A MONTE CARLO TECHNIQUE FOR LARGE-SCALE DYNAMIC TOMOGRAPHY

Mark D. Butala[1], *Richard A. Frazin*[1], *Yuguo Chen*[2], *and Farzad Kamalabadi*[1]

1. Department of Electrical and Computer Engineering and Coordinated Science Laboratory
2. Department of Statistics
University of Illinois at Urbana-Champaign

ABSTRACT

We address the reconstruction of a physically evolving unknown from tomographic measurements by formulating it as a state estimation problem. The approach presented in this paper is the localized ensemble Kalman filter (LEnKF); a Monte Carlo state estimation procedure that is computationally tractable when the state dimension is large. We establish the conditions under which the LEnKF is equivalent to the Gaussian particle filter. The performance of the LEnKF is evaluated in a numerical example and is shown to give state estimates of almost equal quality as the optimal Kalman filter but at a 95% reduction in computation.

Index Terms— Kalman filtering, multidimensional signal processing, recursive estimation, remote sensing

1. INTRODUCTION

The problem of estimating properties of a hidden Markov process is encountered in many applications, including radar tracking of multiple targets [1], time-dependent tomography and interferometric imaging [2], geophysical data assimilation [3], and economic forecasting [4]. These problems can be expressed in state-space form and many recursive methods have been developed in the Bayesian framework to solve the resultant state estimation problem. For certain applications such as medical tomography, the state variation during the measurement interval may be small enough so that the state can be regarded as static. Techniques have also been developed for medical tomography when the state exhibits significant quasi-periodic [5] or rigid-body motion [6]. This paper is concerned with the problem of reconstructing a time varying object from tomographic (line integral) measurements when the state evolution is significant and governed by a complicated physical model, a problem often encountered in remote sensing. In such applications, the dimension of the state can be huge (the state will have over 2 million elements when representing a volumetric image on a $128 \times 128 \times 128$ grid).

The dynamics of a complicated physical system are often described by a nonlinear state-space model. Minimum mean-square error (MMSE) state estimates can be found from the mean of the posterior distribution of the state given the measurements, a difficult problem when either the dimension of the state is large (resulting in the evaluation of a complicated integral) or an analytic expression for the posterior distribution function is unavailable (which is most often the case in practice). The particle filter (PF) circumvents these pitfalls by approximating the posterior distribution with a set of weighted samples (particles) drawn from a simpler auxiliary distribution function, a technique called importance sampling [7]. Then, the (potentially) complicated integral is reduced to a summation over the particles that converges to the MMSE state estimate as the number of particles is increased. Because of problems with the rate of convergence resulting from degeneracy of the particles [7], the PF has been limited to applications where the dimension of the state is small [8], and, even in this case, the PF can be computationally expensive. If the posterior distribution is approximated by a single Gaussian, then the state estimation problem can be solved by the less complex Gaussian particle filter (GPF) [9]. However, the GPF must store and manipulate the full state covariance matrix (for the 128^3 element grid example, about 8 TB of storage is required if the elements are single precision floating point numbers) and also suffers from convergence issues associated with all Monte Carlo integration techniques.

More computationally efficient methods are possible when the state-space model is linear (or if the nonlinear model can be linearized) and the state and measurements are jointly Gaussian. In this case, the Kalman filter (KF) is optimal and an analytic expression exists for propagating the posterior of the state from one time index to the next. Unfortunately, as is the case for the GPF, the KF cannot be applied to problems with large state dimension because it requires the explicit formation of the state covariance matrix. Nevertheless, under the linear Gaussian model, the posterior distribution can be sampled directly and a recursive update on the samples exists given the closed form expression for the propagation of the posterior. The ensemble Kalman filter (EnKF) exploits these properties to form state estimates that approach those given by the (optimal) KF as the number of samples is increased without requiring the formation of the state covariance [3]. The main drawback of the EnKF is that it has similar convergence issues as the PF and GPF. However, for certain applications (as demonstrated in this paper), one can mitigate these issues and obtain further computational simplifications by restricting state estimate updates to a neighborhood about the physical location of the measurements [10].

In this paper, we first establish the conditions that result in the equivalence of the GPF and EnKF. Next, we address the problems associated with the EnKF and large state dimension by incorporating localization. Then, we present a numerical example that demonstrates the excellent performance of the localized ensemble Kalman filter (LEnKF) in estimating a physically evolving state from tomographic measurements.

2. THE GAUSSIAN PARTICLE FILTER

The nonlinear state-space dynamic signal model is defined as:

$$\begin{aligned} \boldsymbol{x}_{i+1} &= f_i(\boldsymbol{x}_i, \boldsymbol{u}_i) \\ \boldsymbol{y}_i &= h_i(\boldsymbol{x}_i, \boldsymbol{v}_i), \end{aligned} \tag{1}$$

where the vector $\boldsymbol{x}_i \in \mathbb{R}^N$ is the state at time index i with $1 \leq i \leq T$, $f_i(\cdot)$ is the (known) state transition function, $\boldsymbol{u}_i$ is the state process noise that accounts for any mismodeling of the state dynamics, $\boldsymbol{y}_i \in \mathbb{R}^M$ is the vector of measurements, $h_i(\cdot)$ is the (known) measurement function, and $\boldsymbol{v}_i$ is the measurement noise process that accounts for measurement uncertainty. The statistical properties of both noise processes are known. In the subsequent development, we define $\boldsymbol{y}_{1:i}$ to represent the set of measurements $\{\boldsymbol{y}_k\}_{k=1}^{i}$ and $\boldsymbol{x}_{i|j}$ as shorthand for either a random vector drawn from the distribution function $p(\boldsymbol{x}_i|\boldsymbol{y}_{1:j})$ or a sample of the importance sampling distribution $\pi(\boldsymbol{x}_i|\boldsymbol{y}_{1:j})$, depending on context.

The GPF [9] solves the state estimation problem under signal model (1) and is summarized in Fig. 1. The following notation is used: $\mathcal{N}(\boldsymbol{a}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the value of the mean $\boldsymbol{\mu}$ covariance $\boldsymbol{\Sigma}$ Gaussian distribution function evaluated at $\boldsymbol{a}$, $\pi(\cdot)$ represents the (application dependent) importance sampling distribution [7] and $\pi(\boldsymbol{x}_{i|i}^l|\boldsymbol{y}_{i:i})$ is the value of the importance sampling distribution evaluated at $\boldsymbol{x}_{i|i}^l$, $\boldsymbol{\mu}_{i|i}$ and $\boldsymbol{\mu}_{i+1|i}$ ($\boldsymbol{\Sigma}_{i|i}$ and $\boldsymbol{\Sigma}_{i+1|i}$) are estimates of the posterior mean (covariance) of $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i+1}$ given $\boldsymbol{y}_{1:i}$, L is the number of particles, and $(\cdot)^T$ indicates matrix transposition.

The GPF has many desirable properties [9]. However, the algorithm becomes computationally intractable as the state dimension N increases because the state covariance, e.g., $\boldsymbol{\Sigma}_{i|i}$, must be explicitly stored and manipulated in addition to the problem that L may have to increase at an exponential rate relative to N to obtain satisfactory state estimates [8].

3. THE ENSEMBLE KALMAN FILTER

The linear state-space dynamic signal model is:

$$\begin{aligned} \boldsymbol{x}_{i+1} &= \boldsymbol{F}_i \boldsymbol{x}_i + \boldsymbol{u}_i \\ \boldsymbol{y}_i &= \boldsymbol{H}_i \boldsymbol{x}_i + \boldsymbol{v}_i, \end{aligned} \tag{2}$$

Initialize: Assume $\boldsymbol{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_{1|0}, \boldsymbol{\Sigma}_{1|0})$
for $i = 1$ **to** T **do**
(Measurement update)
Draw samples: $\boldsymbol{x}_{i|i}^l \overset{\text{iid}}{\sim} \pi(\boldsymbol{x}_i|\boldsymbol{y}_{1:i}),\ 1 \leq l \leq L$
Compute weights:
$\overline{w}_i^l = p(\boldsymbol{y}_i|\boldsymbol{x}_{i|i}^l)\mathcal{N}(\boldsymbol{x}_{i|i}^l; \boldsymbol{\mu}_{i|i-1}, \boldsymbol{\Sigma}_{i|i-1})/\pi(\boldsymbol{x}_{i|i}^l|\boldsymbol{y}_{1:i})$
Normalize weights: $w_i^l = \overline{w}_i^l / \sum_{l=1}^L \overline{w}_i^l$
Filtered estimate: $\boldsymbol{\mu}_{i|i} = \sum_{l=1}^L w_i^l \boldsymbol{x}_{i|i}^l$
$\boldsymbol{\Sigma}_{i|i} = \sum_{l=1}^L w_i^l (\boldsymbol{x}_{i|i}^l - \boldsymbol{\mu}_{i|i})(\boldsymbol{x}_{i|i}^l - \boldsymbol{\mu}_{i|i})^T$
(Time update)
Draw samples: $\boldsymbol{x}_{i+1|i}^l \overset{\text{iid}}{\sim} p(\boldsymbol{x}_{i+1}|\boldsymbol{x}_i = \boldsymbol{x}_{i|i}^l)$
Predicted estimate: $\boldsymbol{\mu}_{i+1|i} = \sum_{l=1}^L w_i^l \boldsymbol{x}_{i+1|i}^l$
$\boldsymbol{\Sigma}_{i+1|i} = \sum_{l=1}^L w_i^l \boldsymbol{d}_i^l (\boldsymbol{d}_i^l)^T$, $\boldsymbol{d}_i^l = \boldsymbol{x}_{i+1|i}^l - \boldsymbol{\mu}_{i+1|i}$
end

Fig. 1. The GPF algorithm.

where the (known) functions $f_i(\cdot)$ and $h_i(\cdot)$ in (1) have been replaced by the (known) state transition and observation matrices $\boldsymbol{F}_i$ and $\boldsymbol{H}_i$, and the mutually independent state and measurement noise processes are zero mean and have (known) covariances $\boldsymbol{Q}_i \equiv \mathrm{E}\left[\boldsymbol{u}_i \boldsymbol{u}_i^T\right]$ and $\boldsymbol{R}_i \equiv \mathrm{E}\left[\boldsymbol{v}_i \boldsymbol{v}_i^T\right]$.

If $\boldsymbol{x}_1$, $\boldsymbol{u}_{1:T}$, and $\boldsymbol{v}_{1:T}$ are jointly Gaussian then the optimal solution under (2) is the KF and closed-form, recursive expressions exist for updating the mean and covariance of the state from one time step to the next (a complete description of the state because it is, in this case, a Gauss-Markov process). The EnKF, a Monte Carlo approximation to the KF, can also solve this state estimation problem [3] and is summarized in Fig. 2. This form of the algorithm is sequential, meaning the measurements are processed one at a time, and is applicable for the case when the measurement noise is uncorrelated, i.e., when $\boldsymbol{R}_i$ is a diagonal matrix (though a whitening filter can always be applied to the measurements to ensure this property). The following notation is used: m indicates the index of the mth element $y_{i,m}$ of the measurement vector $\boldsymbol{y}_i$, the matrix $\widehat{\boldsymbol{P}}_{i,m}$ is an approximation to the state covariance matrix after the assimilation of the set of measurements $\boldsymbol{y}_{1:i-1} \cup \{y_{i,k}\}_{k=1}^{m-1}$, $\boldsymbol{h}_{i,m}$ is the row vector equal to the mth row of $\boldsymbol{H}_i$, $r_{i,m}$ is the mth diagonal element of $\boldsymbol{R}_i$, and $\widehat{\boldsymbol{x}}_{i|i}$ is the best estimate of $\boldsymbol{x}_i$ given $\boldsymbol{y}_{1:i}$. Note that it is the column vector $\widehat{\boldsymbol{P}}_{i,m}\boldsymbol{h}_{i,m}^T$ and not the state covariance matrix itself that is required in the calculation of the Kalman gain vector estimate.

Both the GPF and the EnKF approximate the posterior distribution of the state given the measurements by a single Gaussian and it is, as a result, not surprising that the two methods are equivalent under some set of conditions. Specifically, the GPF and the EnKF give the same state estimates under signal model (2), jointly Gaussian $\boldsymbol{x}_1$, $\boldsymbol{u}_{1:T}$, and $\boldsymbol{v}_{1:T}$, and if importance sampling is not used in the GPF, i.e., if, in the

Initialize: Draw $\boldsymbol{x}^l_{1|0} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{\mu}_{1|0}, \boldsymbol{\Sigma}_{1|0})$
for $i = 1$ **to** T **do**
(Measurement update) (Note: $\boldsymbol{x}^l_{i,0} \equiv \boldsymbol{x}^l_{i|i-1}$)
for $m = 1$ **to** M **do**
$\overline{\boldsymbol{x}}_{i,m-1} = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{x}^l_{i,m-1}$
$\widetilde{\boldsymbol{x}}^l_{i,m-1} = \boldsymbol{x}^l_{i,m-1} - \overline{\boldsymbol{x}}_{i,m-1}$
$\widehat{\boldsymbol{P}}_{i,m}\boldsymbol{h}^T_{i,m} = \frac{1}{L-1}\sum_{l=1}^{L} \widetilde{\boldsymbol{x}}^l_{i,m-1}(\boldsymbol{h}_{i,m}\widetilde{\boldsymbol{x}}^l_{i,m-1})$
Compute Kalman gain estimate:
$\widehat{\boldsymbol{k}}_{i,m} = (\widehat{\boldsymbol{P}}_{i,m}\boldsymbol{h}^T_{i,m})/[\boldsymbol{h}_{i,m}(\widehat{\boldsymbol{P}}_{i,m}\boldsymbol{h}^T_{i,m}) + r_{i,m}]$
Update ensemble members:
Draw $y^l_{i,m} \overset{\text{iid}}{\sim} \mathcal{N}(y_{i,m}, r_{i,m})$
$\boldsymbol{x}^l_{i,m} = \boldsymbol{x}^l_{i,m-1} + \widehat{\boldsymbol{k}}_{i,m}(y^l_{i,m} - \boldsymbol{h}_{i,m}\boldsymbol{x}^l_{i,m-1})$
end
Filtered estimate: $\widehat{\boldsymbol{x}}_{i|i} = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{x}^l_{i,M}$
(Time update) (Note: $\boldsymbol{x}^l_{i|i} \equiv \boldsymbol{x}^l_{i,M}$)
$\boldsymbol{x}^l_{i+1|i} = \boldsymbol{F}_i\boldsymbol{x}^l_{i|i} + \boldsymbol{u}^l_i, \quad \boldsymbol{u}^l_i \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_i)$
end

Fig. 2. The sequential EnKF algorithm.

measurement update, samples are drawn as $\boldsymbol{x}^l_{i|i} \overset{\text{iid}}{\sim} p(\boldsymbol{x}_i|\boldsymbol{y}_{1:i})$ (which is Gaussian under (2)) and the weights are chosen as $\overline{w}^l_i = 1$. Another property shared with the GPF is that the number of ensemble members L may become prohibitive for large state dimension N to obtain satisfactory state estimates.

4. LOCALIZATION

When the state dimension is large, the principle of localization must be applied to the EnKF for the sake of computational tractability. Without localization, each measurement can affect the state estimate at all locations and in the absence of mismodeling, approximations, or any other uncertainties, the global influence of each measurement is justified. However, the propagation of many small errors globally over the state estimate can accumulate as the measurements are processed, potentially causing filter divergence unless the number of particles is increased to offset the error. By localizing, each measurement is restricted to only affect the state estimate in a neighborhood about the elements of the state contributing to the measurement. For example, each measurement in tomography is a line integral and localization can be achieved by restricting updates to state elements intersected by and neighboring the measurement ray.

The LEnKF is obtained from the EnKF by replacing the Kalman gain estimation step in Fig. 2 with:

$$\widehat{\boldsymbol{k}}_{i,m} = \boldsymbol{c}_{i,m} \circ (\widehat{\boldsymbol{P}}_{i,m}\boldsymbol{h}^T_{i,m})/[\boldsymbol{h}_{i,m}(\widehat{\boldsymbol{P}}_{i,m}\boldsymbol{h}^T_{i,m}) + r_{i,m}], \quad (3)$$

where $\boldsymbol{c}_{i,m} \in \{0,1\}^N$ and $\circ$ is the Hadamard product, i.e., the element by element matrix product. If the nth element of $\boldsymbol{c}_{i,m}$ is equal to 1 then the nth element of each ensemble member $\boldsymbol{x}^l_{i,m}$ is updated in the usual manner by the EnKF, otherwise it is not affected by the mth element of the measurement $\boldsymbol{y}_i$. Thus, the nonzero elements of $\boldsymbol{c}_{i,m}$ are chosen to correspond to the local neighborhood affected by a measurement. Because only a small subset of the estimate is updated, localization offers significant computational savings but at the cost of introducing a bias into the measurement update. For certain applications, the error introduced by the bias is offset by the reduction in accumulated errors.

5. NUMERICAL EXAMPLE

To demonstrate the potential of the LEnKF, a highly dynamic movie of the collapse of a cold, self-gravitating, magnetized nebular cloud is reconstructed from tomographic measurements. The state is evolved with the magnetohydrodynamics (MHD) equations, the coupled set of Newton's fluid and Maxwell's electrodynamic equations. The movie is composed of 64 frames each containing 33×33 pixels, four of which are displayed in Fig. 3 (left). Each measurement is a line integral through the computational domain plus a small (0.1%) amount of white noise, i.e., $\boldsymbol{R}_i$ is diagonal. A total of 47 line integral measurements are uniformly sampled per time step from the tomographic projection at angle θ_i. The projection angle θ_i is swept uniformly through 360° over the 64 frames. The state noise process is modeled by a banded matrix $\boldsymbol{Q}_i$ (with 4 bands above and below the diagonal) to provide a degree of temporal smoothness while the measurement equation is augmented with a derivative matrix $\boldsymbol{D}$ to provide a degree of spatial smoothness through regularization [2]. The state transition matrix is chosen as $\boldsymbol{F}_i = \boldsymbol{I}$, i.e., the state evolution is modeled as being completely stochastic in nature. The result of this experiment is given in Fig. 3 (right), where the state at time index $i = 44$ is compared to the (optimal) Kalman filter, EnKF, and LEnKF reconstructions. A quantitative comparison of these methods is given in Table 1, where L is the number of particles, error is measured as $\sum_{i=1}^{T} \|\boldsymbol{x}_i - \widehat{\boldsymbol{x}}_{i|i}\|_2/\|\boldsymbol{x}_i\|_2$, and the time compares how much computation was required to reconstruct the entire 64 frame movie on a 2.4 GHz workstation. From Fig. 3 (right) and Table 1 it is clear that the LEnKF provides a faithful estimate of the state that is nearly as good as the one given by the optimal Kalman filter but at significantly reduced computational cost. Additionally, the LEnKF is much cheaper than the EnKF and provides a much better state estimate for equal number of particles (though the quality of the EnKF estimate would surpass that of the LEnKF given enough particles).

6. SUMMARY

In this paper, we have made an explicit connection between the GPF and the EnKF. Also, the principle of localization was

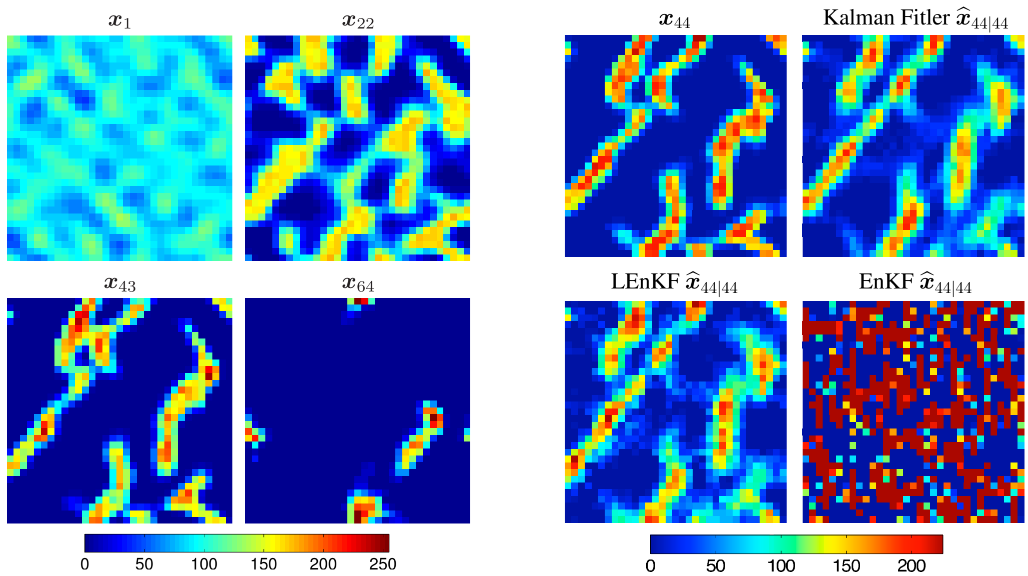

Fig. 3. The hidden state x_i at four time indices (left) and the results of the numerical example (right).

Method	L	Error	Time
KF	N/A	1.1	60 m
EnKF	256	16	60 m
LEnKF	256	1.2	3 m

Table 1. Quantitative results of the numerical example.

applied to the EnKF to give the LEnKF, a state estimation procedure for large dimension applications. Finally, the LEnKF was demonstrated in a numerical example where a hidden state evolving from a physical model was reconstructed from tomographic measurements.

Acknowledgments: The authors would like to thank J. Anderson (NCAR) for helpful discussions on localization and C. Gammie (UIUC) for the MHD simulation. This work has been supported in part by the NSF under grants ATM 0135073 and ATM 0620550 CMG and by NASA under grant NNG-04GG32G to the University of Illinois.

7. REFERENCES

[1] N. Gordon, D. Salmond, and C. Ewing, "Bayesian state estimation for tracking and guidance using the bootstrap filter," *AIAA Journal of Guidance, Control, and Dynamics*, vol. 18, pp. 1434–1443, 1995.

[2] R. A. Frazin, M. D. Butala, A. Kemball, and F. Kamalabadi, "Time-dependent reconstruction of nonstationary objects with tomographic or interferometric measurements," *Astrophysical Journal Letters*, vol. 635, pp. L197–L200, 2005.

[3] 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. 10 143–10 162, 1994.

[4] J. Geweke, "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica*, vol. 57, pp. 1317–1339, 1989.

[5] N. Aggarwal and Y. Bresler, "Spatio-temporal modeling and adaptive acquisition for free-breathing cardiac magnetic resonance imaging," in *Second SIAM Conference on Imaging in Science*, 2004.

[6] T. Ernst, O. Speck, L. Itti, and L. Chang, "Simultaneous correction for interscan patient motion and geometric distortions in echoplanar imaging," *Magnetic Resonance in Medicine*, vol. 42, pp. 201–205, 1999.

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

[8] B. Li, T. Bengtsson, and P. Bickel, "Curse-of-dimensionality revisited: Collapse of importance sampling in very high-dimensional systems," University of California, Berkeley, Department of Statistics, Tech. Rep. 696, 2005.

[9] J. H. Kotecha and P. M. Djurić, "Gaussian particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2592–2601, 2003.

[10] P. L. Houtekamer and H. L. Mitchell, "A sequential ensemble Kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 129, pp. 123–137, 2001.