an importance function, given as

$$\mathbf{X}_k^{(\ell)} \sim q(\mathbf{X}_k^{(\ell)}|\mathbf{X}_{k-1}^{(\ell)}, \mathbf{Y}_{1:k}). \quad (10)$$

The importance weights of the particles are then evaluated by

$$w_k^{(\ell)} = w_{k-1}^{(\ell)} \frac{p(\mathbf{Y}_k|\mathbf{X}_k^{(\ell)})p(\mathbf{X}_k^{(\ell)}|\mathbf{X}_{k-1}^{(\ell)})}{q(\mathbf{X}_k^{(\ell)}|\mathbf{X}_{k-1}^{(\ell)}, \mathbf{Y}_{1:k})}. \quad (11)$$

Usually, the optimal importance function, i.e., $q(\mathbf{X}_k^{(\ell)}|\mathbf{X}_{k-1}^{(\ell)}, \mathbf{Y}_{1:k}) = p(\mathbf{X}_k^{(\ell)}|\mathbf{X}_{k-1}^{(\ell)}, \mathbf{Y}_{1:k})$ is able to provide the minimum estimation variance. After the resampling scheme, the posterior distribution of the state is thus approximated by

$$p(\mathbf{X}_k|\mathbf{Y}_{1:k}) \approx \sum_{\ell=1}^{L} \tilde{w}_k^{(\ell)} \delta_{\mathbf{X}_k^{(\ell)}}(\mathbf{X}_k), \tag{12}$$

where $\delta(\cdot)$ is a Dirac-delta function, and $\tilde{w}_k^{(\ell)}$ is a normalized weight. The centralized PF approach has been developed for AVS based acoustic source tracking in [50,31,38]. It has been shown that PF tracking approach is able to provide better accuracy of source position estimation than that provided by the WLS and RWLS based localization approaches [38].

2.3. Distributed AVS network

In the centralized PF approach, all collected signals are needed to be transmitted to the central processor. This is cumbersome since both the communication and the computation cost of the system can be very high. In this paper, we consider a sensor network comprising of N nodes. Following the AVS signal model (4) in Section 2.1, each node makes a measurement $\mathbf{Y}_k^n$ at discrete time instants k. The global measurement vector is given by

$$\begin{bmatrix} \mathbf{Y}_k^1 \\ \vdots \\ \mathbf{Y}_k^N \end{bmatrix} = \begin{bmatrix} \mathbf{A}^1(\mathbf{X}_k) \\ \vdots \\ \mathbf{A}^N(\mathbf{X}_k) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_k^1 \\ \vdots \\ \boldsymbol{\epsilon}_k^N \end{bmatrix}, \tag{13}$$

where $\boldsymbol{\epsilon}_k^n$ is the uncertainty in the local measurement models. The nodes of the network are modeled as vertices of the graph $\mathcal{G} = (\boldsymbol{\nu}, \mathcal{E})$, namely as the elements of the node set $\boldsymbol{\nu} = \{1, \dots, N\}$. The edge set $\mathcal{E} \subseteq \boldsymbol{\nu} \times \boldsymbol{\nu}$ represents the network's communication constraints, i.e., if node n can send information to node u then $(n,u) \in \mathcal{E}$. For graph $\mathcal{G}$, the maximum degree $\Delta_{\mathcal{G}} = \max_n D^{(n)}$, where $D^{(n)}$ is the number of neighbor nodes for the node n.

Assume that the local observations made at different sensor nodes conditioned on the state variables are independent among each other. According to the Bayesian update equation (9), the posterior distribution of the global state estimates $\mathbf{X}_{1:k}$ can be written as

$$p(\mathbf{X}_{1:k}|\mathbf{Y}_{1:k}) = \frac{\prod_{n=1}^{N} p(\mathbf{Y}_k^n|\mathbf{X}_k)}{p(\mathbf{Y}_k|\mathbf{Y}_{1:k-1})} p(\mathbf{X}_{1:k}|\mathbf{Y}_{1:k-1}). \tag{14}$$

where $p(\mathbf{X}_{1:k}|\mathbf{Y}_{1:k-1}) = p(\mathbf{X}_k|\mathbf{X}_{k-1})p(\mathbf{X}_{1:k-1}|\mathbf{Y}_{1:k-1})$. Further, we assume that the statistics or the state can be estimated at the local nodes, i.e., the posterior distribution is available at each local node. According to Bayes' rule, the likelihood at each node $p(\mathbf{Y}_k^n|\mathbf{X}_k)$ can thus be obtained as

$$\begin{aligned} p(\mathbf{Y}_k^n|\mathbf{X}_k) &= p\left(\mathbf{Y}_k^n|\mathbf{X}_k, \mathbf{Y}_{1:k-1}^n\right) \\ &= \frac{p(\mathbf{X}_k|\mathbf{Y}_{1:k}^n)}{p(\mathbf{X}_k|\mathbf{Y}_{1:k-1}^n)} p\left(\mathbf{Y}_k^n|\mathbf{Y}_{1:k-1}^n\right). \end{aligned} \tag{15}$$

Hence, the posterior distribution of global estimates can be written as

$$\begin{aligned} p(\mathbf{X}_{1:k}|\mathbf{Y}_{1:k}) = &\frac{\prod_{n=1}^{N} p(\mathbf{X}_k|\mathbf{Y}_{1:k}^n)}{\prod_{n=1}^{N} p(\mathbf{X}_k|\mathbf{Y}_{1:k-1}^n)} \\ &\times p(\mathbf{X}_k|\mathbf{X}_{k-1}) p(\mathbf{X}_{1:k-1}|\mathbf{Y}_{1:k-1}). \end{aligned} \tag{16}$$

The distributed estimation approach differs from the centralized approach in that the local posteriors of the states are estimated (distributively) at each local node. These posterior distributions are then fused together to obtain a global estimation.

It is worth mentioning that since dynamic sources are considered, the source is usually assumed to be at a position $\mathbf{X}_k^{\mathrm{p}}$ and moving with a velocity $\mathbf{X}_k^{\mathrm{v}}$. The new state $\mathbf{X}_k$ is thus constructed by cascading the position part with the velocity part, given as

$$\mathbf{X}_k = [(\mathbf{X}_k^{\mathrm{p}})^T, (\mathbf{X}_k^{\mathrm{v}})^T]^T. \tag{17}$$

The constant velocity (CV) model [51,29,30] is employed here to model the source dynamics, given as

$$\mathbf{X}_k = \mathbf{F}\mathbf{X}_{k-1} + \mathbf{G}\mathbf{V}_k. \tag{18}$$

The coefficient matrix $\mathbf{F}$ and $\mathbf{G}$ are defined respectively as

$$\mathbf{F} = \mathbf{I}_M \otimes \begin{bmatrix} \mathbf{I}_3 & \Delta T \mathbf{I}_3 \\ \mathbf{0} & \mathbf{I}_3 \end{bmatrix}; \quad \mathbf{G} = \mathbf{I}_M \otimes \begin{bmatrix} \frac{\Delta T^2}{2} \mathbf{I}_3 \\ \Delta T \mathbf{I}_3 \end{bmatrix}, \tag{19}$$

where $\mathbf{I}_q$ denotes the qth-order identity matrix, and ΔT represents the time period in seconds between the previous and the current time steps, and $\otimes$ denotes the Kronecker product. $\mathbf{V}_k$ is a zero-mean Gaussian process that describes the turbulence in modeling the source velocity and assumed to have a variance matrix $\boldsymbol{\Sigma}_{\mathrm{v}}$. In the next section, a distributed unscented particle filter approach is developed to estimate the positions of multiple sources.

3. Distributed tracking algorithm for AVS network

In this section, a distributed unscented particle filtering (DUPF) approach is proposed for 3-D source position tracking. First, at each AVS node, an unscented information filter (UIF) based PF is implemented to obtain a local state estimation. A consensus filter is then employed to fuse the local posteriors to derive the global estimates.

3.1. Unscented information filter at local nodes

Assume that at time step $k-1$, the global estimates of the source state $\widehat{\mathbf{X}}_{k-1}$ and the covariance $\widehat{\mathbf{P}}_{k-1}$ are available. At each AVS node n, rather than updating $\widehat{\mathbf{X}}_{k-1}$ and $\widehat{\mathbf{P}}_{k-1}$ based on the local measurements, the UIF updates the Fisher information matrix $\tilde{\mathcal{I}}_{k-1}^n$ and the information state vector $\tilde{\mathcal{Z}}_{k-1}^n$. These two information terms are defined as

$$\tilde{\mathcal{I}}_{k-1}^n \triangleq (\widehat{\mathbf{P}}_{k-1}^n)^{-1}; \tag{20}$$

$$\tilde{\mathcal{Z}}_{k-1}^n \triangleq \tilde{\mathcal{I}}_{k-1}^n \widehat{\mathbf{X}}_{k-1}^n. \tag{21}$$

The core steps of the UIF can be summarized as follows [42]:

1. Calculate a set of $(2n_x+1)$ deterministic samples (sigma points) $\mathcal{S} = \{\mathcal{W}^j, \mathcal{X}_{k-1}^j\}_{j=0}^{2n_x}$ where n_x denotes the number of state variables, based on the following scheme:

   $$\mathcal{X}_{k-1}^{nj} = \widehat{\mathbf{X}}_{k-1}^n \pm \left\{ \sqrt{(n_x+\kappa)\widehat{\mathbf{P}}_{k-1}^n} \right\}^j, \tag{22}$$

   where term $\left\{ \sqrt{(n_x+\kappa)\widehat{\mathbf{P}}_{k-1}^n} \right\}^j$ corresponds to the jth column of the square root of matrix $(n_x+\kappa)\widehat{\mathbf{P}}_{k-1}^n$ and the initial condition is given by $\mathcal{X}_k^{n,0} = \widehat{\mathbf{X}}_{k-1}^n$. The corresponding weights for the sigma points $\{\mathcal{W}^j\}_{j=1}^{2n_x}$ are given by $\mathcal{W}^j = 1/(2(n_x+\kappa))$, where κ is a scaling parameter and the initial condition for the sigma points is $\mathcal{W}^0 = \kappa/(n_x+\kappa)$.

2. Generate the predicted sigma points

$$\mathcal{X}_{k|k-1}^{n,j} = \mathbf{F}\mathcal{X}_{k-1}^{n,j} + \mathbf{G}\mathbf{v}_k, \quad \text{for } j = 0, \ldots, 2n_x. \tag{23}$$

3. Propagate $\mathcal{X}_{k|k-1}^{n,j}$ through the measurement function (5) to generate the predicted observation sigma points

$$\mathcal{Y}_{k|k-1}^{n,j} = \mathbf{A}^n(\mathcal{X}_{k|k-1}^{n,j})\overline{\mathbf{S}}_k^n, \quad \text{for } i = 0, \ldots, 2n_x, \tag{24}$$

where $\overline{\mathbf{S}}_k$ is a maximum likelihood estimation of the source signal, given as

$$\overline{\mathbf{S}}_k^n = \left(\left(\mathbf{A}^n(\mathcal{X}_{k|k-1}^{n,j})\right)^H \mathbf{A}^n(\mathcal{X}_{k|k-1}^{n,j})\right)^{-1} \times \left(\mathbf{A}^n(\mathcal{X}_{k|k-1}^{n,j})\right)^H \mathbf{Y}_k^n. \tag{25}$$

4. Estimate the predicted source state $\overline{\mathbf{X}}_{k|k-1}^n$, the error covariance matrix $\overline{\mathbf{P}}_{k|k-1}^n$, and the predicted measurement sequence $\mathcal{Y}_{k|k-1}^n$ as follows:

$$\overline{\mathbf{X}}_{k|k-1}^n = \sum_{j=0}^{2n_x} \mathcal{W}^j \mathcal{X}_{k|k-1}^{n,j}; \tag{26}$$

$$\overline{\mathbf{P}}_{k|k-1}^n = \sum_{j=0}^{2n_x} \mathcal{W}^j (\mathcal{X}_{k|k-1}^{n,j} - \overline{\mathbf{X}}_{k|k-1}^n) \times (\mathcal{X}_{k|k-1}^{n,j} - \overline{\mathbf{X}}_{k|k-1}^n)^T; \tag{27}$$

$$\mathcal{Y}_{k|k-1}^n = \sum_{j=0}^{2n_x} \mathcal{W}^j \mathcal{Y}_{k|k-1}^{n,j}. \tag{28}$$

5. The cross-covariance $\overline{\mathbf{P}}_{k|k-1}^{n,\mathcal{X}\mathcal{Y}}$ between the predicted measurement and the predicted state estimates is computed as

$$\overline{\mathbf{P}}_{k|k-1}^{n,\mathcal{X}\mathcal{Y}} = \sum_{j=0}^{2n_x} \mathcal{W}^j (\mathcal{X}_{k|k-1}^{n,j} - \overline{\mathbf{X}}_{k|k-1}^n) \times (\mathcal{Y}_{k|k-1}^{n,j} - \widehat{\mathbf{Y}}_k^n)^T. \tag{29}$$

6. The predicted information matrix and the state vector are

$$\overline{\mathcal{I}}_{k|k-1}^n = \left(\overline{\mathbf{P}}_{k|k-1}^n\right)^{-1}, \tag{30}$$

$$\overline{\mathcal{Z}}_{k|k-1}^n = \overline{\mathcal{I}}_{k|k-1}^n \overline{\mathbf{X}}_{k|k-1}^n. \tag{31}$$

7. The information contribution equations are given as

$$\mathcal{K}_k^n = (\mathbf{H}_k^n)^T (\mathbf{R}_k^n)^{-1} \left(\Xi_k^n + \mathbf{H}_k^n \overline{\mathbf{X}}_{k|k-1}^n\right), \tag{32}$$

$$\mathcal{G}_k^n = (\mathbf{H}_k^n)^T (\mathbf{R}_k^n)^{-1} \mathbf{H}_k^n, \tag{33}$$

where

$$(\mathbf{H}_k^n)^T = \left(\overline{\mathbf{P}}_{k|k-1}^n\right)^{-1} \overline{\mathbf{P}}_{k|k-1}^{n,\mathcal{X}\mathcal{Y}}, \tag{34}$$

$$\Xi_k^n = \mathbf{Y}_k^n - \mathbf{A}^n(\overline{\mathbf{X}}_{k|k-1}^n)\overline{\mathbf{S}}_k^n. \tag{35}$$

Here $\overline{\mathbf{S}}_k^n$ is obtained according to (25) by replacing $\mathcal{X}_{k|k-1}^{n,j}$ with $\overline{\mathbf{X}}_{k|k-1}^n$.

8. Update the Fisher information matrix and the information state vector

$$\tilde{\mathcal{I}}_k^n = \mathcal{G}_k^n + \overline{\mathcal{I}}_{k|k-1}^n, \tag{36}$$

$$\tilde{\mathcal{Z}}_k^n = \mathcal{K}_k^n + \overline{\mathcal{Z}}_{k|k-1}^n. \tag{37}$$

9. Finally, the statistics estimated at the nth local node can be obtained as

$$\tilde{\mathbf{P}}_k^n = \left(\tilde{\mathcal{I}}_k^n\right)^{-1}, \tag{38}$$

$$\tilde{\mathbf{X}}_k^n = \tilde{\mathbf{P}}_k^n \tilde{\mathcal{Z}}_k^n. \tag{39}$$

The UIF can only provide a coarse estimation of the posterior. A PF is then applied to obtain a more accurate state estimation. After implementing the UIF, the optimum importance function at the local node can be approximated by

$$q(\mathbf{X}_k^n | \mathbf{X}_{k-1}^n, \mathbf{Y}_k^n) \approx \mathcal{N}(\tilde{\mathbf{X}}_k^n, \tilde{\mathbf{P}}_k^n). \tag{40}$$

The particles are thus drawn according to (40), given as

$$\mathbf{X}_k^{n,(\ell)} \sim q(\mathbf{X}_k^n | \mathbf{X}_{k-1}^n, \mathbf{Y}_k^n). \tag{41}$$

The importance weight is calculated as

$$w_k^{n,(\ell)} = w_{k-1}^{n,(\ell)} \frac{p(\mathbf{Y}_k^n | \mathbf{X}_k^{n,(\ell)}) p(\mathbf{X}_k^{n,(\ell)} | \mathbf{X}_{k-1}^{n,(\ell)})}{q(\mathbf{X}_k^{n,(\ell)} | \mathbf{X}_{k-1}^{n,(\ell)}, \mathbf{Y}_k^n)}. \tag{42}$$

The detailed expression of transition density $p(\mathbf{X}_k^{n,(\ell)} | \mathbf{X}_{k-1}^{n,(\ell)})$ and likelihood $p(\mathbf{Y}_k^n | \mathbf{X}_k^{n,(\ell)})$ in (42) will be given in Section 3.3. After resampling, the local statistics are estimated by

$$\overline{\mathbf{X}}_k^n = \sum_{\ell=1}^{L} w_k^{n,(\ell)} \mathbf{X}_k^{n,(\ell)}; \tag{43}$$

$$\overline{\mathbf{P}}_k^n = \sum_{\ell=1}^{L} w_k^{n,(\ell)} (\mathbf{X}_k^{n,(\ell)} - \overline{\mathbf{X}}_k^n)(\mathbf{X}_k^{n,(\ell)} - \overline{\mathbf{X}}_k^n)^T. \tag{44}$$

In the next section, a consensus filter will be introduced to fuse these local statistics to obtain a global estimation of the source state and the covariance matrix.

3.2. *Consensus filter for global estimation*

In the distributed estimation, the local particles and their associated weights are based only on the local observations $\mathbf{Y}_k^n$. This results in inconsistent state estimates $\mathbb{E}(\mathbf{X}_k | \mathbf{Y}_k^n)$ across the network. A consensus step is introduced to provide the consistency of the local estimates. Based on the model described in Section 2.1, we model the local posterior with a Gaussian distribution and develop a consensus-based distributed counterpart of the optimal decentralized fusion rule. We go one step further and use such Gaussian approximations in the context of local UIF as the proposal distribution. At time instant $k-1$, all nodes are assumed to have reached a consensus with a single set of global estimates $\widehat{\mathbf{X}}_{k-1}$ and $\widehat{\mathbf{P}}_{k-1}$ across the network. A new measurement $\mathbf{Y}_k^n$ is now available at the local node n, for $1 \le n \le N$. Based on Eqs. (43) and (44), the node n computes the state $\overline{\mathbf{X}}_k^n$ and its corresponding error covariance $\overline{\mathbf{P}}_k^n$ (local statistics).

The second step is the fusion step used to compute a consistent set of values for the global statistics $\widehat{\mathbf{X}}_k$ and $\widehat{\mathbf{P}}_k$ at time k from the local statistics. To combine the local statistics $\{\overline{\mathbf{X}}_k^n, \overline{\mathbf{P}}_k^n\}_{n=1}^N$ into a common set of global statistics, $\widehat{\mathbf{X}}_k$ and $\widehat{\mathbf{P}}_k$ across the network, based on the Chong–Mori–Chang track-fusion theorem [52], the following rules are derived:

$$\left[\widehat{\mathbf{P}}_k\right]^{-1} = \mathcal{I}_{k|k-1}^n + \underbrace{\sum_{n=1}^N \left\{\left[\overline{\mathbf{P}}_k^n\right]^{-1} - \mathcal{I}_{k|k-1}^n\right\}}_{\mathbf{P}_c(\infty)}; \tag{45}$$

$$\widehat{\mathbf{X}}_k = \left[\widehat{\mathbf{P}}_k\right]^{-1}\left[\mathcal{Z}_{k|k-1}^n + \underbrace{\sum_{n=1}^N \left\{\left[\overline{\mathbf{P}}_k^n\right]^{-1}\overline{\mathbf{X}}_k^n - \mathcal{Z}_{k|k-1}^n\right\}}_{\mathbf{x}_c(\infty)}\right]. \tag{46}$$

In Eqs. (45) and (46), $\{\mathbf{x}_c(\infty), \mathbf{P}_c(\infty)\}$ are obtained by iterating the following average consensus equations:

$$\mathbf{X}_c^n(t+1) = \mathbf{X}_c^n(t) + \epsilon \sum_{j \in \aleph^{(n)}} (\mathbf{X}_c^j(t) - \mathbf{X}_c^n(t)); \tag{47}$$

$$\mathbf{P}_c^n(t+1) = \mathbf{P}_c^n(t) + \epsilon \sum_{j \in \aleph^{(n)}} (\mathbf{P}_c^j(t) - \mathbf{P}_c^n(t)), \tag{48}$$

until they converge to $\{\mathbf{X}_c(\infty), \mathbf{P}_c(\infty)\}$, where $\epsilon \in (0, 1/\Delta_{\mathcal{G}})$. The initial conditions are

$$\mathbf{P}_c^n(t=0) = \left[\overline{\mathbf{P}}_k^n\right]^{-1} - \mathcal{I}_{k|k-1}^n; \tag{49}$$

$$\mathbf{X}_c^n(t=0) = \left[\overline{\mathbf{P}}_k^n\right]^{-1}\overline{\mathbf{X}}_k^n - \mathcal{Z}_{k|k-1}^n. \tag{50}$$

Note that the consensus approach in Eqs. (47) and (48) is a distributed algorithm where each node communicates only with its neighbor nodes. Convergence of the consensus algorithms has been extensively studied [53], e.g., it has been shown [53] that achieving consensus in a finite number of iterations is possible even for time-invariant topologies. In this paper, we consider the case where it is possible to communicate sufficiently fast so that consensus is reached between two successive observations. This is a common practice in consensus-based distributed implementation of the PF. See [46] for extension to the scenario where the consensus is not reached within two consecutive time iterations.

3.3. Tracking algorithm

The transition density $p(\mathbf{X}_k^n|\mathbf{X}_{k-1}^n)$ and the likelihood $p(\mathbf{Y}_k^n|\mathbf{X}_k^n)$ in (42) can be obtained from the state dynamic model (18) and the measurement model (4), respectively. Note that the particle indices are dropped off without any ambiguity in this section. Also, ignoring the node indices, the transition PDF can be written as

$$\begin{aligned} p(\mathbf{X}_k|\mathbf{X}_{k-1}) &= p(\mathbf{X}_k^p, \mathbf{X}_k^v|\mathbf{X}_{k-1}^p, \mathbf{X}_{k-1}^v) \\ &= p(\mathbf{X}_k^p|\mathbf{X}_{k-1}^p, \mathbf{X}_{k-1}^v)\,p(\mathbf{X}_k^v|\mathbf{X}_{k-1}^v), \end{aligned} \tag{51}$$

where $p(\mathbf{X}_k^p|\mathbf{X}_{k-1}^p, \mathbf{X}_{k-1}^v)$ and $p(\mathbf{X}_k^v|\mathbf{X}_{k-1}^v)$ can be, respectively, given by

$$\begin{aligned} &p(\mathbf{X}_k^p|\mathbf{X}_{k-1}^p, \mathbf{X}_{k-1}^v) \\ &= \mathcal{N}(\mathbf{X}_k^p; \mathbf{X}_{k-1}^p + \Delta T \mathbf{X}_{k-1}^v, \Delta T^4/4\mathbf{Q}); \end{aligned} \tag{52}$$

$$p(\mathbf{X}_k^v|\mathbf{X}_{k-1}^v) = \mathcal{N}(\mathbf{X}_k^v; \mathbf{X}_{k-1}^v, \Delta T^2\mathbf{Q}). \tag{53}$$

According to [3], the concentrated likelihood function can be written as

$$\begin{aligned} &p(\mathbf{Y}_k^n|\mathbf{X}_k) = (\pi^{-4T_0})\exp(-4T_0) \\ &\quad \det(\mathbf{\Pi}_k^n \widehat{\mathbf{\Lambda}}_k^n \mathbf{\Pi}_k^n + \widehat{\sigma}_n^2 \mathbf{\Pi}_k^{n,0})^{-T_0}, \end{aligned} \tag{54}$$

where

$$\mathbf{\Pi}_k^n = \mathbf{A}_n(\mathbf{X}_k^p)(\mathbf{A}_n^H(\mathbf{X}_k^p)\mathbf{A}_n(\mathbf{X}_k^p))^{-1}\mathbf{A}_n^H(\mathbf{X}_k^p); \tag{55}$$

$$\mathbf{\Pi}_k^{n,0} = \mathbf{I} - \mathbf{\Pi}_k^n; \tag{56}$$

$$\widehat{\sigma}_n^2 = \frac{1}{4-M}\mathrm{tr}\left(\mathbf{\Pi}_k^{n,0}\widehat{\mathbf{\Lambda}}_k^n\right); \tag{57}$$

$$\widehat{\mathbf{\Lambda}}_k^n = \frac{1}{T_0}\mathbf{Y}_k^n(\mathbf{Y}_k^n)^H, \tag{58}$$

where the superscript H denotes the conjugate transpose. The transition density $p(\mathbf{X}_k^n|\mathbf{X}_{k-1}^n)$ and the likelihood $p(\mathbf{Y}_k^n|\mathbf{X}_k^n)$ at the nth local node can be derived by replacing their counterparts in (51) and (54), respectively with $\mathbf{X}_k^n$ and $\mathbf{X}_{k-1}^n$.

For the first time step in implementing the distributed tracking algorithm, the positions are initialized by using a maximum likelihood estimation of the total likelihood given as

$$\widehat{\mathbf{X}}_0^p = \arg\max_{\mathbf{X}} \prod_{n=1}^N p(\mathbf{Y}_k^n|\mathbf{X}_k). \tag{59}$$

The velocity part of the state is initialized around the ground truth velocity and an initial covariance $\mathbf{\Sigma}_0$ is given to describe the errors in the position and velocity state estimation.

3.4. Analysis of the communication and computational cost

The complete tracking algorithm is summarized in Algorithm 1. The centralized particle filtering approach is an optimal one since it makes full use of correlations between the signals received at different AVSs. However, it requires the transmission of all raw data to a central processor either directly or indirectly via multi-hop relay. Such a centralized scheme is not practical in terms of communication cost since for wideband acoustic signal, a frame of signal is processed at each time step. For example, assume that the frame length is T_0. That means $4T_0$ signal samples have to be transmitted from each AVS at each time step. Consider an AVS network with N sensors. The central processor has to receive $4NT_0$ samples at each iteration of the filter. Such a high data rate is not achievable in applications with low or intermittent bandwidth connectivity. Also, synchronization among different AVSs is not a trivial task. The proposed approach needs only to transmit the first- and the second-order statistics among connected neighboring sensors rather than transmitting all raw measurements to the CP. The number of real values needed to be transmitted at each local node can be expressed as $(\kappa_x+\kappa_x^2)IM$, where I is the number of consensus iterations, M is the number of sources, and κ_x is the dimension of the state for each source. In this paper, since the state is the 3-D position and the corresponding velocity, we have $\kappa_x=6$. In contrast to the CPF, the data that needs to be transmitted does not depend on the dimensions of the measurements. Hence, the proposed DUPF approach is particularly favorable for the case of high-dimensional measurements.

Algorithm 1. DUPF tracking algorithm.

Initialization: estimate $\mathbf{X}_0^{\mathrm{p}}$ according to (59) and let $\mathbf{\Sigma}_0 = \mathbf{\Sigma}_v = 0.01\mathbf{I}_{3M}$;
for $\ell = 1, \ldots, L$, draw particles $\mathbf{X}_0^{(i)} \sim \mathcal{N}(\widehat{\mathbf{X}}_0, \mathbf{\Sigma}_0)$;
set the initial weight $\widetilde{w}_0^{(\ell)} = 1/L$;
Compute the inverse of covariance matrix $\mathbf{R}_k^n$ and $\mathbf{Q}_k^{\mathrm{p}}$.
for $k \leftarrow 1$ **to** K **do**
  for $n \leftarrow 1$ **to** N **do**
    1) implement UIF according to step 1) to step 9) in Section 3.1;
    for $\ell \leftarrow 1$ **to** L **do**
      2) draw samples according to (41);
      3) compute the transition and likelihood densities according to (51) and (54) respectively;
      4) compute the importance weight according to (42);
    end
    5) normalise the weight $\widetilde{w}_k^{(\ell)} = w_k^{(\ell)} / \sum_{\ell=1}^{L} w_k^{(\ell)}$;
    6) resample the particles according to the weights;
    7) calculate the local statistics according to (43) and (44);
  end
  8) estimate the global statistics using consensus filter.
end

each time step is aligned as a vector with dimensions of $4T_0 \times 1$, i.e., $\widehat{\mathbf{Y}}_k^n = \mathrm{Vec}(\mathbf{Y}_k^n)$ with $\mathrm{Vec}(\cdot)$ denoting the vector transformation. Consequently, the dimensionality of the covariance matrix associated with measurement process is large. Directly applying the UKF will result in high computational complexity since an inverse operation of a large covariance matrix is involved in the update equations. The UIF introduced in this work can be regarded as a special member of the UKF family that propagates the inverse of the state variance matrix rather than the state variance matrix itself. While it requiring the inverse of the state variance only, the UKF needs to calculate the inverse of the measurement variance matrix with a dimension of $4T_0 \times 4T_0$. Hence, the size of the matrix to be inverted can be restricted to $6M \times 6M$ where M is the number of

The analysis of the computational complexity in terms of consensus filter can be found in the literature [53]. Usually, the consensus for global statistics can be reached after a small number of iterations, for which the computation cost is trivial and can be ignored. The main computation cost arises from the filtering process at each local node. In the PF algorithm, the unscented Kalman filter (UKF) is often employed to approximate the optimum sampling. When T_0 snapshots are used to estimate the source position, the measurement at each node and at

sources. Usually, $6M \ll 4T_0$. Compared to the extended Kalman filter, calculation of the Hessian matrix and the nonlinear approximation error due to higher order truncation can, therefore, be avoided.

In the state-space model, the noise variances are usually assumed to be constants, e.g., noise process in source dynamics $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{\Sigma})$ and in measurement $\epsilon_k^n \sim \mathcal{N}(0, \mathbf{\Gamma})$. Hence, $(\mathbf{R}_k^n)^{-1} = \mathbf{\Gamma}^{-1}$ can be calculated at the beginning of the implementation. It is not necessary to calculate $\mathbf{R}_k^n$ at each iteration. Note that, the equations

Link to Full-Text Articles :

http://www.sciencedirect.com/science/article/pii/S016516841400454X