An Overview of Sequential Bayesian Filtering in Ocean Acoustics

Caglar Yardim, *Member, IEEE,* Zoi-Heleni Michalopoulou, *Senior Member, IEEE,* and Peter Gerstoft
To Appear, IEEE JOE, Jan 2011, DOI:10.1109/JOE.2010.2098810

***Abstract*— Sequential filtering provides a suitable framework for estimating and updating the unknown parameters of a system as data become available. The foundations of sequential Bayesian filtering with emphasis on practical issues are first reviewed covering both Kalman and particle filter approaches. Filtering is demonstrated to be a powerful estimation tool, employing prediction from previous estimates and updates stemming from physical and statistical models that relate acoustic measurements to the unknown parameters. Ocean acoustic applications are then reviewed focusing on source tracking, estimation of environmental parameters evolving in time or space, and frequency tracking. Spatial arrival time tracking is illustrated with Shallow Water 06 data.**

***Index Terms*— Sequential Monte Carlo methods, ocean acoustics, acoustic signal processing, acoustic tracking, extended Kalman filter, unscented Kalman filter, ensemble Kalman filter, particle filter, sequential importance resampling.**

I. Introduction

A common feature of inverse problems in ocean acoustics is that underlying physical parameters are estimated from measured acoustic data. Examples include source localization [1]–[4], geoacoustic inversion [5]–[9], and marine mammal signal processing [10]. In a Bayesian framework, prior knowledge and acoustic models are combined with a likelihood function to provide posterior probability density functions (PDFs) of parameters-of-interest. This formulation was first proposed in source localization [11], [12]. Geoacoustic inversion was subsequently approached in a similar fashion, estimating, in addition to source location, ocean environment parameters and their uncertainties [13]–[15]. Often, such parameters evolve in time or space, with acoustic data arriving on-line at consecutive steps. Information on parameter value evolution and uncertainty at preceding steps can be invaluable for the determination of future estimates but is frequently ignored.

Depending on the data and the problem, Bayesian approaches can be used to address inversion or tracking problems (a brief comparison is given in Appendix I). Sequential Bayesian filtering, tying together information on parameter evolution, a function relating acoustic field measurements to the unknown quantities, and a statistical description of the random perturbations in the field measurements, offers a suitable framework for the solution of such problems.

When functions relating (i) parameters between consecutive steps and (ii) data and parameters are linear and noise is additive and Gaussian, the Kalman filter (KF) is the optimal estimator in terms of minimizing mean squared error. The KF propagates expectations and covariances of the unknown parameters from step to step, fully characterizing posterior PDFs. Some early examples used in source tracking in the ocean with KFs are given in [16], [17]. Nonlinear functions require variations or generalization of the standard filter. Implementation in ocean acoustic problems of a straightforward extension, the extended Kalman filter (EKF), that linearizes mildly nonlinear functions, has been pioneered in a series of papers [18]–[27]. More recently, unscented Kalman filters (UKFs) have achieved better root-mean-square (RMS) error and convergence performance than EKFs by selecting deterministic points called sigma points to represent parameter PDFs [28], [29]. For large parameter vectors the Ensemble KF (EnKF) is efficient [30]. Highly nonlinear systems and complex noise processes require numerical methods for the computation of posterior PDFs. These approaches are termed particle filters (PFs) and have been encountered in ocean acoustic applications in [29], [31]–[33].

In this tutorial paper we briefly review the foundations of sequential filters, starting from the well-known KF and its variants and proceeding with PFs. These techniques formulate sequential estimation using physical relations between unknown parameters and measurements embedded in noise environments of a diverse nature. We examine and compare filters and present examples, illustrating practical challenges and solutions. A more detailed presentation of these methods, including theoretical derivations, can be found in many excellent signal processing papers [34]–[39] and texts on sequential Monte Carlo methods [40]–[43].

Once basic principles and methods of sequential filtering are discussed, the focus is shifted to an overview of sequential filtering implemented for parameter estimation of dynamical systems in ocean acoustics. We present applications in target tracking, wave estimation, geoacoustic inversion, frequency, and arrival time tracking.

The paper is organized as follows: First a general background about the state-space formulation for the estimation of evolving parameters in dynamical systems is given in Section II together with the basics of Bayesian filtering. The KF framework is introduced in Sections III together with its extensions such as the EKF, UKF, and the ensemble Kalman (EnKF)

Manuscript received ? ?, 2010. This work was supported by the Office of Naval Research Code 32, under grants N00014-09-1-0313, N00015-05-1-0264, N00014-05-1-0262, and N00014-10-1-0073.

Caglar Yardim and Peter Gerstoft are with the Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093–0238 USA (email: cyardim@ucsd.edu, gerstoft@ucsd.edu).

Zoi-Heleni Michalopoulou is with the Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, New Jersey 071028 USA (email: michalop@njit.edu).

filters. The PFs are presented in Section IV. The filter equations are derived starting from the basic importance sampling (IS) concepts, moving to sequential IS, finally deriving the commonly used PF, often referred to as sequential importance resampling (SIR). Advanced PFs are briefly mentioned in Section IV and are presented more extensively in Appendix II. Practical challenges arising in sequential Bayesian filtering are discussed in Section V and ocean acoustic applications are discussed in Section VI. Although Section VI is dedicated to solving estimation problems in the ocean, specific issues in ocean acoustics as they relate to filtering are discussed throughout the paper.

II. Background

A. State-space model

In dynamical systems, a major goal is to estimate parameters that evolve sequentially with time or space. As data become available, the unknown parameters forming a state vector, are estimated sequentially using collective data history and prior knowledge on evolution of the state. Tracking in time or space presents us with problems that can be readily formalized in the following framework.

Let $\mathbf{y}_k$ be the measurement vector (for example, pressure along a hydrophone array in ocean acoustics) at step k and $\mathbf{x}_k$ represent the state vector (for example, location of a source emitting signal $\mathbf{y}_k$), where $k = 1, \ldots, K$. The size of the state vector is n_x. Our goal is to sequentially estimate $\mathbf{x}_k$ as data measurements $\mathbf{y}_k$ become available.

Two equations define a state-space model:

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_k) \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{w}_k) \tag{2}$$

The *state equation*, Eq. (1), describes the evolution or transition of $\mathbf{x}_k$ with k and assumes that states follow a first order Markov process. Function $\mathbf{f}_k$ is a known function relating the state vector at step k to that at step $k-1$. Variable $\mathbf{v}_k$ is the process or state noise and has a known PDF $p(\mathbf{v}_k)$.

The *measurement equation* (or observation equation), Eq. (2), relates measurements $\mathbf{y}_k$ to state vector $\mathbf{x}_k$ through a known function $\mathbf{h}_k$. Variable $\mathbf{w}_k$ is the measurement noise with a PDF $p(\mathbf{w}_k)$.

Noise $\mathbf{v}_k$ and $\mathbf{w}_k$ can be additive, multiplicative, or incorporated in the state and measurement through complex functions of $\mathbf{f}_k$ and $\mathbf{h}_k$, respectively. In addition to the state vector $\mathbf{x}_k$ and data $\mathbf{y}_k$, the functions $\mathbf{f}_k$ and $\mathbf{h}_k$, and noise components $\mathbf{v}_k$ and $\mathbf{w}_k$ can also change with k.

In most ocean acoustic applications, the noise terms are additive and functions $\mathbf{f}_k$ and $\mathbf{h}_k$ is independent of k. Hence, Eqs. (1–2) are simplified to:

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{v}_k \tag{3}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{w}_k \ . \tag{4}$$

In a typical ocean acoustic example, function $\mathbf{h}_k$ represents some form of a forward or acoustic propagation model and vector $\mathbf{y}_k$ consists of complex-valued acoustic data measurements across a hydrophone array.

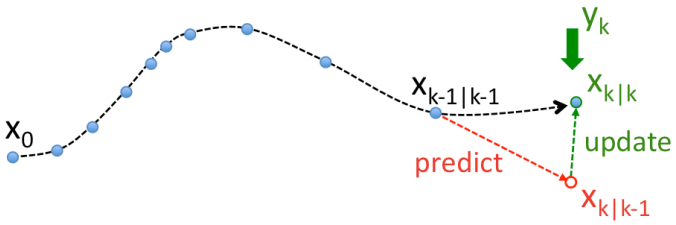

Fig. 1. Sequential Bayesian filtering. From state $\mathbf{x}_{k-1}$, state $\mathbf{x}_k$ is first predicted via the state equation [Eq. (1)], providing $\mathbf{x}_{k|k-1}$. As data $\mathbf{y}_k$ becomes available, measurement equation [Eq. (2)] is employed to update state x_k, providing $\mathbf{x}_{k|k}$.

A sketch of sequential localization of a moving target as described by Eqs. (1) and (2) is shown in Fig. 1 with circles representing target location at step k. Term $\mathbf{x}_{k-1|k-1}$ is the estimate of the target location and velocity at $k-1$ using data obtained up to and including step $k-1$. By exploiting knowledge of target position and trajectory $\mathbf{x}_{k-1|k-1}$, a filter first *predicts* $\mathbf{x}_{k|k-1}$, the target location and velocity at step k. The prediction is then *updated* when new data $\mathbf{y}_k$ become available.

B. Bayesian filtering

Examining the problem from a Bayesian standpoint, we are interested in deriving the posterior PDF for $\mathbf{x}_k$ rather than simple point estimates. The initial PDF, $p(\mathbf{x}_0)$, of the state vector, is assumed to be known; it can be assumed in a common form such as uniform or Gaussian, based on ground-truth information, or estimated using numerical techniques such as a Markov chain Monte Carlo (MCMC) sampler on the initial data [14], [44]. Let $\mathbf{Y}_k = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k]$ be the set of data observed at the first k steps and $\mathbf{X}_k = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k]$ be the sequence of unknown state vectors. It is desirable to estimate the state vector and its uncertainty at all steps via PDF $p(\mathbf{X}_k|\mathbf{Y}_k)$. Due to the computational cost, the problem is simplified by recursively estimating the marginal PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$ from $p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1})$, rather than estimating joint PDF $p(\mathbf{X}_k|\mathbf{Y}_k)$.

With $p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1})$ available, we can predict $p(\mathbf{x}_k|\mathbf{Y}_{k-1})$ through the transition PDF $p(\mathbf{x}_k|\mathbf{x}_{k-1})$. The transition density is determined by the state equation [Eq. (1)] and noise PDF $p(\mathbf{v}_k)$. Due to the first order Markov chain assumption of $\mathbf{x}_k$, $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ does not depend on data $\mathbf{Y}_{k-1}$. Density $p(\mathbf{x}_k|\mathbf{Y}_{k-1})$ can be written as:

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{Y}_{k-1}) &= \int p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{Y}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1}) \mathrm{d}\mathbf{x}_{k-1} \\ &= \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1}) \mathrm{d}\mathbf{x}_{k-1}. \end{aligned} \tag{5}$$

When a new measurement $\mathbf{y}_k$ becomes available, we can update the state $\mathbf{x}_k$ using the likelihood of the state vector

$$l(\mathbf{x}_k) = p(\mathbf{y}_k|\mathbf{x}_k) \tag{6}$$

and the prediction PDF $p(\mathbf{x}_k|\mathbf{Y}_{k-1})$ of Eq. (5). The new estimate of the state vector at k is calculated via Bayes theorem:

$$p(\mathbf{x}_k|\mathbf{Y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Y}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})}. \tag{7}$$

The denominator in Eq. (7) is the normalizing constant of $p(\mathbf{x}_k|\mathbf{Y}_k)$, termed the evidence, and may be expressed as:

$$p(\mathbf{y}_k|\mathbf{Y}_{k-1}) = \int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Y}_{k-1})\mathrm{d}\mathbf{x}_k. \tag{8}$$

Density $p(\mathbf{x}_k|\mathbf{Y}_k)$ contains all information necessary for the computation of marginal PDFs of elements of $\mathbf{x}_k$, moments, and modes. Optimal sequential estimators [45], such as the minimum mean squared error (MMSE) estimator, which calculates the expected value $(\mathrm{E}[\mathbf{x}_k|\mathbf{Y}_k])$ of the state after data have been observed, and the maximum a posteriori (MAP) estimator that provides the likely value of the state also after data measurement, can be readily implemented.

III. Kalman filters

The discussion of Section II-A raises the question of how to track optimally an evolving density $p(\mathbf{x_k}|\mathbf{Y_k})$. Kalman [46] showed that, if a Gaussian distribution can represent this evolving PDF, the optimum Bayesian estimator (one that attains the Bayesian Cramér Rao lower bound (CRLB) [47] with the MMSE) exists and can be computed analytically. This optimum sequential estimator has been termed the Kalman filter (KF). Having a Gaussian posterior at each step implies that:

1) State and measurement noise must be additive. Noise variables $\mathbf{v}_k$ and $\mathbf{w}_k$, and the prior $\mathrm{p}(\mathbf{x}_0)$ must be Gaussian.
2) $\mathbf{f}_k$ and $\mathbf{h}_k$ are known linear functions of the state and measurement vectors.

This reduces the system given in Eqs. (1–2) to:

$$\mathbf{x}_k = \mathbf{F}_k\mathbf{x}_{k-1} + \mathbf{v}_k \tag{9}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k, \tag{10}$$

where $\mathbf{F}_k$ and $\mathbf{H}_k$ replace $\mathbf{f}_k$ and $\mathbf{h}_k$ as matrices representing known linear functions of the state vector $\mathbf{x}_k$. State and measurement noise covariances $\mathbf{Q}_k$ and $\mathbf{R}_k$ are:

$$\begin{aligned} \mathrm{E}\{\mathbf{v}_k\mathbf{v}_i^T\} &= \mathbf{Q}_k\delta_{ki} \\ \mathrm{E}\{\mathbf{w}_k\mathbf{w}_i^T\} &= \mathbf{R}_k\delta_{ki} \\ \mathrm{E}\{\mathbf{v}_k\mathbf{w}_i^T\} &= \mathbf{0} \ \ \forall\, i,k. \end{aligned} \tag{11}$$

Even though the parameters given here are real, it is straightforward to extend these to complex cases. The requirements given above are fairly restrictive, however the resulting filter is straightforward to implement. Because a Gaussian PDF is uniquely and completely defined by its first two moments, mean and covariance, the KF needs only recomputing mean $\widehat{\mathbf{x}}$ and covariance $\mathbf{P}$ at step k, using $\widehat{\mathbf{x}}_{k-1}$, $\mathbf{P}_{k-1}$, and the new measurement $\mathbf{y}_k$.

This is a two-step procedure, repeated at each k:

- *Predict:* This stage predicts the current value of the state given its previous value using Eq. (9). This is represented by $\widehat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$.
- *Update:* This stage updates or corrects the values predicted in the previous stage given the information obtained from measurement $\mathbf{y}_k$. This provides the posterior at k represented by $\widehat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$.

The resulting set of equations for the linear, Gaussian KF is given in Table I [46], where $\mathbf{K}_k$ represents the Kalman gain.

TABLE I
Kalman filter

◊**Predict:**	
$\widehat{\mathbf{x}}_{k\|k-1}$	$= \mathbf{F}_k\widehat{\mathbf{x}}_{k-1\|k-1}$
$\mathbf{P}_{k\|k-1}$	$= \mathbf{F}_k\mathbf{P}_{k-1\|k-1}\mathbf{F}_k^{\mathrm{T}} + \mathbf{Q}_k$
◊**Update:**	
$\widehat{\mathbf{x}}_{k\|k}$	$= \widehat{\mathbf{x}}_{k\|k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\widehat{\mathbf{x}}_{k\|k-1}\right)$
$\mathbf{P}_{k\|k}$	$= \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_{k\|k-1}$
Kalman gain:	
$\mathbf{K}_k$	$= \mathbf{P}_{k\|k-1}\mathbf{H}_k^{\mathrm{T}}\left(\mathbf{R}_k + \mathbf{H}_k\mathbf{P}_{k\|k-1}{\mathbf{H_k}}^{\mathrm{T}}\right)^{-1}$

A. Extended Kalman filter

Although elegant and easy to implement, the KF has strict linearity and Gaussian PDF requirements that make it unsuitable for a large class of systems that cannot be characterized by Eqs. (9–10). An obvious way to extend the KF framework is linearizing functions $\mathbf{f}_k$ and $\mathbf{h}_k$ in Eqs. (1–2) around the current state value. When coupled with a Gaussian prior $\mathrm{p}(\mathbf{x}_0)$ and Gaussian noise terms, the posterior $\mathrm{p}(\mathbf{x}_k|\mathrm{Y}_k)$ will remain Gaussian. Thus, only mean and covariance need to be tracked via slightly modified KF equations.

The new filter, the extended Kalman filter [45], [48] (EKF), linearizes locally the state and measurement equations using the first terms in the Taylor series expansions. For the EKF to perform well, nonlinearities should be small and the underlying densities should be close to Gaussian. However, even under these circumstances, because of the implied approximations, the EKF cannot claim the optimality enjoyed by the KF for linear-Gaussian systems. For large nonlinearities, the mean and covariances will not be mapped correctly as shown in the example of Sec. III-B. Nevertheless, the EKF has been implemented successfully in numerous applications in areas such as radar and sonar target tracking among others. The EKF algorithm is summarized in Table II [45], [48].

B. Unscented Kalman filter

The simplicity of the Kalman framework arises from its Gaussian nature. The filter only needs to carry the mean and covariance to the next step. A nonlinear system will disrupt the flow of evolving Gaussian PDFs, because a Gaussian random variable passing through a nonlinear transformation will lose its Gaussian form and solely mean and covariance will not be adequate in fully defining this density. This realization leads to the development of KFs that use sigma points coupled with unscented transforms (UT) [49]. The UT enables the propagation of the mean and covariance through nonlinear functions. KFs that make use of the UT are called unscented Kalman filters (UKF) [50].

While the EKF enforces linearity through analytic linearization, the UKF enforces a Gaussian distribution while keeping the functional nonlinearity. This means that the filter assumes that any nonlinear transformation of a Gaussian input PDF will always be another Gaussian PDF. Since, under this

TABLE II
EXTENDED KALMAN FILTER

◊**Predict:**	
$\widehat{\mathbf{x}}_{k\|k-1}$	$= \mathbf{f}_k\left(\widehat{\mathbf{x}}_{k-1\|k-1}\right)$
$\mathbf{P}_{k\|k-1}$	$= \mathbf{F}_k\mathbf{P}_{k-1\|k-1}\mathbf{F}_k^{\mathrm{T}} + \mathbf{Q}_k$
◊**Update:**	
$\widehat{\mathbf{x}}_{k\|k}$	$= \widehat{\mathbf{x}}_{k\|k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{h}_k\left(\widehat{\mathbf{x}}_{k\|k-1}\right)\right)$
$\mathbf{P}_{k\|k}$	$= \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_{k\|k-1}$
Kalman gain:	
$\mathbf{K}_k$	$= \mathbf{P}_{k\|k-1}\mathbf{H}_k^{\mathrm{T}}\left(\mathbf{R}_k + \mathbf{H}_k\mathbf{P}_{k\|k-1}\mathbf{H}_k^{\mathrm{T}}\right)^{-1}$
Jacobians:	
$\mathbf{F}_k$	$= \frac{\partial \mathbf{f}_k}{\partial \mathbf{x}_{k-1}}\Big\|_{\mathbf{x}_{k-1}=\widehat{\mathbf{x}}_{k-1\|k-1}}$
$\mathbf{H}_k$	$= \frac{\partial \mathbf{h}_k}{\partial \mathbf{x}_k}\Big\|_{\mathbf{x}_k=\widehat{\mathbf{x}}_{k\|k-1}}$

TABLE III
UNSCENTED KALMAN FILTER. ($\left(\sqrt{\cdot}\right)_i$ IS THE iTH COLUMN OF THE MATRIX SQUARE ROOT [49], [50].)

◊**Predict:**	
$\widehat{\mathbf{x}}_{k\|k-1}$	$= \sum_{i=0}^{2n_{\mathrm{x}}} W_{\mathrm{m}}^i \mathcal{X}_{k\|k-1}^i$
$\mathbf{P}_{k\|k-1}$	$= \mathbf{Q}_k$
	$+ \sum_{i=0}^{2n_{\mathrm{x}}} W_{\mathrm{cov}}^i \left[\mathcal{X}_{k\|k-1}^i - \widehat{\mathbf{x}}_{k\|k-1}\right]\left[\mathcal{X}_{k\|k-1}^i - \widehat{\mathbf{x}}_{k\|k-1}\right]^T$
◊**Update:**	
$\widehat{\mathbf{x}}_{k\|k}$	$= \widehat{\mathbf{x}}_{k\|k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \widehat{\mathbf{y}}_{k\|k-1}\right)$
$\mathbf{P}_{k\|k}$	$= \mathbf{P}_{k\|k-1} - \mathbf{K}_k\left(\mathbf{P}_{yy} + \mathbf{R}_k\right)\mathbf{K}_k^T$
Sigma point generation:	
$\mathcal{X}_{k-1}^0$	$= \widehat{\mathbf{x}}_{k-1\|k-1}$
W_{m}^0	$= \frac{\lambda}{n_{\mathrm{x}}+\lambda}, \qquad W_{\mathrm{cov}}^0 = W_{\mathrm{m}}^0 + \beta + 1 - \alpha^2$
$\mathcal{X}_{k-1}^i$	$= \widehat{\mathbf{x}}_{k-1\|k-1} + \left(\sqrt{(n_{\mathrm{x}}+\lambda)\mathbf{P}_{k-1\|k-1}}\right)_i$
W_{m}^i	$= W_{\mathrm{cov}}^i = \frac{0.5}{n_{\mathrm{x}}+\lambda} \qquad i = 1, \ldots, n_{\mathrm{x}}$
$\mathcal{X}_{k-1}^i$	$= \widehat{\mathbf{x}}_{k-1\|k-1} - \left(\sqrt{(n_{\mathrm{x}}+\lambda)\mathbf{P}_{k-1\|k-1}}\right)_i$
W_{m}^i	$= W_{\mathrm{cov}}^i = \frac{0.5}{n_{\mathrm{x}}+\lambda} \qquad i = n_{\mathrm{x}}+1, \ldots, 2n_{\mathrm{x}}$
Unscented transform:	
$\mathcal{X}_{k\|k-1}^i$	$= \mathbf{f}_k\left(\mathcal{X}_{k-1}^i\right)$
$\mathcal{Y}_{k\|k-1}^i$	$= \mathbf{h}\left(\mathcal{X}_{k\|k-1}^i\right), \quad \widehat{\mathbf{y}}_{k\|k-1} = \sum_{i=0}^{2n_{\mathrm{x}}} W_{\mathrm{m}}^i \mathcal{Y}_{k\|k-1}^i$
$\mathbf{P}_{xy}$	$= \sum_{i=0}^{2n_{\mathrm{x}}} W_{\mathrm{cov}}^i \left[\mathcal{X}_{k\|k-1}^i - \widehat{\mathbf{x}}_{k\|k-1}\right]\left[\mathcal{Y}_{k\|k-1}^i - \widehat{\mathbf{y}}_{k\|k-1}\right]^T$
$\mathbf{P}_{yy}$	$= \sum_{i=0}^{2n_{\mathrm{x}}} W_{\mathrm{cov}}^i \left[\mathcal{Y}_{k\|k-1}^i - \widehat{\mathbf{y}}_{k\|k-1}\right]\left[\mathcal{Y}_{k\|k-1}^i - \widehat{\mathbf{y}}_{k\|k-1}\right]^T$
Kalman gain:	
$\mathbf{K}_k$	$= \mathbf{P}_{xy}\left(\mathbf{P}_{yy} + \mathbf{R}_k\right)^{-1}$

assumption, a Gaussian input results in a Gaussian output, the system is effectively linearized in a statistical rather than analytical sense. This still enables the filter to carry all necessary information by propagating only the mean and covariance as required by the KF.

In the UT a number of weighted sigma points are carefully chosen so that they capture the mean and covariance of the state variable (the sigma points depend on the matrix square root of covariance $\mathbf{P}_k$, see Table III). As the random variable undergoes a nonlinear transformation, these points are propagated through this nonlinear function and are used to reconstruct the new mean and covariance using the UT weights [49]. Hence, unlike with the EKF, mean and covariance to at least second order of the nonlinearity (third if the initial PDF is Gaussian) can be computed accurately.

The UKF uses a recursive formulation where $2n_{\mathrm{x}}+1$ sigma points $\left\{\mathcal{X}^i\right\}_{i=0}^{2n_{\mathrm{x}}}$ and their corresponding weights W^i are generated. These are used with the unscented transform (UT) to perform the mean $\widehat{\mathbf{x}}_k$ and covariance $\mathbf{P}_k$ calculations required in the Kalman framework. The UT weights are computed using three independent parameters: prior knowledge parameter β, scaling parameter κ, and parameter α, which controls the spread of the sigma points around the mean. The second scaling parameter λ can be computed from these parameters: $\lambda = \alpha^2(n_{\mathrm{x}}+\kappa)-n_{\mathrm{x}}$. The algorithm is presented in Table III. Details about the derivation can be found in [49].

Although UKFs are derivative-free, an improvement over EKFs, and efficient relative to more advanced techniques to be discussed later, they have two disadvantages as demonstrated in the example below.

1) Example: EKF vs. UKF vs. PF: Assume that a Gaussian random variable $\mathbf{x}$ with $p(\mathbf{x}) \sim \mathcal{N}(\mu_{\mathbf{x}}, \mathbf{\Sigma}_{\mathbf{x}})$ goes through a nonlinear transformation $\mathbf{y} = \mathbf{h}(\mathbf{x})$ with:
$\mu_{\mathbf{x}} = \begin{bmatrix}0.25\\10\end{bmatrix}$, $\mathbf{\Sigma}_{\mathbf{x}} = \begin{bmatrix}0.1 & 0.2\\0.2 & 0.5\end{bmatrix}$, $\mathbf{y} = \begin{bmatrix}\mathbf{x}^2(1) - \mathbf{x}(1)\mathbf{x}(2)\\\mathbf{x}(2)\cos\mathbf{x}(1)\end{bmatrix}$.
A cloud of 10,000 particles drawn from $p(\mathbf{x})$ is plotted in Fig. 2 together with the mean and 3σ covariance ellipse. When these particles are propagated through the nonlinear equation, the distribution becomes non-Gaussian as shown.

The analytical linearization in EKF uses the Jacobian: $\mathbf{H} = \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} = \begin{bmatrix}2\mathbf{x}(1) - \mathbf{x}(2) & -\mathbf{x}(1)\\-\mathbf{x}(2)\sin\mathbf{x}(1) & \cos\mathbf{x}(1)\end{bmatrix}$. This will result in a Gaussian density with a mean and covariance of $\widehat{\mu}_{\mathbf{y}} = \mathbf{h}(\mu_{\mathbf{x}})$ and $\widehat{\mathbf{\Sigma}}_{\mathbf{y}} = \mathbf{H}\mathbf{\Sigma}_{\mathbf{x}}\mathbf{H}^{\mathrm{T}}$ represented by $\times$ and a dashed ellipsoid in Fig. 2. Both mean and covariance have high errors in their estimates.

Mean and covariance are also obtained by using the sigma point generation and a successive UT given in Table III. The values obtained by UT are very close to their true values with almost identical ellipses. Fig. 2 highlights two main drawbacks of KFs in nonlinear, non-Gaussian problems:

The first is the effect of nonlinearity on the accurate estimation of mean and covariance. This example is 'too nonlinear' for the EKF but not for the UKF for obtaining good mean and covariance estimates. However, there are many cases where the nonlinearity may be severe and may require an even higher order accuracy than the UKF can provide to correctly capture

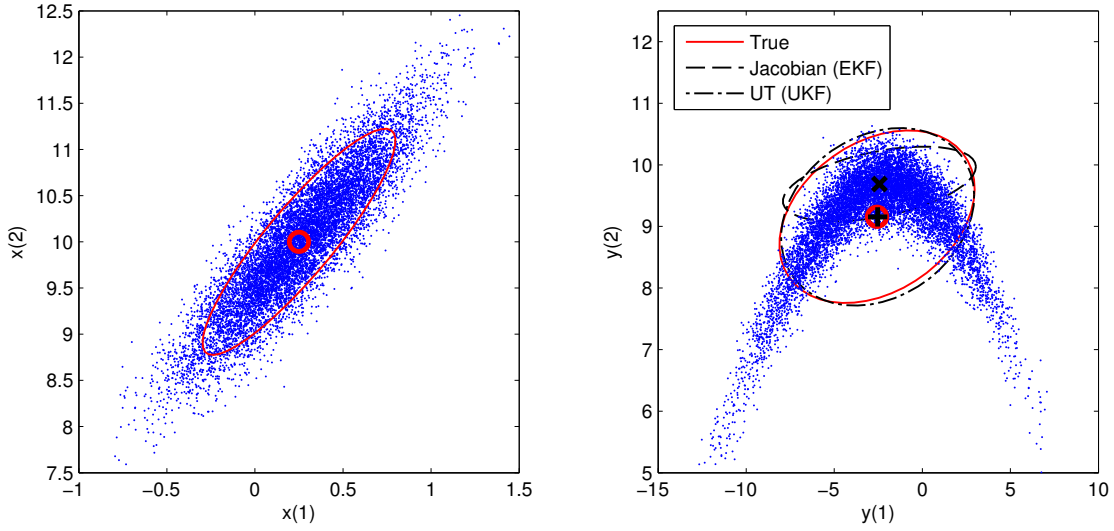

Fig. 2. Mean and covariance propagation of the EKF and UKF in a nonlinear, non-Gaussian system with $\mathbf{y} = \mathbf{h}(\mathbf{x})$. True mean ($\circ$), means from analytical ($\times$) and statistical ($+$) linearizations are given with covariance estimates in terms of 3σ ellipses for the true covariance and the estimates from the Jacobian (EKF) and the UT (UKF).

mean and covariance.

The other is that the densities may be too non-Gaussian to be represented using the first two moments, even if those can be calculated correctly. The non-Gaussian PDF $p(\mathbf{y})$ cannot be sufficiently characterized with mean and covariance alone in this example. In such cases, the only way to accurately propagate complex, non-Gaussian PDFs is to store the cloud of particles that represent the PDF as shown in Fig. 2. This forms the basic philosophy of particle filtering.

The UKF is the first and most widely used sigma-point KF. It is possible to select different sigma points and UT to match more than the first two moments. Moreover, less complicated and more stable versions of UKF with reduced computational cost have also been proposed. [51]

C. Ensemble Kalman filter

The Ensemble KF (EnKF) [52] is a Kalman variant for systems where the state is composed of a large number of variables. This type of problem is encountered frequently in geophysical and hydrological fields, ocean and atmospheric modeling, and data analysis [53]. As discussed in [54] and Sec. V-E, PFs are not appropriate for problems of high dimensionality.

The EnKF is a KF that includes MC sampling. It is possible to argue that the EnKF is a hybrid between a KF and a PF (see Section IV for details on PFs). The EnKF assumes that the large number of state variables and all PDFs are Gaussian, hence the problem is based on a Kalman framework. However, working with a large covariance matrix in a high-dimensional system is computationally inefficient for the classical KF. This drawback is addressed by using an ensemble of points, similar to the particle set in a PF, to replace the covariance matrix with the ensemble covariance matrix using MC sampling. For detailed theory and practical implementation see Ref. [55].

IV. Particle filters

Sometimes, the dynamics of a problem cannot be adequately captured by the Kalman family of filters, because of the nonlinear structure of state and measurement equations and the nature of the noise. When KFs and their extensions fail, the estimation of $\mathbf{x}_k$ via filtering posterior PDFs $p(\mathbf{x}_k|\mathbf{Y}_k)$ may be possible using numerical sequential MC techniques, commonly referred to as PFs.

PFs track the posterior PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$ using a cloud of particles $\{\mathbf{x}_k^i\}_{i=1}^{N_p} = \{\mathbf{x}_k^1, \ldots, \mathbf{x}_k^{N_p}\}$ that evolve with step k. All other estimates can be calculated from $p(\mathbf{x}_k|\mathbf{Y}_k)$, including $\widehat{\mathbf{x}}_k^{MMSE}$ and $\widehat{\mathbf{x}}_k^{MAP}$ solutions, the variance in the estimates, and the marginal posterior PDF $p(\mathbf{x}_k(j)|\mathbf{Y}_k)$ of the jth state variable:

$$p(\mathbf{x}_k(j)|\mathbf{Y}_k) = \int \delta(\mathbf{x}'_k(j) - \mathbf{x}_k(j)) p(\mathbf{x}'_k|\mathbf{Y}_k) \mathrm{d}\mathbf{x}'_k. \quad (12)$$

A. Importance Sampling

Before presenting the details of the PF, we summarize how importance sampling (IS) and MC integration work. IS is a method employed to compute expectations with respect to one density using random samples drawn from another. Assume that we want to compute an integral $I = \int f(\mathbf{x})\mathrm{d}\mathbf{x}$. One way of computing I is using samplers [56] after assuming $\mathbf{x}$ is a random variable, defining $f(\mathbf{x}) = g(\mathbf{x})p(\mathbf{x})$, and rewriting I in the form of an expectation:

$$I \equiv E_p\{g(\mathbf{x})\} = \int_{\mathcal{X}} g(\mathbf{x}) p(\mathbf{x}) \mathrm{d}\mathbf{x} \ , \quad (13)$$

where $g(\mathbf{x})$ is some function of $\mathbf{x}$ with PDF $p(\mathbf{x})$. This integral can be computed numerically via MC integration [56], by drawing N_p independent and identically distributed $\mathbf{x}$ samples

from the sampling or proposal density $p(\mathbf{x})$:

$$\{\mathbf{x}^i\}_{i=1}^{N_p} \sim p(\mathbf{x}) \quad \longrightarrow \quad I \approx \frac{1}{N_p}\sum_{i=1}^{N_p} g(\mathbf{x}^i) \tag{14}$$

This method is commonly referred to as the 'perfect MC' [40]. This is an unbiased integral estimator, which converges to the true value with diminishing error variance as the number of particles drawn from $p(\mathbf{x})$ increases. Another important property of perfect MC is that the convergence is independent of the state dimension, a critical property that IS and many other sampling schemes lack.

In many cases it is too costly or not possible to sample from $p(\mathbf{x})$. Therefore, selection of an appropriate sampling PDF is the critical part in IS. It is customary to use a PDF easy to sample from, such as a Gaussian or uniform. However, in Bayesian tracking and inversion, the integral is already in the form of Eq. (13), where $p(\mathbf{x})$ is a complex PDF, *e.g.* Eq. (12). It is still possible to use a simple function $q(\mathbf{x})$ as the sampling density by rewriting Eq. (13) and using MC integration to get:

$$I = \int_{\mathcal{X}} \left[g(\mathbf{x})\frac{p(\mathbf{x})}{q(\mathbf{x})}\right] q(\mathbf{x})\mathrm{d}\mathbf{x} \equiv E_q\left\{g(\mathbf{x})\frac{p(\mathbf{x})}{q(\mathbf{x})}\right\}$$
$$\{\mathbf{x}^i\}_{i=1}^{N_p} \sim q(\mathbf{x}) \tag{15}$$

The estimate is

$$\widehat{I} = \frac{1}{N_p}\sum_{i=1}^{N_p} g(\mathbf{x}^i)\frac{p(\mathbf{x}^i)}{q(\mathbf{x}^i)} \tag{16}$$

$$= \frac{1}{N_p}\sum_{i=1}^{N_p} w^i g(\mathbf{x}^i) \quad \text{with} \quad w^i = \frac{p(\mathbf{x}^i)}{q(\mathbf{x}^i)} \ , \tag{17}$$

where w^i are the importance weights. The variance is [56]:

$$\mathrm{Var}_q(\widehat{I}) = \left\{E_q\left[\frac{g(\mathbf{x})p(\mathbf{x})}{q(\mathbf{x})}\right]^2 - I^2\right\}/N_p \quad . \tag{18}$$

The variance in the estimate is minimum if $q(\mathbf{x})$ is proportional to $g(\mathbf{x})p(\mathbf{x})$ and increases as $q(\mathbf{x})$ deviates from the latter. Using IS requires the selection of a balanced $q(\mathbf{x})$: as easy to sample from as possible without sacrificing the accuracy of the method.

B. Sequential Importance Sampling

Bayesian filtering requires performing successive IS runs at each k. The output of each IS run is used as the prior for the next one. This process is referred to as sequential importance sampling (SIS). Recalling that $\mathbf{Y}_k = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k]$ and $\mathbf{X}_k = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k]$ from Section II-A, it is possible to obtain the filtering PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$ from the full posterior density $p(\mathbf{X}_k|\mathbf{Y}_k)$:

$$p(\mathbf{x}_k|\mathbf{Y}_k) = \int p(\mathbf{X}_k|\mathbf{Y}_k)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2\ldots\mathrm{d}\mathbf{x}_{k-1} \ , \tag{19}$$

$$= \int \delta(\mathbf{x}_k - \mathbf{x}_{k'})p(\mathbf{X}_{k'}|\mathbf{Y}_{k'})\mathrm{d}\mathbf{X}_{k'} \ , \tag{20}$$

which is in the same form as Eq. (13). Assuming a sampling density $q(\mathbf{X}_k|\mathbf{Y}_k)$ and performing IS, we obtain

$$p(\mathbf{x}_k|\mathbf{Y}_k) \approx \sum_{i=1}^{N_p} w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i) \ , \tag{21}$$

$$w_k^i \propto \frac{p(\mathbf{X}_k^i|\mathbf{Y}_k)}{q(\mathbf{X}_k^i|\mathbf{Y}_k)} \ . \tag{22}$$

However, to perform IS sequentially at step k, we utilize the IS results of the previous step $k-1$, that is, we use the cloud of particles and associated weights $\{\mathbf{x}_{k-1}^i, w_{k-1}^i\}_{i=1}^{N_p}$. This is done by selecting $q(\mathbf{X}_k|\mathbf{Y}_k)$:

$$q(\mathbf{X}_k|\mathbf{Y}_k) = q(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{Y}_k)q(\mathbf{X}_{k-1}|\mathbf{Y}_{k-1}) \ . \tag{23}$$

Expanding the full posterior PDF (only the results shown here, for additional details see [41]), we can write:

$$p(\mathbf{X}_k|\mathbf{Y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})}p(\mathbf{X}_{k-1}|\mathbf{Y}_{k-1}) \ , \tag{24}$$

where the weight of the ith particle at step k is computed by inserting Eqs. (23–24) into Eq. (22) and ignoring the constant term $p(\mathbf{y}_k|\mathbf{Y}_{k-1})$:

$$w_k^i \propto \frac{p(\mathbf{y}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i, \mathbf{Y}_k)} \times \frac{p(\mathbf{X}_{k-1}^i|\mathbf{Y}_{k-1})}{q(\mathbf{X}_{k-1}^i|\mathbf{Y}_{k-1})} \tag{25}$$

$$\propto \frac{p(\mathbf{y}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i, \mathbf{Y}_k)}w_{k-1}^i \ . \tag{26}$$

A key issue in PF design is choosing a good proposal density. A simple choice is $q(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{Y}_k) = p(\mathbf{x}_k|\mathbf{x}_{k-1})$, reducing Eq. (26) to

$$w_k^i = p(\mathbf{y}_k|\mathbf{x}_k^i)w_{k-1}^i \ . \tag{27}$$

Note that, while the full posterior PDF $p(\mathbf{X}_k|\mathbf{Y}_k)$ is used in deriving IS Eqs. (19–27), only the likelihood $p(\mathbf{y}_k|\mathbf{x}_k)$ at step k is employed in updating weights w_k^i. The implementation of PF is only concerned with the marginal PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$, as seen in Table IV.

C. Sequential Importance Resampling

Although SIS provides a complete framework for performing sequential Bayesian estimation, its implementation quickly runs into the problem of sampling degeneracy. After a few iterations of successive SIS, the process leads to a cloud containing few particles with large weights and numerous particles with negligible weights. In the extreme case, there will be only one particle left with a large w_k^i. This loss of sample diversity results in poor filter performance.

To counter degeneracy, a second sampling stage is proposed right after the update stage [57]. The purpose of this resampling stage [58] is to create more high-weight particles from the original set of particles $\{\mathbf{x}_k^i\}_{i=1}^{N_p}$. The modified filter is called sequential importance resampling (SIR) [57], [59], [60] and is the most popular PF implementation.

Resampling is easily performed at the end of each step k. Alternatively, resampling is implemented when the effective number of particles N_p^{eff} needed to maintain diversity drops

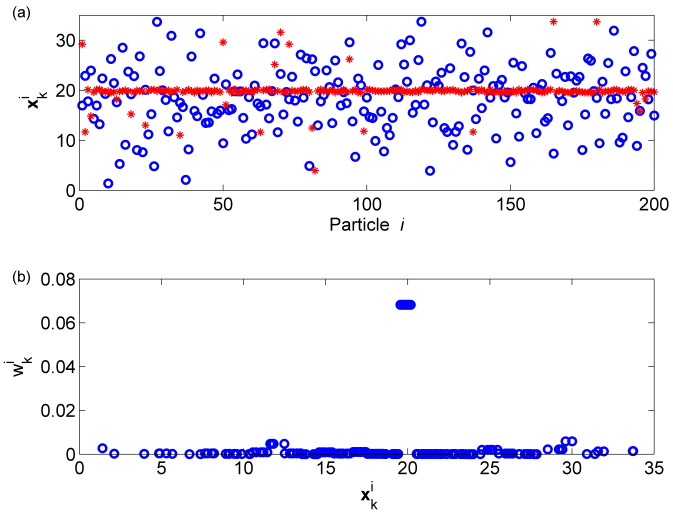

Fig. 3. The need for resampling: (a) Particles before ($\circ$) and after ($*$) resampling, (b) particles before resampling and their weights $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^{N_p}$ for $k = 10$.

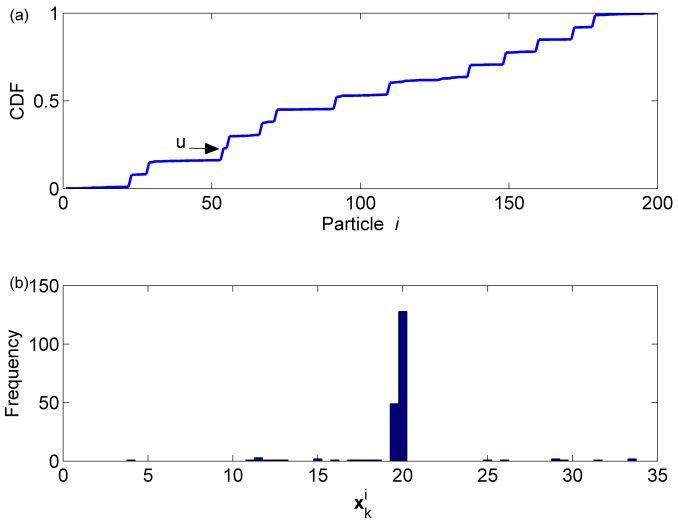

Fig. 4. (a) Resampling via the CDF, (b) histogram of particles after resampling showing large number of newly created particles at high likelihood regions.

below a threshold [61]. The effective number of particles can be estimated as

$$N_p^{eff} = \frac{1}{\sum_{i=1}^{N_p}(w_k^i)^2} \ . \tag{28}$$

The resampling process takes the posterior filtering PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$ represented by the particle set $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^{N_p}$ of SIS at the end of each step and redistributes samples so that all weights of the new particles are the same, that is, $\{\mathbf{x}_k^j, w_k^j = 1/N_p\}_{j=1}^{N_p}$. This results in a larger number of particles in the high likelihood regions, preventing degeneracy. It is best to examine the degeneracy and resampling via an example.

1) Example, need for resampling: Resampling is demonstrated for a SIR PF tracking a target at step $k = 10$ in Figs. 3–4. The filter has $N_p = 200$ particles for state variable $\mathbf{x}_k$ representing target location. Target location is assumed to be between 0–35 with true value $\mathbf{x}_{10} = 20$. The state equation is: $\mathbf{x}_k = \mathbf{x}_{k-1} + \text{round}(\mathbf{v}_k)$, where $\mathbf{v}_k$ is zero mean Gaussian noise. (The example is discussed in more detail in Section IV-E.1).

Fig. 3(a) shows 200 particles (o) that predict $\mathbf{x}_k$. During the update stage, weights for those particles are computed employing a suitable likelihood function. Only four particles have significant weights and provide substantial information on the posterior PDF, see Fig. 3(b). Resampling is needed for the generation of more particles around $\mathbf{x}_{10} = 20$, so that SIS at step $k + 1$ will not be impacted by degeneracy.

Fig. 4 illustrates the implementation of the resampling stage. Resampling creates a new discrete density by first computing the cumulative distribution function (CDF) of $p(\mathbf{x}_k|\mathbf{Y}_k)$ from $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^{N_p}$, see Fig. 4(a). Using a sample-drawing technique that randomly picks samples from a [0,1] uniform density, new particles are generated via the inverse CDF. Regions with large weights are replicated multiple times in the new cloud and particles with small weights are less likely to be selected. The new particles after resampling ($*$) are much more focused than the original set of particles, see Fig. 3(a). The histogram of resampled particles in Fig. 4(b) provides the new estimate of $p(\mathbf{x}_k|\mathbf{Y}_k)$ via the likelihood function. Almost 90% of the particles are now in the high likelihood region.

2) SIR algorithm: We can merge the SIS and resampling to create the SIR algorithm outlined below:

Predict: This stage is a SIS prediction that starts with the cloud of equal weigh particles from the previous step $\{\mathbf{x}_{k-1}^i\}_{i=1}^{N_p}$, and creates a new set of predictions for the current step $\{\mathbf{x}_{k|k-1}^i\}_{i=1}^{N_p}$ by sampling from the transitional density $p(\mathbf{x}_k|\mathbf{x}_{k-1})$. This is done by propagating each $\mathbf{x}_{k-1}^i$ through the state equation Eq. (1) together with a random realization from $\mathbf{v}_k$. This step corresponds to the implementation of Eq (5), see Table IV.

Update: Note that $w_{k-1}^i = 1/N_p$ for all i because of resampling at step $k-1$. Having measured $\mathbf{y}_k$, the weight of each particle is re-evaluated and normalized via Eq. (27):

$$w_k^i = \frac{p(\mathbf{y}_k|\mathbf{x}_{k|k-1}^i)}{\sum_{i=1}^{N_p} p(\mathbf{y}_k|\mathbf{x}_{k|k-1}^i)} \ , \tag{29}$$

where $p(\mathbf{y}_k|\mathbf{x}_{k|k-1}^i)$ is the likelihood function defined by the measurement equation [Eq. (2)], which includes statistical behavior of errors in the data. The weights of Eq. (29) are used in Eq. (21) for expressing the posterior PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$.

Resample: New particles $\{\mathbf{x}_{k|k}^j, w_k^j = 1/N_p\}_{j=1}^{N_p}$ are drawn from a discrete approximation to density $p(\mathbf{x}_k|\mathbf{Y}_k)$ obtained at the *update* stage. This is achieved through the calculation of the CDF of $\mathbf{x}_k$ as explained above. Different types of resampling are presented in [39], [62]. All particle weights are now equal to $1/N_p$.

A single iteration of a ten-particle SIR algorithm is illustrated in Fig. 5. A cylinder with height proportional to the likelihood represents each particle. As usual, the first stage is

TABLE IV
SIR PARTICLE FILTER

◊**Predict:**		
$\{\mathbf{x}^i_{k\|k-1}\}_{i=1}^{N_p}$	$\sim p(\mathbf{x}_k\|\mathbf{x}_{k-1})$	given $\{\mathbf{x}^i_{k-1\|k-1}\}_{i=1}^{N_p}$
$\mathbf{x}^i_{k\|k-1}$	$= \mathbf{f}_k\left(\mathbf{x}^i_{k-1\|k-1}, \mathbf{v}^i_k\right)$	$i = 1, \ldots, N_p$
◊**Update:**		
w^i_k	$= \frac{p(\mathbf{y}_k\|\mathbf{x}^i_{k\|k-1})}{\sum_{i=1}^{N_p} p(\mathbf{y}_k\|\mathbf{x}^i_{k\|k-1})}$	
$p(\mathbf{x}_k\|\mathbf{Y}_k)$	$\approx \sum_{i=1}^{N_p} w^i_k \delta(\mathbf{x}_k - \mathbf{x}^i_{k\|k-1})$	
◊**Resample:**		
$\{\mathbf{x}^i_{k\|k-1}, w^i_k\}_{i=1}^{N_p}$	$\longmapsto \{\mathbf{x}^j_{k\|k}, \frac{1}{N_p}\}_{j=1}^{N_p}$	s.t.
$p(\mathbf{x}_k\|\mathbf{Y}_k)$	$\approx \frac{1}{N_p}\sum_{j=1}^{N_p} \delta(\mathbf{x_k} - \mathbf{x^j_{k\|k}})$	

prediction, where new particles are created from the particles representing the PDF of the previous step, $\mathrm{p}(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1})$, the state noise at step k, and Eq. (1). The likelihood is calculated for each of these new particles using new data $\mathbf{y}_k$. Resampling follows this update stage, where a new set of particles is formed from the previous one. The larger the weight of a particle, the more new particles it generates during resampling.

Resampling sometimes adds significantly to the computational requirements of a PF. In this case, quasi-random resampling methods can be used to decrease computational cost as an alternative to the method explained in Fig. 4. The commonly used efficient resampling algorithm is systematic resampling [40]. In this approach, only one random sample is picked from a [0,$1/N_p$] uniform distribution and the CDF is sliced with N_p equidistant points with $1/N_p$ spacing starting from the random sample. Original particles are then converted to new particles using the inverse CDF as shown on the right in Fig. 5. Cylinders representing each particle are stacked to obtain the CDF. The new particles are selected by randomly choosing $\mathbf{x}^1$ and equally sampling upwards. Further discussion about resampling algorithms is given in Section V-F.

To implement the filtering process for $k = 1$, an initial probability density for state vector $\mathbf{x}_0$ has to be selected. Prior $p(\mathbf{x}_0)$ can be obtained via a sampler. Alternatively, a uniform or Gaussian PDF can be used. The latter requires more steps before stable estimates are obtained. One can sample from $p(\mathbf{x}_0)$ to generate the initial cloud $\{\mathbf{x}^i_0\}_{i=1}^{N_p}$. Particles in this cloud will be propagated through the *predict, update, resample* process for the estimation of the unknown parameters at the subsequent step.

The process at step k is summarized in Table IV.

D. Advanced Particle Filters

Although SIR provides robust tracking performance in many nonlinear/non-Gaussian tracking problems, it has two basic flaws:

1) Sample impoverishment: Since the resampling stage creates many exact replicas of high likelihood particles, it is possible to lose sample diversity as the filter progresses.
2) Sensitivity to outliers: The popularity of the SIR filter lies in the simplicity of the selection of transition prior $\mathbf{p}(\mathbf{x}_k|\mathbf{x}_{k-1})$ as the sampling density. This is, at the same time, a weakness, because the prediction at step k does not use the newly available information in measurement $\mathbf{y}_k$, performing a 'blind' prediction. The filter becomes vulnerable to outliers, especially when the likelihood is peaked relative to the prior [41].

To address these problems, advanced PF variants have been developed. [35], [37], [41] Common ones are briefly discussed in Appendix II in terms of the specific concern they address, their main advantages, and possible drawbacks. Comparative information is provided in Table V, summarizing salient features and differences.

E. Sequential model selection

Inversion in ocean acoustics typically assumes a known number of parameters that need to be estimated. In geoacoustic inversion, an acoustic propagation model is selected that relates sound to a specific number of environmental parameters. Hence, the model is fixed and known. It was recognized early on that the assumption of a fixed model is not optimal [67]. This observation has lead to efforts towards non-sequential model selection in ocean acoustics with approaches such as hypothesis testing [68], Akaike and Bayesian (Schwartz-Rissanen) information criteria [68]–[70], and evidence calculation [15], [71]. A reversible jump MCMC, where the MC can switch between different models, seems most promising and efficient. For an early implementation see Ref. [72], with an extensive derivation and description in [73].

In dynamic problems, the model may be unknown and changing for each step, similarly to other unknown parameters. Therefore, a sequential state-space system may need more than one model as new data become available. In ocean acoustics, this problem may be encountered in:

- Tracking an unknown number of targets [74].
- Tracking targets with different motion models such as slow moving and fast moving targets, targets with large accelerations, maneuvering targets.
- Tracking a changing number of unknown vertical modes of the acoustic field, the number of which changes because of range dependence. Tracking an unknown number of modes in time-frequency representations.
- Switching between ray tracing, normal mode, adiabatic normal mode, near-field normal mode, narrow-angle parabolic equation, and wide-angle parabolic equation models that relate data to state parameters, depending on factors such as distance and environmental complexity between source and receiver.
- Geoacoustic tracking of sediments and their varying properties with an unknown number of layers.
- Arrival time tracking across phones.

Successful tracking under such circumstances requires sequential filtering algorithms that are capable of jumping or

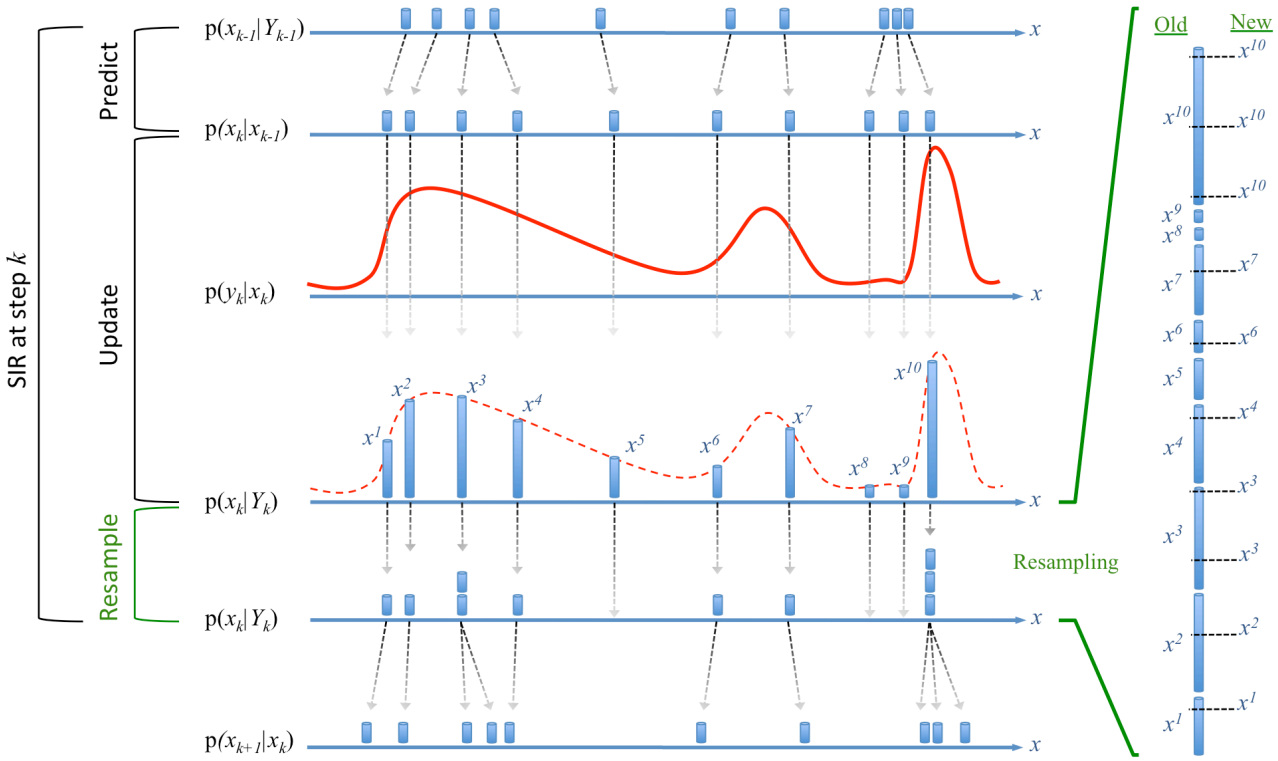

Fig. 5. Particle flow for a SIR algorithm with $N_p = 10$ particles.

TABLE V

FILTER COMPARISON. L: LINEAR G: GAUSSIAN NL: NONLINEAR NG: NON-GAUSSIAN.

Filter	Sect	Problem	PDF	Effort	Reference	Details
KF	Sec III	L/G	μ and Σ	Min.	Kalman [46]	Optimal filter when problem is L/G.
EKF	Sec III-A	mild NL/NG	μ and Σ	Low	Cox [48] Kay [45]	Analytical linearization by using the Jacobians. Reduces the NL/NG problem to a L/G one.
UKF	III-B	moderate NL/NG	μ and Σ sigma pts	Low	Julier [49] Wan [50]	Statistical linearization: Unscented transform propagates μ and Σ. Derivative free, reduces the NL/NG problem to a NL/G one.
EnKF	Sec III-C	moderate NL/NG	μ and Σ ensemble	Med-High	Evensen [52]	Ensemble statistics are used to propagates μ and Σ. Reduces the NL/NG problem to a L/G one.
SIS	Sec IV-B	NL/NG	particles	High	Doucet [40]	Degeneracy problem.
SIR	Sec IV-C	NL/NG	particles	High	Gordon [57]	Resampling avoids degeneracy but results in sample impoverishment. Blind prediction stage, potentially sensitive to outliers and inefficient.
RPF	App II-A	NL/NG	particles	High	Musso [63]	Impoverishment solved using kernel funcs., problems in larger dim. The kernel approx. violates the asymptotic convergence to posterior.
MCMCPF	App II-B	NL/NG	particles	High	Gilks [60]	Impoverishment solved using a MCMC kernel. A MCMC sampler is inserted in the resampling stage.
APF	App II-C	NL/NG	particles	High	Pitt [64]	Blind prediction and degeneracy issues solved using an aux. variable. Resampling in prediction stage, problems with large state noise.
RBPF	App II-D	NL/NG	particles	High	Liu [65] Doucet [59]	Effective when some parameters are conditionally L/G. L/G part of the state is solved using a KF, NL/NG part using a PF.
LPF	App II-E	NL/NG	particles	High	Doucet [59] Merwe [66]	Each particle is treated as a KF, creating a bank of EKF/UKFs. Implementation using UKFs is called unscented PF.

switching between models [63], [75]. These filters are referred to as the multiple model particle filters (MMPF) [41]: The state space is augmented with model order m_k, becoming $[\mathbf{x}_k^{\mathrm{T}} \; m_k]^{\mathrm{T}}$, where parameter m_k takes discrete values corresponding to different models. The particles contain samples both for $\mathbf{x}_k$ and m_k.

At each step, $\{m_k^i\}_{i=1}^{N_p}$, are updated using values m_{k-1}^i and the model order probability transition matrix, creating a mixture of particles that use different models. Once model order parameter m_k^i within the ith particle is determined, a "model-conditioned" SIR is run, where the particles, after updates imposed by the state equation, are propagated through the likelihood, relating measurements and state [41]. Since each model has a different number of parameters, the length of the state vector varies with state. Because MMPFs can jump or switch between different models, they are referred to as jump Markov or switching filters.

1) Example: Multiple-model target tracking: The example of Section IV-C is expanded to include a second target that appears at step $k = 11$. Thus, the single-target model is assigned the model order $m_k = 1$ and the two-target model uses $m_k = 2$.

An important task is the selection of a transition probability matrix, which determines how model order m_{k-1} transitions into m_k. Here:

$$T_m = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}, \tag{30}$$

where it is assumed that the probability of a sudden appearance of a second target or disappearance of an existing target is only 10% at each time step. Each particle will have its own model order m_k.

The state equation for indices $\mathbf{x}_k$ of order m_k include uncorrelated Gaussian noise $\mathbf{v}_k$ with $Q_k = 25$:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \text{round}(\mathbf{v}_k). \tag{31}$$

The measurement equation for observations $\mathbf{y}_k$ includes additive, Gaussian noise $\mathbf{w}_k$ with $\mathbf{R}_k = 0.09\mathbf{I}$ and

$$\mathbf{h}_k(\mathbf{x}_k) = \sum_{l=1}^{m_k} \delta(\mathbf{x}_k(l)), \tag{32}$$

where $k = 1, \ldots, 30$, and the size of measurement vector $\mathbf{y}_k$ is 100. The $\delta(i)$ vector has the value of one at index i and zero otherwise. Model order m_k determines how many targets are present. We assume that $\mathbf{x}_k(1)$ and $\mathbf{x}_k(2)$ represent locations of two targets at step k. The target locations are functions of step k: $\mathbf{x}_k(1) = 2k$ and $\mathbf{x}_k(2) = 3k$. When both targets are present, $\mathbf{y}_k = \delta(2k) + \delta(3k)$.

True state vector values are shown in Fig. 6(a) (solid lines). To estimate state parameters, particles for $\mathbf{x}_1(1)$ and $\mathbf{x}_1(2)$ at $k = 1$ are drawn from a uniform distribution between 1 and 100. The model order for each particle is initialized via a uniform prior allowing discrete values 1 and 2. Particles are propagated through the state equation, are updated via the measurement equation and resampled, and are then perturbed to predict the following step. Circles in Fig. 6(a) illustrate the modes of the posterior PDFs for state variables. The first target is tracked accurately with small deviations from true values because of noise $\mathbf{w}_k$. The second target is missed at $k = 11$ and 12 but is subsequently detected and correctly localized.

Figs. 6(b–e) demonstrate the posterior PDFs of model order m_k at four consecutive steps. At $k = 10$, only one target is present and the model order PDF clearly supports that. A second target enters at $k = 11$ but the PDF of the model order remains concentrated at 1 and the second target is missed. At $k = 12$ the system still favors one target, however, significant probability appears at $m_{12} = 2$. At $k = 13$, the presence of two targets is favored: $m_{13} = 2$. The system learns the presence of the second target slowly because of the low probability of transition between models and the posterior PDF from $k - 1$. Changing T_m alters filter behavior. A T_m with a larger transition probability would be more sensitive to sudden changes, capturing the change at $k = 11$ but the location estimates would have larger RMS errors since more particles would be allowed to "jump".

V. Practical Issues

Sequential MC techniques provide a powerful framework for performing signal processing in non-stationary dynamic systems involving nonlinear equations and non-Gaussian PDFs. Implementation of such methods requires a careful selection of a suitable and computationally efficient filter, noise in state and measurement equations, and the number of particles when a particle filtering approach is the best choice. Some of the practical issues arising in state-space estimation are discussed below.

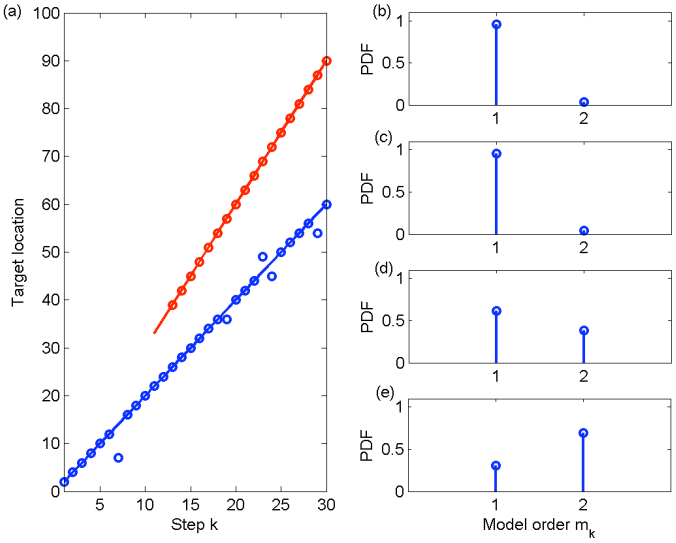

Fig. 6. (a) Target trajectories with a varying model order: true (solid lines) and estimated with a PF (○). Posterior PDF of model order m_k for (b) $k = 10$, (c) $k = 11$, (d) $k = 12$, and (e) $k = 13$.

A. Expressing the problem in a dynamic state space format

Establishing the measurement equation Eq. (2) that relates data $\mathbf{y}_k$ to the state vector $\mathbf{x}_k$ and the state equation Eq. (1) that expresses the dynamic evolution of $\mathbf{x}_k$ is the first step in the implementation of a Bayesian filter. In ocean acoustics, $\mathbf{y}_k$ is often the acoustic field received at an array of hydrophones. Some examples of states $\mathbf{x}_k$ in ocean acoustics are source bearing, location, velocity and acceleration, ocean bathymetry and sound speed profile, geoacoustic parameters, parameters related to acoustic modes and dispersion, frequency, Doppler shift, and multipath arrival time structure along a receiving array.

B. Likelihood function

An essential part in implementing a sequential particle filter is to derive the likelihood used in Eqs. (6) and (29) from the measurement equation. The likelihood function for $\mathbf{x}_k$ is derived based on function $\mathbf{h}_k$ and statistical properties of noise $\mathbf{w}_k$. In ocean acoustic applications, data errors are often considered uncorrelated and Gaussian, giving analytically tractable likelihoods.

As an example, measurements $\mathbf{y}_k$ employed for source localization and geoacoustic inversion contain complex-valued hydrophone data across the array and several frequencies. For a single frequency, a simple measurement equation is of the form:

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{w}_k = s_k \mathbf{d}_k(\mathbf{x}_k) + \mathbf{w}_k \tag{33}$$

where s_k is the known or unknown complex-valued source amplitude and $\mathbf{d}_k(\mathbf{x}_k)$ is the acoustic propagation model. Assuming noise $\mathbf{w}_k$ to be $\mathcal{N}(0, \mathbf{R}_k)$, the likelihood function is of the form

$$l(\mathbf{x}_k) \propto \frac{1}{\|\mathbf{R}_k\|} \exp\left[-(\mathbf{y}_k - s_k\mathbf{d}_k(\mathbf{x}_k))^{\mathrm{H}}\mathbf{R}_k^{-1}(\mathbf{y}_k - s_k\mathbf{d}_k(\mathbf{x}_k))\right] \tag{34}$$

In geoacoustic tracking, the complex-valued source amplitude s_k is usually unknown. The PF likelihood formulation with an unknown source signal [33] follows closely that of the classical geoacoustic inversion likelihood obtained from the Bartlett power objective function [76]. In most cases, likelihoods derived for inversion can also be used in the sequential filtering.

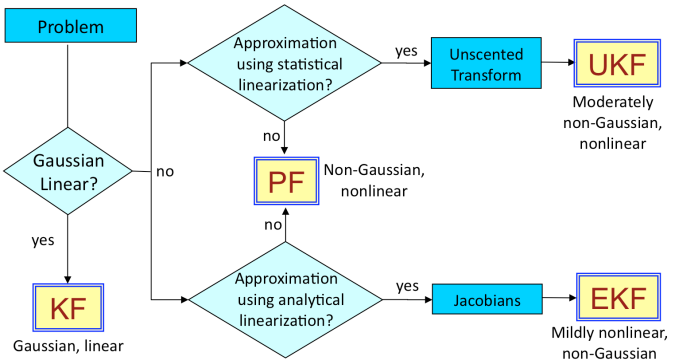

Fig. 7. A quick guide to filter selection.

C. State vector transitions

State noise $\mathbf{v}_k$ in Eq. (1) determines the error in the state vector evolution and is often chosen as additive and Gaussian. Furthermore, parameters in the state vector are usually assumed uncorrelated, *i.e.,* $\mathbf{Q}_k$ is diagonal.

Although state variables are not expected to vary in a way that is not consistent with the state evolution model [Eq. (1)], sudden changes can occur. The filters in such cases need a noise term $\mathbf{v}_k$ with a suitable covariance $\mathbf{Q}_k$ to continue tracking even when the state behaves in an unexpected way. For example, when bathymetry is tracked in a relatively flat ocean environment, the water column depth varies slowly with range. Hence, small perturbations of the state vector can capture these variations. The presence of a canyon, on the other hand, results in a radical bathymetry change. A small variance in the state noise may then cause the filtering process to diverge. Covariance matrix $\mathbf{Q}_k$ should contain values large enough to accommodate the unexpected changes, but at the same time small enough to prevent noisy estimates and poor performance. These trade-offs can make state noise selection a challenging practical problem. An example where state noise and the number of particles need to be carefully selected because of sharp changes in sediment thickness and sound speed is discussed in Section VI-B.

D. Filter selection

Once an estimation problem is defined with appropriate state and measurement equations, a suitable filter needs to be identified, see Fig. 7. For linear $\mathbf{h}$ and $\mathbf{f}$, additive Gaussian noise terms, the KF is the optimal sequential filter. When $\mathbf{h}$ and $\mathbf{f}$ are weakly nonlinear, EKF and UKF provide near-optimal solutions. The EKF has been a popular choice under such circumstances, especially when Jacobian computations of $\mathbf{h}$ and $\mathbf{f}$ are straightforward. However, the UKF is superior to the EFK because of higher accuracy in moment estimation and derivative-free implementation using comparable computational resources.

EnKFs are recommended for high dimensionality of the state vector. When the estimation problem is too complex to be handled via either analytical or statistical linearization, a PF approach is mandated. However, PFs require considerably more computational resources. PFs are not superior to KFs but rather complementary to them. All variants of Bayesian filters have their place and may perform better in terms of RMS error and divergence performance than the other filter types for particular sets of problems.

An example where a PF implementation is required is the problem of atmospheric environmental tracking [77]. The relationship between data and state vector is complex and requires a strongly nonlinear measurement equation with multiple non-additive, non-Gaussian noise terms. Signal $\mathbf{y}_k$ received at a naval radar is a function of the changing sea surface and $\mathbf{x}_k$ consists of the atmospheric conditions to be tracked. Data are related to $\mathbf{x}_k$ as follows:

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{w}_k) = \left\| \sqrt{\mathbf{w}_k(1)}\ P_e^2(\mathbf{x}_k) + \mathbf{w}_k(2) \right\|^2,$$

where $P_e(\mathbf{x}_k)$ is the field calculated using a split-step Fourier transform parabolic equation for inhomogeneous $\mathbf{x}_k$ and $\mathbf{w}_k(1)$ and $\mathbf{w}_k(2)$ represent a compound K-distributed PDF (a gamma PDF modulated by a Rayleigh) and a complex Gaussian PDF, respectively. Neither noise component is additive. Even after eliminating $\mathbf{w}_k(2)$ and reducing $\mathbf{w}_k(1)$ to an additive Gaussian term through a series of approximations, the PF significantly outperforms the EKF and UKF [29], [78]. This example illustrates a case where the additive Gaussian noise assumption in Section V-B is not suitable.

E. Number of PF particles

In a classical inversion, the whole search space needs to be explored. However, in sequential Bayesian filtering the state vector is assumed to follow the state equation. Thus, utilizing the information from the prior step and the known evolution of the states enables a focused search in the state space. Hence, the PF requires a smaller number of particles at each step than the number of samples necessary for a single inversion via, for example, Metropolis-Hastings sampling [14].

The essential number of particles for the filtering process to successfully estimate posterior $p(\mathbf{x}_k|\mathbf{Y}_k)$ is problem-dependent. There are four main factors that determine N_p:

- *Problem complexity:* Nonlinearities in the state and measurement equations, complexity of the underlying posterior PDF, errors in the state and measurement equations, and desired filter output impact the selection of N_p. It

takes fewer particles to track the median/MAP than the marginal PDF. Highly non-Gaussian, long tailed, peaked, multi-modal PDFs require a large number of particles. If the state $\mathbf{x}_k$ differs from the evolution in Eq. 1, the actual state will deviate from predicted $\mathbf{x}_{k|k-1}$. A larger N_p is then needed for tracking.

- *Accuracy:* Desired accuracy in track estimates is a major factor in selecting the number of particles. Increasing the number of particles initially provides a large performance improvement in a PF. It has been shown in [79] that the error is of the order of $\mathcal{O}(N_p^{-1})$. However, after a problem dependent N_p is reached, the performance stays relatively flat and increasing the computational cost provides only marginal benefits. An example in geoacoustic tracking is given in [29].
- *State dimension:* A major factor determining N_p is the state dimension n_x. MC integration is independent of the state dimension for N_p statistically independent particles. However, for sequential filtering the resampling stage creates multiple copies of high likelihood particles, making N_p tightly related to the number of state variables [80]. The problem is further complicated by the fact that the importance density in SIR determines convergence. A good sampling density results in N_p increasing linearly with n_x [81]. The required N_p increases exponentially with n_x when the importance density is poorly chosen (often referred to as 'the curse of dimensionality'). Special circumstances facilitate the reduction of the effective size of the state dimension without compromising accuracy. Rao-Blackwellized PFs [82] do just that and are briefly discussed in Appendix II. Note that the effects of a high state dimension on N_p are of importance for EnKF applications [54].
- *Computational cost:* An upper limit for N_p is determined by the maximum number of particles that can be processed with limited computational resources, which is important especially for real-time filters.

It is possible to have filters with adaptive N_p schemes [83]. Coupling PFs with error metrics, the sample size can be adapted depending on the error estimate. Such an implementation is particularly important during the initial steps, when there is only little prior information for $\mathbf{x}_0$. To address the significant uncertainty, a large number of particles can be sampled, with the cloud size decreasing in subsequent steps.

F. Computational cost

Computational cost is related to the overall CPU time and number of operations a filter has to perform at each step. KFs are generally computationally efficient. EnKFs and PFs can be computationally intensive. The cost is problem specific, depending in particular on the computational complexity of the state transition function $\mathbf{f}$, the propagation model $\mathbf{h}$, and the ability to sample from the noise PDFs. Following the example in Section V-D, obtaining a Gaussian sample is much simpler than obtaining a K-distributed sample. Similarly, the propagation model may require linear matrix multiplications or a complex, nonlinear parabolic equation acoustic propagation model.

The KF has the lowest possible computational cost. The computational complexity of KFs and their variations mainly lie in the inversion of matrices necessary for the calculation of the Kalman gain. The EKF needs only slightly more computation, depending on whether the Jacobians are calculated analytically or numerically. The UKF needs $2n_x + 1$ sigma points, and EnKFs and PFs use N_p particles. PFs are critically affected by the number of particles N_p and the implementation of resampling.

Three stages contribute to the computational cost of a SIR PF with N_p particles. The prediction stage uses $\mathcal{O}(N_p)$ computations to obtain $\mathbf{x}_{k|k-1}$. This is followed by $\mathcal{O}(N_p)$ likelihood and weight calculations during the update stage, which are followed by a resampling stage. The resampling algorithm described in the example of Section IV-C is called multinomial [57]. Less computationally expensive resampling algorithms that operate on the order of $\mathcal{O}(N_p)$ have been proposed such as residual resampling [84] and systematic resampling [40]. The systematic resampling shown in Fig. 5 is the commonly used $\mathcal{O}(N_p)$ resampling algorithm.

Overall, the computational cost of a PF is of the order of $\mathcal{O}(N_p)$. Although the complexity of filtering methods is expressed as a function of the number of particles drawn within each step, it is important to consider the computational cost imposed by the propagation model $\mathbf{h}$ in the measurement equation. In many ocean acoustic applications, the number of operations required for propagation model calculations is significant relative to the filtering. This should be taken into account when evaluating the computational efficiency of sequential filtering approaches. Selecting an algorithm that requires few propagation model calculations without affecting track quality is important.

VI. Ocean Acoustic Applications

Signal processing in ocean acoustics frequently involves complicated and rapidly changing ocean environments in coastal shallow water. The dynamics that characterize ocean acoustic applications are inherently nonlinear, non-Gaussian, and non-stationary processes that quickly vary with space and time. Sequential filtering is able to adapt to these environmental changes. The introduction of newer filters such as UKFs, EnKFs, and PFs that handle nonlinear/non-Gaussian problems, coupled with an increase in computational power, enabled a steady increase in the number of KF and PF approaches developed for ocean acoustic applications. This section is composed of four ocean acoustic examples where tracking filters have been used.

A. Target localization and tracking with filtering and ocean acoustics

Sequential filters have been applied extensively to target tracking [41]. There is a fair amount of literature about underwater AUV positioning, navigation, and target tracking [35], [85]. EKFs are frequently employed in target tracking in radar and sonar applications and used in all references cited below. These algorithms involve a state equation that describes the motion of the target in addition to ocean acoustic observations

contained in the measurement equation. The state vector may consist of target range, azimuth, (x,y,z) coordinates, velocity, and acceleration depending on the complexity of the target motion [35].

Due to the difficulty of incorporating the complex ocean physics in the estimation process, early work in tracking involved simple ocean models that progressively became more sophisticated. This evolution can be seen in the following series of papers. Culver and Hodgkiss [16] apply an EKF to track the position of an array of freely drifting floats. They use a constant velocity track model for the float positions. The measurements consist of observed travel times between the floats. The ocean acoustic model is responsible for separating the direct arrivals from the surface and bottom reflected arrivals. Instead of filtering out the reflected arrivals, El-Hawary and Mbamalu [17] create a more complete ocean model, relating the difference in travel times between direct and reflected arrivals to target range and depth, and solve the tracking problem using various techniques, including EKFs.

A unique technique is implemented by Candy and Sullivan [18], [22] in that the propagation model (wave equation) is integrated into both the state and measurement equations, as opposed to using acoustic propagation model just in the measurement equation, common to other ocean acoustic applications. This enables realistic ocean modeling by computing the acoustic field in a normal mode framework. In acoustics, a full propagation model is implemented by marching the field in depth (*e.g.*, a normal mode solution of the Helmholtz equation) or range (*e.g.* parabolic equation). In [18], [22], the state vector consists of the modal functions and their horizontal wavenumbers and the state equation marches in depth based on a measured sound speed profile. The measurement equation relates the measured complex-valued acoustic pressure on a vertical array to the vector (solutions of the wave equation). This method solves normal mode propagation as an EKF and enables passive source localization in range and depth. Hence, coupling a non-linear optimization algorithm with the EKF-based ocean acoustic model can solve the source localization problem in a complex ocean environment [21].

Bayesian filtering can track the bearing of a target. Sullivan *et al.* [25] improve target bearing tracking by applying an EKF to a broadband acoustic field measured on a short array towed by an autonomous submerged vehicle. The moving array causes a Doppler shift in the received signal frequency, which is a function of both bearing and source frequency. By casting the problem as a joint bearing-frequency tracking problem, they track bearing and the rate of change of bearing together with source frequency.

An alternative Bayesian tracking approach has been explored extensively [86], [87] for tracking a source in the presence of ocean and geoacoustic variability. Normally, tracking is performed sequentially using SMC. In this paper, the track is parameterized in terms of an analytic equation with unknown parameters (e.g. a straight line with unknown slope) and estimated using non-sequential MC methods. Environmental parameters are included in the inversion to obtain the full posterior PDF. One difference is that they are constant and hence are not tracked.

He *et al.* [26] track with an EKF a target in a horizontal x-y coordinate system illuminated with an active source and a 2-D planar receiver array. The state parameters are the x-y position and x-y velocity. The tracking stability is improved by utilizing the waveguide invariant [88], which is approximately one for shallow water propagation with reflecting boundaries. The proposed invariance constraint enforces consistency between frequency and range and smoothness in frequency variations with time along an identified striation. Adding the invariance constraint to the EKF improves tracking.

B. Environmental parameter estimation, geoacoustic inversion and tracking

The relation between the complex ocean environment and acoustic wave propagation involves nonlinear processes. An acoustic propagation model such as normal modes or the parabolic equation is needed depending on the problem. Due to complexity and nonlinearities, environmental sequential Bayesian inversion applications require filters that can handle such problems. Advances in filtering techniques using physics-based environmental inversion are shown in a series of papers discussed below.

Candy and Sullivan [19] give an early application of environmental Bayesian filtering, where an EKF-based normal mode code is used [18] to estimate the ocean sound speed profile (SSP) from noisy acoustic measurements at a single frequency. The EKF marches in depth, constructing the mode shapes together with the SSP as a function of depth. They then apply the algorithm to the Hudson Canyon experiment data [20] and generalize it for broadband processing [23]. These early applications use an EKF. More advanced KFs such as the UKF, EnKF, and PFs enable estimation and tracking in more complex environments varying temporally and/or spatially both in range and depth.

Carriere *et al.* [28] use the UKF to perform range-independent SSP prediction enabling the temporal tracking of the profile as shown in Fig. 8(a). The SSP is constructed using empirical orthogonal functions (EOF). The coefficients of each EOF are tracked in time using both an EKF and an UKF. The results confirm that a UKF in range-independent SSP tracking is superior to the EKF. However, the SSP in [28] is a range-independent profile. Temporally tracking a range-dependent SSP requires a much larger number of state parameters to estimate the evolution of the SSP in both depth and range.

To solve this problem, Carriere *et al.* [30] replace the UKF with an EnKF that is efficient for larger state spaces as discussed in Section III-C. They develop an EnKF processor for tracking the range-dependent sound speed structure from multi-frequency sound transmission between a point source and a vertical array. The ocean sound speed is parameterized in terms of range-dependent EOF coefficients. The acoustic measurement equation is augmented by sea-surface sound speed measurements. They demonstrate the approach from a 13-day experiment in the Ligurian Sea for a 20 km long transect in a water depth of about 100 m. The introduction of the EnKF decreases the track RMS error significantly and reduces track divergence. A 4D spatio-temporal environmental

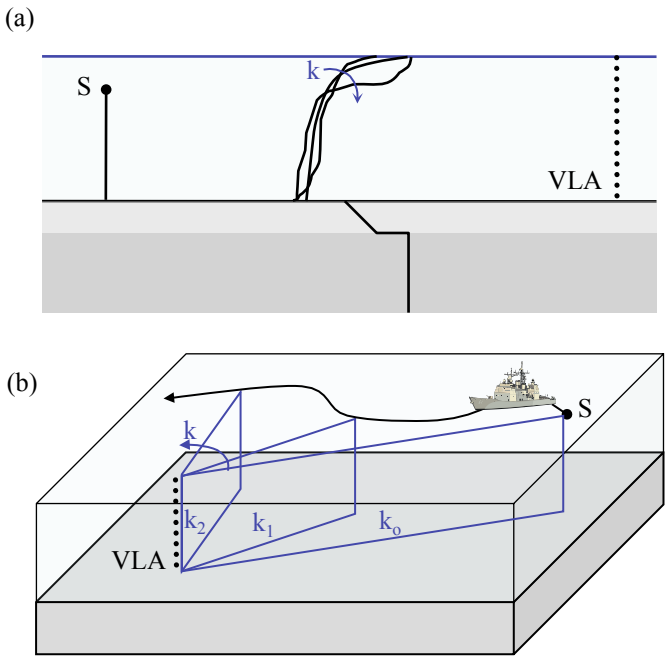

Fig. 8. Geoacoustic environmental tracking: (a) Temporal tracking of the ocean sound speed profile for a fixed-receiver and a fixed-source, (b) tracking of the changing environment between the receiver and a moving source. Here shown for a vertical line array (VLA) of receivers.

tracking of the SSP using full-field tomography is also given in [27] using the EKF.

Another application of Bayesian filtering is in geoacoustic inversion and tracking. Just like the water column SSP, bottom parameters such as sound SSPs, thickness, attenuation, and density of sedimentary layers affect acoustic propagation. Yardim *et al.* [29] show how to track these geoacoustic parameters using both vertical and horizontal array data as shown in Fig. 8(b). The paper also gives a detailed analysis of performance comparison between the EKF, UKF, and SIR PFs with changing particle numbers and it was shown the more computationally demanding PF performed best followed by UKF and then EKF.

Yardim *et al.* [33] demonstrate sequential geoacoustic tracking on real data and extend geoacoustic parameter tracking by including the unknown source parameters, effectively tracking the source in an unknown and changing environment. This requires inclusion of the source depth, range, and speed together with the geoacoustic parameters and a source motion model similar to that in target tracking filters. The algorithm is tested on the SWellEX-96 data [89] with bathymetry ranging from 100 to 250 m and a range-dependent sedimentary layer.

C. Frequency tracking

Time-frequency representations provide a wealth of information on the time-varying evolution of the frequency content of non-stationary signals. Short time Fourier transforms (STFT) and time varying auto-regressive (TVAR) methods have been frequently employed in e.g., audio processing [90] and speech processing [91].

PFs offer an excellent framework for instantaneous frequency extraction in such problems. Data measurements (time slices of the STFT or TVAR spectra) are related to the varying frequencies through a measurement equation. The state vector consists of peak frequencies that vary smoothly between consecutive time steps as given by the state equation. In addition to peak frequencies, the state vector includes the corresponding peak amplitudes and the model order (how many frequency components are present); see Section IV-E for the treatment of a varying or unknown model order. TVAR offers different choices for state-space formulation, with the state vector including either auto-regressive coefficients or poles and moduli. Poles, for example, represent frequencies of an observed signal, related to the measured data in a nonlinear fashion. The unknown number of frequencies also dictates a PF implementation rather than that of a KF or KF variant.

Zorych and Michalopoulou [31] track group velocity dispersion curves using PFs, utilizing a normal mode code for sound propagation. Each modal dispersion curve is treated as a 'moving target', whose 'location' (here frequency) evolves with time. Synthetic signals at receiving phones from a 20-km range in a 120-m waveguide are processed with a STFT, the magnitude of which is expressed as a weighted sum of Gaussian pulses representing the modes present in the signal. This parametric formulation provides the measurement equation for the PF framework. Frequencies and modal amplitudes are the state parameters.

Geoacoustic inversion can employ modal arrival times at specific frequencies as observed data. The PF method described above can enhance inversion by computing uncertainty on extracted arrival times, which can be propagated through the inversion process for the calculation of uncertainty in geoacoustic estimates [92]. Compiling particle information of a specific frequency in each mode provides a posterior PDF of arrival times for this frequency.

Instead of estimating wave dispersion using time-frequency representations, Candy and Chambers [24] assume a known dispersion relation in a time-varying waveguide. Under this assumption, their goal is to obtain accurate estimates of the propagating wave from noisy field measurements. This problem is addressed in a state-space framework. The modal wavenumber is the state variable and the state equation describes the wavenumber evolution in time between consecutive states. The measurement equation relates field measurements to wavenumber via the dispersion relation. State transition and measurement functions are nonlinear and an EKF is proposed, calculating Jacobians for the two functions. At each step, MAP wavenumber estimates are employed in the calculation of an enhanced acoustic field through the measurement equation. Wavenumber tracking and internal wave estimation of synthetic data demonstrates the approach.

Dispersion aside, frequency line tracking has been successfully implemented in other ocean acoustic problems as well. For example, KFs are employed in [93] for the extraction of striations from active sonar spectrograms. This striation identification is useful in verifying target tracks provided by active sonar measurements. Striations relate intensity, target range, and source frequency through the known waveguide

invariance [88]. KFs are designed that trace frequency lines in time using smoothness constraints on frequency evolution. The measurement equations relate measured frequency and intensity to the state vector, which also includes frequency and intensity as well as the difference in frequency between consecutive pings.

D. Spatial arrival time tracking

In ocean acoustics, multipath arrival times have been successfully used for source location, bathymetry estimation, and geoacoustic inversion [17], [94]. This approach is especially helpful for short-range, high frequency sound propagation in a shallow water waveguide. Such circumstances allow us to observe and quantify interactions of sound paths with the propagation medium and boundaries. The quality of source localization and environmental parameter inversion depends on the accuracy of arrival time estimates.

A PF was employed for time-delay estimation [95], where arrival times at two receivers were tracked in time for signals propagating in a time-varying environment. The measurement equation related measurements to arrival times and the state equation allowed limited and controlled variations between time states. In addition to evolution in time, sequential Bayesian filtering is employed to trace spatial variations of arrival times.

In [32] a PF extracted accurate travel times (state variables) from acoustic signals (measurements). Spatial rather than temporal sequential variability of arrivals was exploited. Following the process of [32], we applied the PF arrival time estimation approach to data from the Shallow Water 06 (SW06) experiment. The data were received on August 27, 2006, at the 16 element MPL-VLA1 vertical array, just before geoacoustic inversion was performed [96], [97]. A linear frequency modulated signal was transmitted with frequency content between 100 and 900 Hz; the sampling rate was 50 kHz. The source depth was 25 m and the range to VLA was 230 m. Received signals were match-filtered to produce the time series of Fig. 9(a). Only the 14 lower phones were used here, as the SNR at the top two phones was low due to large ocean waves.

The purpose is to estimate the arrival times of three paths at each phone (direct path and surface and bottom reflections), because accurate estimation is needed for inversion based on extracted path arrivals. The state variables for the PF are the arrival times for the three paths of Fig. 9(a). Similarly to the example of Section IV-E, both state and measurement equations include uncorrelated Gaussian noise, $\mathbf{v}_k$ and $\mathbf{w}_k$, respectively. The state equation is given in Eq. (31). The measurement equation is:

$$\mathbf{y}_k = \sum_{l=1}^{m_k} \alpha_l \, \mathrm{sinc}\left(2b\left[t - \mathbf{x}_k\left(l\right)\right]\right) + \mathbf{w}_k \tag{35}$$

The sinc pulse results from the autocorrelation of the linearly frequency modulated source waveform of bandwidth b, which was match-filtered similarly to the received time series.

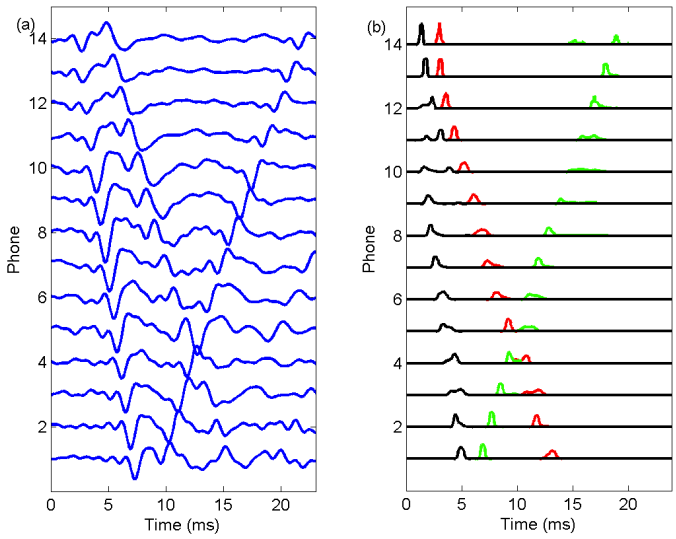

Fig. 9. (a) Observed time series at range 230 m from the VLA showing direct, surface- and bottom-reflected arrivals. (b) PDF of peak arrival time for each of these arrivals.

The likelihood function for the unknown arrival times at receiver k is:

$$p(\mathbf{y}|\mathbf{x}_k(1),\ldots,\mathbf{x}_k(m_k)) \propto \tag{36}$$
$$\exp(-\frac{1}{2\sigma^2}||\mathbf{y}_k - \sum_{l=1}^{m_k} \alpha_l \, \mathrm{sinc}\left(2b\left[t - \mathbf{x}_k\left(l\right)\right]\right)||^2).$$

Data covariance matrix $\mathbf{R}_k$ is assumed diagonal with all non-zero elements equal to σ^2. Quantities α_l are amplitudes of the arriving paths that are here assumed to be known. The sequential filtering is started from the top (phone 14). Initially, order m_k is assumed to be the same for all phones and equal to three.

Figure 9(b) demonstrates the PDFs of peak arrival times as calculated by the PF. The spread of the PDFs as well as the number of modes express the uncertainty in arrival time estimation. The third arrival is the weakest in terms of amplitude. As a result, the PDF for this arrival at the top phone is spread and bimodal. The uncertainty is reduced at lower phones, where prior information builds and sequentially propagates from one phone to the next.

We next expanded the state vector to include the order, namely the number of arrivals that are present in the time series. That is, m_k could be 1, 2, or 3. We used only a portion of the time series of Fig. 9(a), up to 14 ms, excluding the bottom echo from phones 14–9. At the beginning of the filtering process, the prior on the number of arrivals was uniform allowing values 1, 2, and 3; the transition matrix probability between model orders (numbers of arrivals) is given by

$$T_m = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.2 & 0.6 \end{pmatrix}. \tag{37}$$

PDFs of peak arrival times at each phone are shown in Figure 10(a). Two arrivals are detected at phones 14–8. The third arrival appears in the time series at receiver 8. The PF misses it there but captures it at phones 7–1.

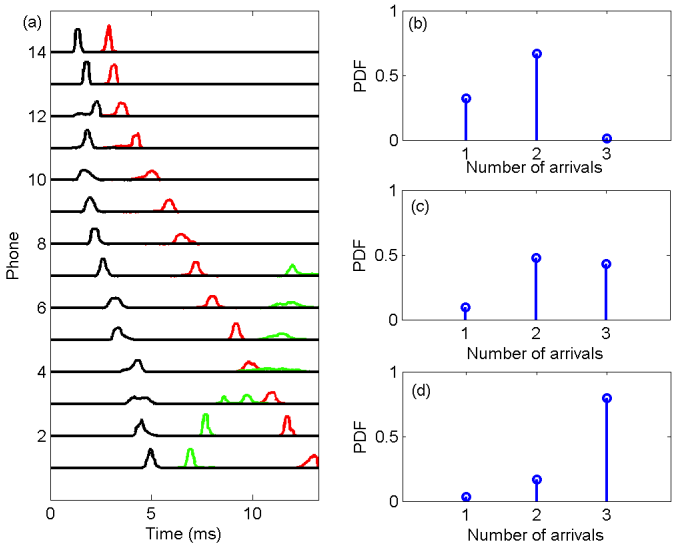

Fig. 10. (a) PDF of arrival times for two or three arrivals. Model order is varying with each hydrophone. (b) PDF of model order (number of arrivals) present at phones 9, 8, and 7.

Due to the sequential nature of the Bayesian estimation process, where the previous estimate influences the current, the approach is biased towards the prior solution. The transition from two to three arrivals in Figures 10(b, c, d) illustrates the predict-and-update process. The posterior PDF on the order at phone 9 exhibits a significant probability for the presence of two arrivals, which is the actual case. At phone 8, when the third arrival appears, its presence appears plausible, but lags behind the probability for two arrivals, because of the information that is propagated from the previous phone. At phone 7, the order is comfortably estimated to be three, with little probability corresponding to the order of two. That is, after two updating steps, the estimation process "learns" that an order jump has occurred.

Increased uncertainty is evident in the surface and bottom reflected paths near their crossing in the neighborhood of the 5th phone. The arrivals can still be well estimated as modes of the computed PDFs, with the direction of the paths being correctly identified after the crossing.

VII. Conclusions

The theory for sequential Bayesian filtering of acoustic signals is summarized focusing on classical Kalman filters (KF) and particle filters (PF). Filtering tracks the evolving state parameters as new data become available. It is based on two equations, state equation (transitions between states) and measurement equation (relation between data and states for one measurement).

KFs and their extensions are simple in their implementation and computationally efficient. PFs are computationally onerous but make fewer assumptions than the KF family, solve complex problems, and are still straightforward to implement. Hence, KFs and PFs complement each other enabling us to successfully address a wide spectrum of problems involving dynamical systems.

Ocean acoustic measurements in many applications are by nature sequential. Thus, the powerful sequential Bayesian filtering approaches can be applied to a broad range of acoustic problems. In the examples that were presented in this paper, we described how classical target tracking has evolved into the more complex problem of environmental tracking. Model selection is another important example as well as frequency component extraction from acoustic measurements. We showed how PFs can be successfully applied to SW06 data for the extraction of multipath arrival times.

Appendix I
Choice of Posterior PDF

It is important to identify which posterior PDFs of the unknown state variables $\mathbf{x}_k$ are calculated in problems such as the ones encountered in ocean acoustics, depending on the available data measurements $\mathbf{y}_k$ and limitations on complexity. There may be significant computational differences between these approaches.

1) $p(\mathbf{x}_k|\mathbf{y}_k)$; *inversion*. This is the marginal posterior PDF calculated in ocean acoustics where a state equation is not utilized to predict the subsequent steps. The estimate of $\mathbf{x}_k$ does not use information from previous or future measurements. It is computationally inefficient to perform each inversion independently with k. Further, the estimates can be inconsistent from step to step.
2) $p(\mathbf{X}_k|\mathbf{Y}_k)$; *full posterior*. The dimension of $\mathbf{X}_k$ can become large, as the dimension of $\mathbf{X}_k$ is the product of number of data measurements and states, making the estimation of the full posterior PDF practically impossible. An example of such an approach in ocean acoustics is given in [86], where $p(\mathbf{X}_K|\mathbf{Y}_K)$ has been computed for a small k ($k = 9$) in a non-sequential way.
3) $p(\mathbf{x}_k|\mathbf{Y}_k)$; *filter*: Bayesian filtering is the topic of this paper. Filtering enables all the previous and current measurements to be used in estimating $\mathbf{x}_k$ using "predict" and "update" steps.
4) $p(\mathbf{x}_k|\mathbf{Y}_{k+l})$, $l \geq 1$; *smoother*. A smoother is appropriate in applications where all data have already been observed and are readily available. Therefore, both past and "future" measurements can be exploited for the calculation of marginal PDF $p(\mathbf{x}_k|\mathbf{Y}_{k+l})$. Smoothing is more complex and computationally expensive than filtering. Although inclusion of future data improves estimation in comparison to a one-way filtering approach, the increase in computational cost sometimes makes smoothing less desirable than filtering.
5) $p(\mathbf{x}_{k+l}|\mathbf{Y}_k)$, $l \geq 1$; *predictor*. Ideal for applications where the acoustic measurements collected up to and including the current step are needed to predict some future value of state $\mathbf{x}_{k+l}$. The most common form where $l = 1$ is known as the one-step ahead predictor.

Appendix II
Advanced Particle Filters

A. Regularized particle filters

The SIR disadvantage associated with the resampling stage is a result of the discrete posterior PDF, designated by a

sum of delta functions with corresponding particle weights as given in Eq. (21). The resampling stage can only select from these discrete values, creating a large number of identical particles. An alternative is to replace the delta functions in Eq. (21) with sharp continuous kernel functions [63]. Since the PDF is now continuous, resampling results in non-identical particles in the high likelihood regions, regularizing/smoothing the particle distribution. This type of filter is called the regularized PF (RPF) and resolves the sample impoverishment problem. Kernel approximations are not appropriate for large dimensional state spaces and the introduction of a kernel violates the asymptotic convergence of the particle set to the underlying posterior PDF as the particle number increases.

B. Markov Chain Monte Carlo particle filters

The Markov chain MC PF (MCMC-PF) [60] is similar to the RPF. It is designed to solve the problem of sample impoverishment by inserting a MCMC sampler in the resampling stage. Such samplers are designed to draw samples from complex PDFs. The histogram of MCMC particles can be shown to converge asymptotically to the desired PDF [56], [98].

Similarly to the PRF, the MCMC kernel and sampling create the necessary sample diversity. Moreover, due to the nature of MCMC, the resulting set of particles still satisfies the asymptotic convergence as opposed to the RPF.

C. Auxiliary particle filter

The auxiliary PF (APF), or auxiliary SIR (ASIR), is introduced by Pitt and Shephard [64] to reduce sample impoverishment and prevent the blind prediction of SIR. This is implemented by delaying the resampling stage that would normally follow the update stage in classical SIR. Instead, the resampling is incorporated into the prediction stage of the next step. The APF introduces an auxiliary variable to state vector $\mathbf{x}_k$ that combines the resampling at $k-1$ with the prediction stage at k. Therefore, particles $\mathbf{x}_{k|k-1}$ at the prediction stage are conditioned on the current prior and the likelihood, resulting in a more efficient filter. APF may be adversely affected in the presence of large process noise [41].

D. Rao-Blackwellized particle filters

A main drawback of any PF algorithm is the large amount of particles needed to estimate accurately the posterior. Moreover, this already large number increases quickly as the state-space dimension rises, especially when the importance sampling density used in the PF is different from the true PDF. It is, therefore, desirable to reduce the number of unknown parameters that the PF needs to track. This might be possible in a large number of practical problems (target tracking in particular) [35]. In many cases, not all of the unknown parameters in the state space are nonlinear and non-Gaussian. These include target/source velocity and acceleration components. If some subset of the state vector can be modeled as linear/Gaussian conditioned on the rest of the parameters, it might be possible to create tracking filters that are both more efficient and accurate than a standard PF.

This is done through Rao-Blackwellization. The Rao-Blackwell theorem shows how to transform an arbitrary estimator into an estimator that is optimal relative to a selected criterion, in the mean squared error (MSE) sense [45]. In cases where conditional linear/Gaussian assumption holds, one can Rao-Blackwellize a standard PF. This process involves splitting the state space into two parameter sets: linear/Gaussian and nonlinear/non-Gaussian parameters. We can apply the optimal Kalman filter for the marginally linear/Gaussian parameters at each step, concentrating the powerful yet costly PF on the reduced, nonlinear/non-Gaussian portion of the state space. Since the Rao-Blackwellization procedure results in a mixture of Kalman and PF sections and reduces the PF state space via marginalization, the filter is known as Mixture PF, Marginalized PF, or Rao-Blackwellized PF (RBPF) [41]. A detailed analysis of computational cost and estimate quality of RBPF is given in [82].

E. Linearized particle filter

Another interesting variation is the linearized PF (LPF). Instead of using the transitional PDF as the IS density, each particle is treated as a KF and the Gaussian PDFs at the output of each KF is employed as the IS density for the predict stage of the LPF. This creates a bank of KFs, one for each particle and a significant increase in the computational cost for each particle. However, by running a KF on the particles of the previous step, the LPF incorporates both the prior and the current measurement in its predict stage similar to APF does. Since the algorithm uses various forms of KFs, each KF effectively performs a local linearization around its particle. Hence, it is also called the local linearization PF [41]. Successful implementations using both EKFs [59] and UKFs [66] show superior performance compared to SIR. Even though the cost for each particle calculation increases, the total number of particles needed to obtain good track performance decrease.

REFERENCES

[1] M. J. Hinich, "Maximum-likelihood signal processing for a vertical array," *J. Acoust. Soc. Am.*, vol. 54, no. 2, pp. 499–503, 1973.

[2] H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," *J. Acoust. Soc. Am.*, vol. 59, no. 2, pp. 368–373, 1976.

[3] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE J. Oceanic Eng.*, vol. 18, no. 4, pp. 401–424, oct 1993.

[4] A. Tolstoy, *Matched Field Processing for Underwater Acoustics*. Singapore: World Scientific, 1993.

[5] S. D. Rajan, J. F. Lynch, and G. V. Frisk, "Perturbative inversion methods for obtaining bottom geoacoustic parameters in shallow water," *J. Acoust. Soc. Am.*, vol. 82, no. 3, pp. 998–1017, 1987.

[6] M. D. Collins, W. A. Kuperman, and H. Schmidt, "Nonlinear inversion for ocean-bottom properties," *J Acoust. Soc. Am.*, vol. 92, no. 5, pp. 2770–2783, 1992.

[7] M. Siderius, P. L. Nielsen, and P. Gerstoft, "Range-dependent seabed characterization by inversion of acoustic data from a towed receiver array," *J Acoust. Soc. Am.*, vol. 112, no. 4, pp. 1523–1535, 2002.

[8] N. R. Chapman, S. Chin-Bing, D. King, and R. B. Evans, "Benchmarking geoacoustic inversion methods for range-dependent waveguides," *IEEE J. Oceanic Eng.*, vol. 28, no. 3, pp. 320–330, july 2003.

[9] R. A. Koch and D. P. Knobles, "Geoacoustic inversion with ships as sources," *J Acoust. Soc. Am.*, vol. 117, no. 2, pp. 626–637, 2005.

[10] A. M. Thode, G. L. D'Spain, and W. A. Kuperman, "Matched-field processing, geoacoustic inversion, and source signature recovery of blue whale vocalizations," *J. Acoust. Soc. Am.*, vol. 107, no. 3, pp. 1286–1300, 2000.

[11] A. M. Richardson and L. W. Nolte, "A posteriori probability source localization in an uncertain sound speed, deep ocean environment," *J. Acoust. Soc. Am.*, vol. 89, no. 5, pp. 2280–2284, MAY 1991.

[12] S. L. Tantum and L. W. Nolte, "Tracking and localizing a moving source in an uncertain shallow water environment," *J. Acoust. Soc. Am.*, vol. 103, no. 1, pp. 362–373, JAN 1998.

[13] P. Gerstoft and C. F. Mecklenbräuker, "Ocean acoustic inversion with estimation of *a posteriori* probability distributions," *J. Acoust. Soc. Am.*, vol. 104, no. 2, pp. 808–819, 1998.

[14] S. E. Dosso, "Quantifying uncertainty in geoacoustic inversion I. A fast Gibbs sampler approach," *J. Acoust. Soc. Am.*, vol. 111, no. 1, pp. 129–142, 2002.

[15] D. Battle, P. Gerstoft, W. S. Hodgkiss, W. A. Kuperman, and P. L. Nielsen, "Bayesian model selection applied to self–noise geoacoustic inversion," *J. Acoust. Soc. Am.*, vol. 116, pp. 2043–2056, 2004.

[16] R. L. Culver and W. S. Hodgkiss, "Comparison of Kalman and least squares filters for locating autonomous very low frequency acoustic sensors," *IEEE J. Oceanic Eng.*, vol. 13, no. 4, pp. 282–290, Oct 1988.

[17] F. El-Hawary and G. A. N. Mbamalu, "Underwater target tracking via the IRWLS filtering approach," *IEE Proc. F, Radar and Signal Processing*, pp. 459–469, 1991.

[18] J. V. Candy and E. J. Sullivan, "Ocean acoustic signal processing: A model-based approach," *J. Acoust. Soc. Am.*, vol. 92, no. 6, pp. 3185–3201, 1992.

[19] ——, "Sound velocity profile estimation: A system theoretic approach," *IEEE J. Oceanic Eng.*, vol. 18, no. 3, pp. 240–252, Jul 1993.

[20] ——, "Model-based processor design for a shallow water ocean acoustic experiment," *J. Acoust. Soc. Am.*, vol. 95, no. 4, pp. 2038–2051, 1994.

[21] ——, "Passive localization in ocean acoustics: A model-based approach," *J. Acoust. Soc. Am.*, vol. 98, no. 3, pp. 1455–1471, 1995.

[22] ——, "Model-based identification: an adaptive approach to ocean-acoustic processing," *IEEE J. Oceanic Eng.*, vol. 21, no. 3, pp. 273–289, Jul 1996.

[23] ——, "Broadband model-based processing for shallow ocean environments," *J. Acoust. Soc. Am.*, vol. 104, p. 275, 1998.

[24] J. V. Candy and D. H. Chambers, "Model-based dispersive wave processing: A recursive Bayesian solution," *J. Acoust. Soc. Am.*, vol. 105, no. 6, pp. 3364–3374, 1999.

[25] E. J. Sullivan, J. D. Holmes, W. M. Carey, and J. F. Lynch, "Broadband passive synthetic aperture: Experimental results," *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. EL49–EL54, 2006.

[26] C. He, J. E. Quijano, and L. M. Zurk, "Enhanced Kalman filter algorithm using the invariance principle," *IEEE J. Oceanic Eng.*, vol. 34, no. 4, pp. 575–585, october 2009.

[27] O. Carrière, J.-P. Hermand, J.-C. Le Gac, and M. Rixen, "Full-field tomography and Kalman tracking of the range-dependent sound speed field in a coastal water environment," *J. Mar. Syst.*, vol. 78, no. 4, pp. S382–S392, 2009.

[28] O. Carrière, J. P. Hermand, M. Meyer, and J. V. Candy, "Dynamic estimation of the sound-speed profile from broadband acoustic measurements," *Oceans*, 2007.

[29] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Tracking of geoacoustic parameters using Kalman and particle filters," *J. Acoust. Soc. Am.*, vol. 125, no. 2, pp. 746–760, 2009.

[30] O. Carrière, J.-P. Hermand, and J. V. Candy, "Inversion for time-evolving sound-speed field in a shallow ocean by ensemble Kalman filtering," *IEEE J. Oceanic Eng.*, vol. 34, no. 4, pp. 586–602, Oct. 2009.

[31] I. Zorych and Z.-H. Michalopoulou, "Particle filtering for dispersion curve tracking in ocean acoustics," *J. Acoust. Soc. Am.*, vol. 124, no. 2, pp. EL45–EL50, 2008.

[32] R. Jain and Z.-H. Michalopoulou, "A particle filtering approach for multipath arrival time estimation from acoustic time series," *J. Acoust. Soc. Am.*, vol. 126, no. 4, pp. 2249–2249, 2009.

[33] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Geoacoustic and source tracking using particle filtering: Experimental results," *J. Acoust. Soc. Am.*, vol. 128, no. 1, pp. 75–87, 2010.

[34] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, pp. 174–188, 2002.

[35] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P. J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Processing*, vol. 50, pp. 425–437, 2002.

[36] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, 2003.

[37] O. Cappé, S. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. IEEE*, vol. 95, no. 5, pp. 899–924, 2007.

[38] J. V. Candy, "Bootstrap particle filtering," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 73–85, 2007.

[39] A. Doucet and A. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," Department of Computer Science, University of British Columbia, Tech. Rep., 2008.

[40] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York: Springer, 2001.

[41] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Boston, MA: Artech House, 2004.

[42] O. Cappe, E. Moulines, and T. Ryd én, *Inference in Hidden Markov Models*. Springer, 2005.

[43] J. V. Candy, *Bayesian Signal Processing: Classical, Modern and Particle Filtering Methods*. New Jersey: John Wiley & Sons, 2009.

[44] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Estimation of radio refractivity from radar clutter using Bayesian Monte Carlo analysis," *IEEE Trans. Antennas Propagat.*, vol. 54, no. 4, pp. 1318–1327, 2006, doi:10.1109/TAP.2006.872673.

[45] S. M. Kay, *Fundamentals of Statistical Signal Processing – Volume I: Estimation Theory*. New Jersey: Prentice-Hall, 1993.

[46] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[47] H. L. van Trees, *Detection, Estimation and Modulation Theory*. New York: John Wiley & Sons, 1968.

[48] H. Cox, "On the estimation of state variables and parameters for noisy dynamic systems," *IEEE Trans. Automat. Contr.*, vol. 9, no. 1, pp. 5–12, 1964.

[49] S. Julier, J. Uhlmann, and H. F. Durrant-White, "A new method for nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 477–482, 2000.

[50] E. A. Wan and R. van der Merve, "The unscented Kalman filter", in S. Haykin, *Kalman Filtering and Neural Networks*. New York: John Wiley & Sons, 2001.

[51] J. T. Ambadan and Y. Tang, "Sigma-point Kalman filter data assimilation methods for strongly nonlinear systems," *Journal of the Atmospheric Sciences*, vol. 66, no. 2, pp. 261–285, 2009.

[52] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, pp. 10–10, 1994.

[53] P. L. Houtekamer and H. L. Mitchell, "Ensemble Kalman filtering," *Quarterly Journal of the Royal Meteorological Society*, vol. 131, no. 613, pp. 3269–3290, 2005.

[54] P. J. van Leeuwen, "Particle filtering in geophysical systems," *Monthly Weather Review*, vol. 137, pp. 4089–4114, 2009.

[55] G. Evensen, "The ensemble Kalman filter: Theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, no. 4, pp. 343–367, 2003.

[56] J. J. K. Ó Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*. New York: Springer–Verlag, 1996.

[57] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc. F, Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[58] A. F. M. Smith and A. E. Gelfand, "Bayesian statistics without tears: A sampling-resampling perspective," *American Statistician*, vol. 46, pp. 84–88, 1992.

[59] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[60] W. R. Gilks and C. Berzuini, "Following a moving target—Monte Carlo inference for dynamic Bayesian models," *J. R. Stat. Soc. B*, vol. 63, pp. 127–146, 2001.

[61] A. Kong, J. S. Liu, and W. H. Wong, "Sequential imputations and Bayesian missing data problems," *J. Amer. Statist. Assoc.*, vol. 89, no. 425, pp. 278–288, 1994.

[62] M. Bolić, P. M. Djurić, and S. Hong, "Resampling algorithms for particle filters: A computational complexity perspective," *EURASIP J. Appl. Signal Process.*, pp. 2267–2277, 2004.

[63] C. Musso, N. Oudjane, and F. LeGland, "Improving regularised particle filters" , in A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New Jersey: Springer, 2001.

[64] M. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filter," *J. Acoust. Soc. Am.*, vol. 94, pp. 590–599, 1999.

[65] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Statist. Assoc.*, pp. 1032–1044, 1998.

[66] R. Van der Merwe, A. Doucet, N. De Freitas, and E. Wan, "The unscented particle filter," *Advances in Neural Information Processing Systems*, pp. 584–590, 2001.

[67] P. Gerstoft and D. F. Gingras, "Parameter estimation using multifrequency range-dependent acoustic data in shallow water," *J. Acoust. Soc. Am.*, vol. 99, no. 5, pp. 2839–2850, 1996.

[68] C. F. Mecklenbräuker, P. Gerstoft, J. F. Böhme, and P.-J. Chung, "Hypothesis testing for geoacoustic environmental models using likelihood ratio," *J. Acoust. Soc. Am.*, vol. 105, no. 3, pp. 1738–1748, 1999.

[69] Z.-H. Michalopoulou and M. Picarelli, "Gibbs sampling for time-delay- and amplitude estimation in underwater acoustics," *J Acoust. Soc. Am.*, vol. 117, no. 2, pp. 799–808, 2005.

[70] J. Dettmer, S. E. Dosso, and C. Holland, "Model selection and Bayesian inference for high-resolution seabed reflection inversion," *J. Acoust. Soc. Am.*, vol. 125, no. 2, pp. 706–716, 2009.

[71] J. Dettmer, C. W. Holland, and S. E. Dosso, "Analyzing lateral seabed variability with Bayesian inference of seabed reflection data," *The Journal of the Acoustical Society of America*, vol. 126, no. 1, pp. 56–69, 2009. [Online]. Available: http://link.aip.org/link/?JAS/126/56/1

[72] Z.-H. Michalopoulou, "Multiple source localization using a maximum a posteriori Gibbs sampling approach," *J Acoust. Soc. Am.*, vol. 120, no. 5, pp. 2627–2634, 2006.

[73] J. Dettmer, S. E. Dosso, and C. W. Holland, "Trans-dimensional geoacoustic inversion," *J Acoust. Soc. Am.*, vol. 13X, pp. X–X, 2011.

[74] J. R. Larocque, J. P. Reilly, and W. Ng, "Particle filters for tracking an unknown number of sources," *IEEE Trans. Signal Processing*, vol. 50 (12), pp. 2926–2937, 2002.

[75] J. Vermaak, S. J. Godsill, and P. Perez, "Monte Carlo filtering for multi-target tracking and data association," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, pp. 309–332, 2005.

[76] C. F. Mecklenbräuker and P. Gerstoft, "Objective functions for ocean acoustic inversion derived by likelihood methods," *J. Comp. Acoust.*, vol. 8, pp. 259–270, 2000.

[77] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Sensitivity analysis and performance estimation of refractivity from clutter techniques," *Radio Science*, vol. 44, p. RS1008, 2009.

[78] ——, "Tracking refractivity from clutter using Kalman and particle filters," *IEEE Trans. Antennas Propagat.*, vol. 56, no. 4, pp. 1058–1070, 2008.

[79] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practitioners," *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 736–746, 2002.

[80] T. Bengtsson, P. Bickel, and B. Li, "Curse-of-dimensionality revisited: Collapse of the particle filter in very large scale systems," *IMS Collections: Probability and Statistics*, vol. 2, pp. 316–334, 2008.

[81] F. Daum and J. Huang, "Curse of dimensionality and particle filters," in *Proc. IEEE Aerospace Conf.*, 2003.

[82] R. Karlsson, T. Schön, and F. Gustafsson, "Complexity analysis of the marginalized particle filter," *IEEE Trans. Signal Processing*, vol. 53, no. 11, pp. 4408–4411, 2005.

[83] D. Fox, "Adapting the sample size in particle filters through KLD-sampling," *The International Journal of Robotics Research*, vol. 22, no. 12, pp. 985–1003, 2003.

[84] J. S. Liu, R. Chen, and T. Logvinenko, "A theoretical framework for sequential importance sampling and resampling" , in A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New Jersey: Springer, 2001.

[85] L. Stutters, H. Liu, C. Tiltman, and D. J. Brown, "Navigation technologies for autonomous underwater vehicles," *IEEE Trans. Sys., Man, and Cybern.*, vol. 38, no. 4, pp. 581–589, 2008.

[86] S. E. Dosso and M. J. Wilmut, "Uncertainty estimation in simultaneous bayesian tracking and environmental inversion," *J Acoust. Soc. Am.*, vol. 124, no. 1, pp. 82–97, 2008.

[87] D. Tollefsen and S. E. Dosso, "Three-dimensional source tracking in an uncertain environment," *J Acoust. Soc. Am.*, vol. 125, no. 5, pp. 2909–2917, 2009.

[88] G. L. D'Spain and W. A. Kuperman, "Application of waveguide invariants to analysis of spectrograms from shallow water environments that vary in range and azimuth," *J. Acoust. Soc. Am.*, vol. 106, no. 5, pp. 2454–2468, 1999.

[89] N. O. Booth, A. T. Abawi, P. W. Schey, and W. S. Hodgkiss, "Detectability of low-level broad-band signals using adaptive matched-field processing with vertical aperture arrays," *IEEE J. Oceanic Eng.*, vol. 25, no. 3, pp. 296–313, 2000.

[90] C. Dubois and M. Davy, "Joint detection and tracking of time-varying harmonic components: A flexible Bayesian approach," *IEEE Trans. Audio. Speech, Lang. Processing*, vol. 15, no. 4, pp. 1283–1295, may 2007.

[91] K. Kalgaonkar and M. Clements, "Vocal tract area based formant tracking using particle filter," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, April 2008, pp. 3405–3408.

[92] C.-F. Huang, P. Gerstoft, and W. S. Hodgkiss, "Validation of statistical estimation of transmission loss in the presence of geoacoustic inversion uncertainty," *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. 1932–1941, 2006.

[93] T. G. Oesterlein, C. He, J. E. Quijano, R. L. Campbell Jr., L. M. Zurk, and M. Siderius, "Extraction of time-frequency target features," in *IEEE Conference on Signals, Systems, and Computers*, November 2009.

[94] P. Pignot and N. R. Chapman, "Tomographic inversion of geoacoustic properties in a range-dependent shallow-water environment," *J. Acoust. Soc. Am.*, vol. 110, pp. 1338–1348, 2001.

[95] E. A. Lehmann, "Particle filtering approach to adaptive time-delay estimation," in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, 2006, pp. 1129–1132.

[96] C.-F. Huang, P. Gerstoft, and W. Hodgkiss, "Effect of ocean sound speed uncertainty on matched-field geoacoustic inversion," *J Acoust. Soc. Am.*, vol. 124, no. 4, pp. EL162–EL168, 2008.

[97] Y.-M. Jiang and N. R. Chapman, "Bayesian geoacoustic inversion in a range dependent shallow water environmen," *J Acoust. Soc. Am.*, vol. 124, no. 4, pp. EL155–EL161, 2008.

[98] C. Yardim, P. Gerstoft, and W. S. Hodgkiss, "Statistical maritime radar duct estimation using a hybrid genetic algorithm–Markov chain Monte Carlo method," *Radio Science*, vol. 42, p. RS3014, 2007.

Caglar Yardim (S'98-M'07) received the B.S. and the M.S. degrees in electrical engineering from the Middle East Technical University, Ankara, Turkey in 1998 and 2001, respectively, and the Ph.D. degree in electrical engineering from the University of California, San Diego (UCSD) in 2007. Since 2010 he has been a Scientist at the Marine Physical Laboratory, UCSD. His research interests include signal processing, propagation, optimization, modeling, and inversion of electromagnetic and acoustic signals.

Dr. Yardim was the recipient of the Best Student Paper Award at the 2007 IEEE Radar Conference and URSI Young Scientist Award in 2008. He is a member of IEEE, Acoustical Society of America.

Zoi-Heleni Michalopoulou received the Diploma in Electrical Engineering from the National Technical University of Athens in 1988, and an MS and Ph.D. from Duke University in 1990 and 1993, respectively, both in Electrical Engineering. She is currently a Professor of Mathematical Sciences at the New Jersey Institute of Technology. She is a Senior Member of IEEE and a Fellow of the Acoustical Society of America. Her research interests include underwater acoustics, Bayesian modeling, inverse problems, array signal processing, and spectroscopy.

Peter Gerstoft received the Ph.D. from the Technical University of Denmark, Lyngby, Denmark, in 1986. From 1987-1992 he was with Ødegaard and Danneskiold-Samsøe, Copenhagen, Denmark. From 1992-1997 he was at Nato Undersea Research Centre, La Spezia, Italy. Since 1997, he has been with the Marine Physical Laboratory, University of California, San Diego. His research interests include modeling and inversion of acoustic, elastic and electromagnetic signals.

Dr. Gerstoft is a Fellow of Acoustical Society of America and an elected member of the International Union of Radio Science, Commission F.