Construction of the proposal distribution for the particle filter using the ensemble transform Kalman filter

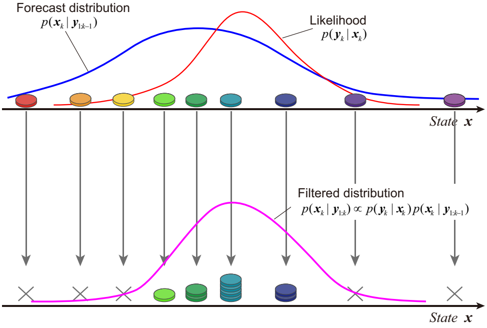
S. Nakano

The Institute of Statistical Mathematics

18 January 2013

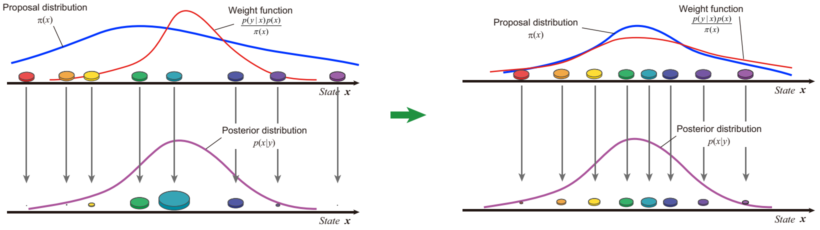
Particle filter

- The particle filter represents a probability density function (PDF) by a Monte Carlo approximation.
- The ensemble representing a posterior PDF is obtained by resampling a forecast ensemble.
- It is applicable even to the cases with non-linear or non-Gaussian observations.
- However, it requires a huge number of particles to avoid the problem due to ensemble degeneracy.



Particle filter

- The particle filter algorithm is based on the importance sampling method, which represents the posterior PDF by weighted sample.
- If we choose a good proposal distribution similar to the posterior PDF, the imbalance of weights among the particles can be reduced, and therefore we could achieve high accuracy and high computational efficiency.
- The prior (forecast) PDF is usually used as the proposal PDF. This makes the algorithm simple. But, a large discrepancy often exists between the prior PDF and the posterior PDF.



Improvement of the proposal

Several studies have proposed a way to improve the proposal distribution.

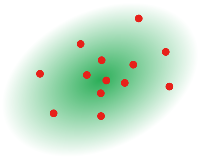
- Auxiliary particle filter (Pitt and Shephard, 1999).
- Nudge ensemble members toward the observation by considering the system noise (model error) (Chorin et al., 2010; van Leeuwen, 2010; 2011).
- Weighted ensemble Kalman filter (Papadakis et al., 2010).

Ensemble transform Kalman filter

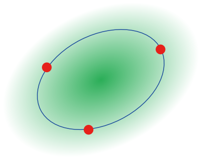
- We use the ensemble transform Kalman filter (ETKF) (Bishop et al, 2001) to obtain a proposal distribution for importance sampling, which is similar to the weighted ensemble Kalman filter.
- In the ETKF, the first and second order moments of the PDF is represented by an ensemble.
- The ETKF algorithm is derived on the basis of the linear Gaussian observation model.
- It therefore ignores non-Gaussian features of the PDF.
- The aim of using the importance sampling method is to represent non-Gaussian features of the PDF that is not considered by the ETKF.

Remark

- The importance sampling method is based on a Monte Carlo representation which assumes that the ensemble size is of the order of the exponential of the state dimension.
- On the other hand, the ensemble Kalman filters (esp. ETKF) is used with a limited ensemble size (much less than the state dimension).
- If the ensemble size N is smaller than the rank of the state covariance matrix, the ensemble would form a simplex in an $(N-1)$-dimensional subspace; that is, it provides a spherical simplex representation of the PDF (Wang et al., 2004).
- The ETKF uses a conceptually different representation from the importance sampling.

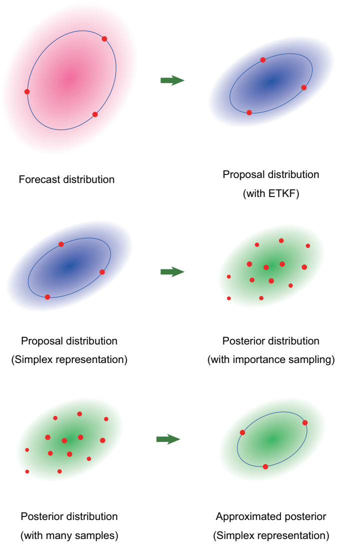


Monte Carlo representation



Simplex representation

Overview



- We consider cases in which the forecast PDF is represented by a simplex representation with a limited-size ensemble.
- To allow nonlinear or non-Gaussian observation models, the simplex representation is converted into a Monte Carlo representation. Then the importance sampling method is applied.
- Finally, the importance sampling result is converted into a simplex representation again.

Some definitions

- Suppose that the forecast distribution is represented by an ensemble $\{\boldsymbol{x}_{k|k-1}^{(1)}, \ldots, \boldsymbol{x}_{k|k-1}^{(N)}\}$.
- The mean of the forecast distribution is obtained as:

$$\bar{\boldsymbol{x}}_{k|k-1} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}_{k|k-1}^{(i)}.$$

- We define a matrix $\mathsf{X}_{k|k-1}$ and $\mathsf{Y}_{k|k-1}$ as

$$\mathsf{X}_{k|k-1} = \frac{1}{\sqrt{N}} \begin{pmatrix} \delta\boldsymbol{x}_{k|k-1}^{(1)} & \cdots & \delta\boldsymbol{x}_{k|k-1}^{(N)} \end{pmatrix}, \quad \mathsf{Y}_{k|k-1} = \frac{1}{\sqrt{N}} \begin{pmatrix} \delta\boldsymbol{y}_{k|k-1}^{(1)} & \cdots & \delta\boldsymbol{y}_{k|k-1}^{(N)} \end{pmatrix},$$

where $\delta\boldsymbol{x}_{k|k-1}^{(i)} = \boldsymbol{x}_{k|k-1}^{(i)} - \bar{\boldsymbol{x}}_{k|k-1}$ and $\delta\boldsymbol{y}_{k|k-1}^{(i)} = H_k(\boldsymbol{x}_{k|k-1}^{(i)}) - \overline{H_k(\boldsymbol{x}_{k|k-1})}$, respectively, and we assumed the following observation model

$$\boldsymbol{y}_k = H_k(\boldsymbol{x}_k) + \boldsymbol{w}_k.$$

- The covariance matrix of the forecast (predictive) distribution is written as $\mathsf{V}_{k|k-1} = \mathsf{X}_{k|k-1}\mathsf{X}_{k|k-1}^T$.

Ensemble transform Kalman filter (ETKF) (Bishop et al., 2001)

- The mean of the filtered distribution is obtained according to the Kalman filter algorithm:

$$\bar{\boldsymbol{x}}^{\dagger}_{k|k} = \bar{\boldsymbol{x}}_{k|k-1} + \mathsf{K}_k\left(\boldsymbol{y}_k - \mathsf{H}_k\boldsymbol{x}_{k|k-1}\right)$$

- The square root of the covariance matrix is also calculated as $\mathsf{X}^{\dagger}_{k|k} = \mathsf{X}_{k|k-1}\mathsf{T}_k$, where the matrix T_k is designed to satisfy $\mathsf{V}^{\dagger}_{k|k} = \mathsf{X}^{\dagger}_{k|k}\mathsf{X}^{\dagger T}_{k|k}$ and $\mathsf{X}^{\dagger}_{k|k}\mathbf{1} = \mathbf{0}$, where $\mathbf{1} = (1 \ \cdots \ 1)^T$. The latter condition is required to preserve the mean of the PDF (Wang et al., 2004; Livings et al., 2008).
- Using the following eigen-value decomposition

$$\mathsf{Y}_{k|k-1}\mathsf{R}_k^{-1}\mathsf{Y}_{k|k-1} = \mathsf{U}_k\Lambda_k\mathsf{U}_k^T,$$

the matrices K_k and T_k are obtained as follows:

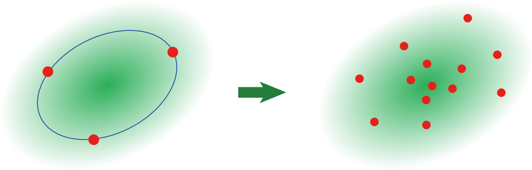
$$\mathsf{K}_k = \mathsf{X}_{k|k-1}\mathsf{U}_k(\mathrm{I}_N + \Lambda_k)^{-1}\mathsf{U}_k^T\mathsf{Y}_{k|k-1}^T\mathsf{R}_k^{-1},$$
$$\mathsf{T}_k = \mathsf{U}_k(\mathrm{I}_N + \Lambda_k)^{-\frac{1}{2}}\mathsf{U}_k^T,$$

where R_k is the covariance matrix of the observation noise.

Sampling from the ETKF estimate

- The ETKF estimates the filtered (posterior) distribution as a Gaussian distribution $\mathcal{N}(\bar{\boldsymbol{x}}^{\dagger}_{k|k}, \mathsf{V}^{\dagger}_{k|k})$.
- However, it does not actually calculate the covariance matrix $\mathsf{V}^{\dagger}_{k|k}$ itself. Instead, a square root of the covariance matrix $\mathsf{X}^{\dagger}_{k|k}$ is calculated.
- Using the matrix $\mathsf{X}^{\dagger}_{k|k}$, we can easily generate a large number of random numbers obeying $\mathcal{N}(\bar{\boldsymbol{x}}^{\dagger}_{k|k}, \mathsf{V}^{\dagger}_{k|k})$ using the following generative model:

$$\boldsymbol{x}_k = \bar{\boldsymbol{x}}^{\dagger}_{k|k} + \mathsf{X}^{\dagger}_{k|k} z_k, \quad \text{where } z_k \sim \mathcal{N}(\mathbf{0}, \mathrm{I}_N).$$



Simplex representation

Monte Carlo representation

Importance sampling

- Since the posterior distribution $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ is written as

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) = \frac{p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})}{\pi(\boldsymbol{x}_k)}\pi(\boldsymbol{x}_k) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})\pi(\boldsymbol{x}_k)}\pi(\boldsymbol{x}_k),$$

  the posterior $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ can be represented by the importance sampling using the sample drawn from $\pi(\boldsymbol{x}_k)$:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) \approx \sum_{j=1}^{M} \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{\pi(j)})p(\boldsymbol{x}_k^{\pi(j)}|\boldsymbol{y}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})\pi(\boldsymbol{x}_k^{\pi(j)})}\delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{\pi(j)}).$$

- In the normal particle filter, the forecast $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})$ is used as $\pi(\boldsymbol{x}_k)$.
- On the other hand, we use the estimate of $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ obtained by the ETKF as the proposal $\pi(\boldsymbol{x}_k)$.

Importance sampling

- If we obtain the proposal $\pi(\boldsymbol{x}_k)$ by the ETKF, we can generate a large number of particles from $\pi(\boldsymbol{x}_k)$ according to the following generative model:

$$\boldsymbol{x}_k^{\pi(j)} = \bar{\boldsymbol{x}}_{k|k}^{\dagger} + \mathsf{X}_{k|k}^{\dagger} \boldsymbol{z}_k^{(j)} \quad \left(\boldsymbol{z}_k^{(j)} \sim \mathcal{N}(\mathbf{0}, \mathrm{I}_N)\right).$$

- In order to approximate the posterior $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ using the importance sampling method as follows:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) \approx \sum_{j=1}^{M} \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{\pi(j)}) p(\boldsymbol{x}_k^{\pi(j)}|\boldsymbol{y}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1}) \pi(\boldsymbol{x}_k^{\pi(j)})} \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{\pi(j)}),$$

we need to calculate

$$\frac{p(\boldsymbol{x}_k^{\pi(j)}|\boldsymbol{y}_{1:k-1})}{\pi(\boldsymbol{x}_k^{\pi(j)})}$$

for each particle $\boldsymbol{x}_k^{\pi(j)}$. (We can obtain $p(\boldsymbol{y}_k|\boldsymbol{x}_k^{\pi(j)})$ from the observation model.)

Importance weight

- According to the generative model

$$\boldsymbol{x}_k^{\pi(j)} = \bar{\boldsymbol{x}}_{k|k}^{\dagger} + \mathsf{X}_{k|k}^{\dagger}\boldsymbol{z}_k^{(j)} \quad \left(\boldsymbol{z}_k^{(j)} \sim \mathcal{N}(\mathbf{0}, \mathrm{I}_N)\right),$$

  $\pi(\boldsymbol{x}_{k|k}^{\pi,(j)})$ can be associated with the probability density for $\boldsymbol{z}_k^{(j)}$, $p(\boldsymbol{z}_k^{(j)})$.

- The probability density $p(\boldsymbol{z}_k^{(j)})$ is proportional to $\exp\left(-\|\boldsymbol{z}_k^{(j)}\|^2/2\right)$.
- Considering that $\mathsf{X}_{k|k}^{\dagger}$ satisfies the mean-preserving condition $\mathsf{X}_{k|k}^{\dagger}\mathbf{1} = \mathbf{0}$, the component parallel to $\mathbf{1}$ is projected onto a null space. We therefore obtain

$$\pi(\boldsymbol{x}_{k|k}^{\pi,(j)}) \propto \exp\left[-\frac{1}{2}\left(\left\|\boldsymbol{z}_k^{(j)}\right\|^2 - \frac{(\mathbf{1}^T\boldsymbol{z}_k^{(j)})^2}{N}\right)\right].$$

We consider that a sample from the forecast $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1})$ is generated according to the following model:

$$\boldsymbol{x}_k = \bar{\boldsymbol{x}}_{k|k-1} + \mathsf{X}_{k|k-1}\boldsymbol{z}_k \quad \left(\boldsymbol{z}_k \sim \mathcal{N}(\mathbf{0}, I_N)\right).$$

We can then evaluate the probability density that $\boldsymbol{x}_{k|k}^{\pi,(j)}$ is drawn from the forecast distribution as follows:

$$\begin{aligned}
\boldsymbol{x}_{k|k}^{\pi,(j)} &= \overline{\boldsymbol{x}}_{k|k}^{\dagger} + \mathsf{X}_{k|k}^{\dagger}\boldsymbol{z}_k^{(j)} = \overline{\boldsymbol{x}}_{k|k-1} + \mathsf{K}_k\left(\boldsymbol{y}_k - \overline{\boldsymbol{h}_k(\boldsymbol{x}_{k|k-1})}\right) + \mathsf{X}_{k|k-1}\mathsf{T}_k\boldsymbol{z}_k^{(j)} \\
&= \overline{\boldsymbol{x}}_{k|k-1} + \mathsf{X}_{k|k-1}\left[\mathsf{U}_k(\mathrm{I}_N + \Lambda_k)^{-1}\mathsf{U}_k^T\mathsf{Y}_{k|k-1}^T\mathsf{R}^{-1}\left(\boldsymbol{y}_k - \overline{\boldsymbol{h}_k(\boldsymbol{x}_{k|k-1})}\right) + \mathsf{T}_k\boldsymbol{z}^{(j)}\right] \\
&= \overline{\boldsymbol{x}}_{k|k-1} + \mathsf{X}_{k|k-1}\boldsymbol{\zeta}_k^{(j)}
\end{aligned}$$

where

$$\boldsymbol{\zeta}_k^{(j)} = \mathsf{U}_k(\mathrm{I}_N + \Lambda_k)^{-1}\mathsf{U}_k^T\mathsf{Y}_{k|k-1}^T\mathsf{R}^{-1}\left(\boldsymbol{y}_k - \overline{\boldsymbol{h}_k(\boldsymbol{x}_{k|k-1})}\right) + \mathsf{T}_k\boldsymbol{z}^{(j)}.$$

We therefore obtain

$$p(\boldsymbol{x}_{k|k}^{\pi,(j)}|\boldsymbol{y}_{1:k-1}) \propto \exp\left[-\frac{1}{2}\left(\left\|\boldsymbol{\zeta}_k^{(j)}\right\|^2 - \frac{(\mathbf{1}^T\boldsymbol{\zeta}_k^{(j)})^2}{N}\right)\right].$$

As seen previously, the posterior distribution is approximated as

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) \approx \sum_{j=1}^{M} \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k^{\pi(j)})p(\boldsymbol{x}_k^{\pi(j)}|\boldsymbol{y}_{1:k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1})\pi(\boldsymbol{x}_k^{\pi(j)})}\delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{\pi(j)}).$$

If we generate the proposal sample according to the following model:

$$\boldsymbol{x}_{k|k}^{\pi,(j)} = \bar{\boldsymbol{x}}_{k|k-1} + \mathsf{X}_{k|k-1}\boldsymbol{z}_k^{(j)} \quad \left(\boldsymbol{z}_k^{(j)} \sim \mathcal{N}(\boldsymbol{0}, \mathrm{I}_N)\right),$$

the weight for each particle can be given as follows:

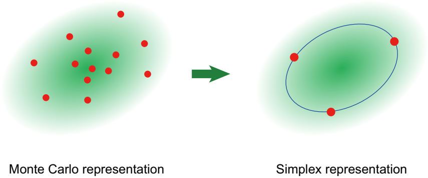
$$\beta_k^{(j)} \propto \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_{k|k}^{\pi,(j)}) \ \exp\left[-\frac{1}{2}\left(\left\|\boldsymbol{\zeta}_k^{(j)}\right\|^2 - \frac{(\boldsymbol{1}^T\boldsymbol{\zeta}_k^{(j)})^2}{N}\right)\right]}{\exp\left[-\frac{1}{2}\left(\left\|\boldsymbol{z}_k^{(j)}\right\|^2 - \frac{(\boldsymbol{1}^T\boldsymbol{z}_k^{(j)})^2}{N}\right)\right]}.$$

We then obtain a new approximation of the posterior PDF:

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) \approx \sum_{j=1}^{M} \beta_k^{(j)}\delta(\boldsymbol{x}_k - \boldsymbol{x}_k^{\pi(j)}).$$

Ensemble reconstruction

- Using the weight $\beta_k^{(j)}$, we can obtain a random sample from the posterior $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ with the rejection sampling method or the independent chain Metropolis-Hastings method.
- However, we consider the case in which a large ensemble size is not allowed. A small-size ensemble generated randomly would not give a good approximation of $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$.
- To avoid the errors due to the randomness, we construct a simplex approximation that represents the first and second order moments of the posterior.



Moments on the z-space

If we calculate the mean and the covariance on the z-space:

$$\bar{z}_k = \sum_{i=1}^{M} \beta_k^{(j)} z_k^{(j)}, \qquad \mathsf{V}_{z,k|k} = \sum_{i=1}^{M} \beta_k^{(j)} (z_k^{(j)} - \bar{z}_k)(z_k^{(j)} - \bar{z}_k)^T,$$

the mean and the covariance of the filtered distribution $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$ are given as follows:

$$\bar{\boldsymbol{x}}_{k|k} = \bar{\boldsymbol{x}}_{k|k}^{\dagger} + \mathsf{X}_{k|k}^{\dagger} \bar{z}_k, \qquad \mathsf{V}_{k|k} = \mathsf{X}_{k|k} \mathsf{X}_{k|k}^T = \mathsf{X}_{k|k}^{\dagger} \mathsf{V}_{z,k|k} \mathsf{X}_{k|k}^{\dagger T}$$

where $\bar{\boldsymbol{x}}_{k|k}^{\dagger}$ and $\mathsf{X}_{k|k}^{\dagger}$ provide the estimate by the ETKF.

To avoid the bias of the ensemble mean, the new $\mathsf{X}_{k|k}$ should also satisfy

$$\mathsf{X}_{k|k} \mathbf{1} = \mathbf{0}.$$

We define the following matrix

$$\mathsf{A} = \mathrm{I}_N - \frac{1}{N}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix},$$

which obviously satisfies

$$\mathsf{A}\mathbf{1} = 0.$$

The covariance matrix $\mathsf{V}_{k|k}$ can then be written as follows:

$$\begin{aligned} \mathsf{V}_{k|k} &= \mathsf{X}^{\dagger}_{k|k}\mathsf{V}_{z,k|k}\mathsf{X}^{\dagger T}_{k|k} \\ &= \mathsf{X}^{\dagger}_{k|k}\mathsf{A}\,\mathsf{V}_{z,k|k}\mathsf{A}^{T}\mathsf{X}^{\dagger T}_{k|k} \end{aligned}$$

because obviously $\mathsf{X}^{\dagger}_{k|k} = \mathsf{X}^{\dagger}_{k|k}\mathsf{A}$.

When we calculate the eigen-value decomposition of the matrix $\mathsf{A}\,\mathsf{V}_{z,k|k}\mathsf{A}$ as

$$\mathsf{A}\,\mathsf{V}_{z,k|k}\mathsf{A} = \mathsf{U}_{z,k}\Gamma_k\mathsf{U}_{z,k}^T,$$

the matrix $\mathsf{U}_{z,k}$ contains an eigen-vector which is parallel to $\mathbf{1}$ and corresponds to zero eigen-value. Therefore, if we define $\mathsf{X}_{k|k}$ as

$$\mathsf{X}_{k|k} = \mathsf{X}_{k|k}^{\dagger}\mathsf{U}_{z,k}\Gamma_k^{\frac{1}{2}}\mathsf{U}_{z,k}^T,$$

it satisfies both of the necessary conditions:

$$\mathsf{X}_{k|k}\mathsf{X}_{k|k}^T = \mathsf{X}_{k|k}^{\dagger}\mathsf{V}_{z,k|k}\mathsf{X}_{k|k}^{\dagger T},$$
$$\mathsf{X}_{k|k}\mathbf{1} = \mathbf{0}.$$

Ensemble reconstruction

Finally, we obtain ensemble perturbations:

$$\begin{pmatrix} \delta \boldsymbol{x}_{k|k}^{(1)} & \cdots & \delta \boldsymbol{x}_{k|k}^{(N)} \end{pmatrix} = \sqrt{N}\, \mathsf{X}_{k|k}.$$

We then obtain the filtered ensemble:

$$\boldsymbol{x}_{k|k}^{(i)} = \bar{\boldsymbol{x}}_{k|k} + \delta \boldsymbol{x}_{k|k}^{(i)}.$$

Remark

- Using the generative model, $\boldsymbol{x}_k^{\pi(j)} = \bar{\boldsymbol{x}}_{k|k}^{\dagger} + \mathsf{X}_{k|k}^{\dagger}\boldsymbol{z}_k^{(j)}$, the ensemble members are generated in the subspace spanned by the ensemble members.
- We could consider a small uncertainty in the complement space as follows
$$\boldsymbol{x}_k^{\pi,(j)} = \bar{\boldsymbol{x}}_{k|k}^{\dagger} + \mathsf{X}_{k|k}^{\dagger}\boldsymbol{z}_k^{(j)} + \boldsymbol{\varepsilon}_k^{(j)},$$
where $\boldsymbol{\varepsilon}_k^{(j)}$ is a random sample representing the uncertainty of the orthogonal complement space. But, this may invoke 'the curse of dimensionality'.
- As far as we ignore the complement space, we can convert between the importance sampling result and a spherical simplex representation through the calculation in the small subspace spanned by the forecast ensemble members. This would help reduce the computational cost.

Experiment

We performed experiments using the Lorenz 96 model (Lorenz and Emanuel 1998):

$$\frac{dx_l}{dt} = (x_{l+1} - x_{l-2})x_{l-1} - x_l + f$$

where $x_{-1} = x_{L-1}$, $x_0 = x_L$, and $x_{L+1} = x_1$. We take the dimension of a state vector L to be 40 and the forcing term f to be 8. One time step was assumed to be 0.01.

It was assumed that x_l can be observed only if l is an even number. This means that the half of the variables x_l are observable.

The following observation model is considered:

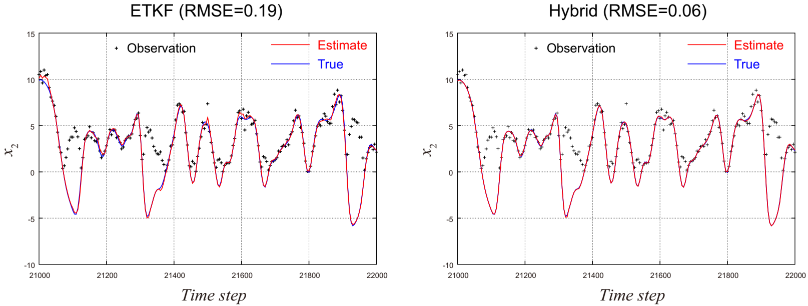
$$y_k^l = \log|x_k^l| + w_k^l \quad \left(w_k^l \sim \mathcal{N}(0, 0.0225)\right)$$

We use covariance inflation with the inflation factor 1.1.

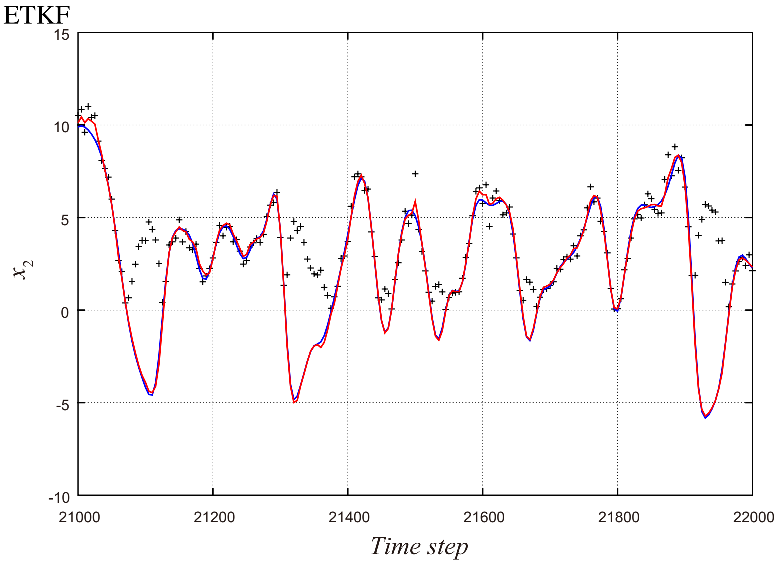
Result

An estimate of an observed variable

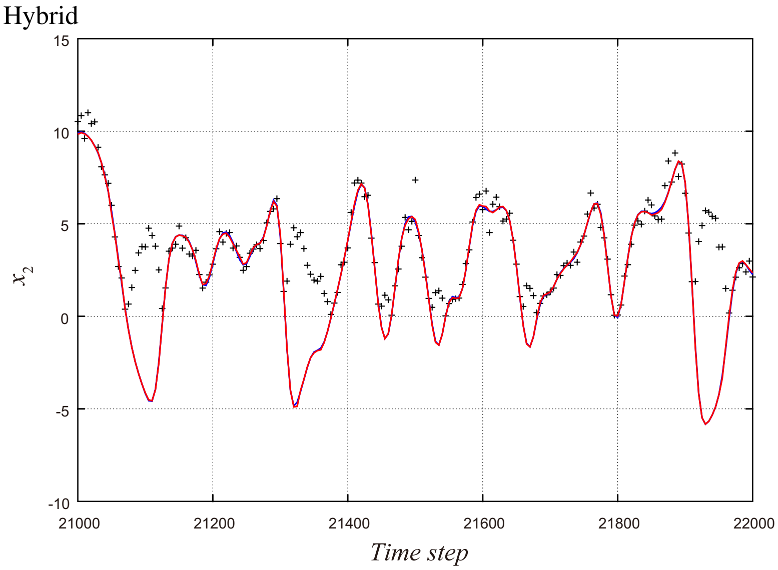
- With 30 ensemble members (and 1920 particles for importance sampling)
- RMSE: 0.06 (with the hybrid algorithm), 0.19 (with the ETKF)



Result


ETKF
15
10
5
0
-5
-10
21000
21200
21400
21600
21800
22000
x_2
Time step

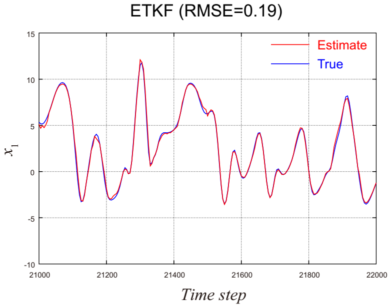
Result

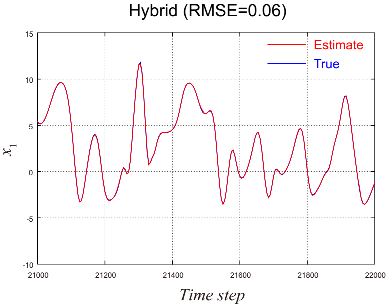

Hybrid
15
10
5
0
-5
-10
21000
21200
21400
21600
21800
22000
x_2
Time step

Result

An estimate of an unobserved variable

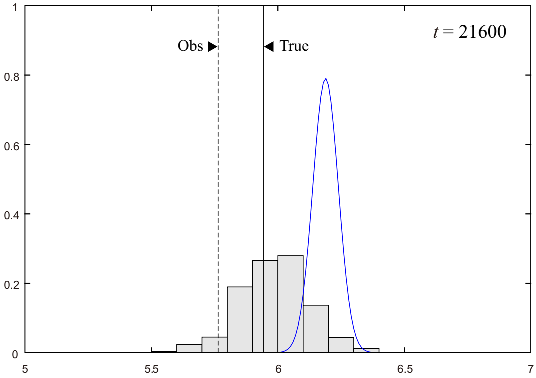
- With 30 ensemble members (and 1920 particles for importance sampling)
- RMSE: 0.06 (with the hybrid algorithm), 0.19 (with the ETKF)





Result

- Blue line: the result with the ETKF
- Histogram: the result with the hybrid algorithm
- Solid vertical line: the true state
- Dashed vertical line: the observed value



We also consider the following observation model:

$$y_k^l = \left| x_k^j + w_k^j \right| \quad \left(w_k^j \sim \mathcal{N}(0, 1) \right),$$

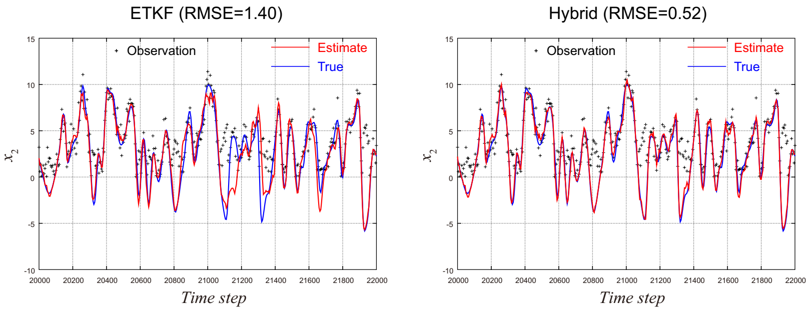
The system noise is assumed as follows:

$$v_k^l \sim \mathcal{N}(0, 0.0625).$$

Result

An estimate of an observed variable

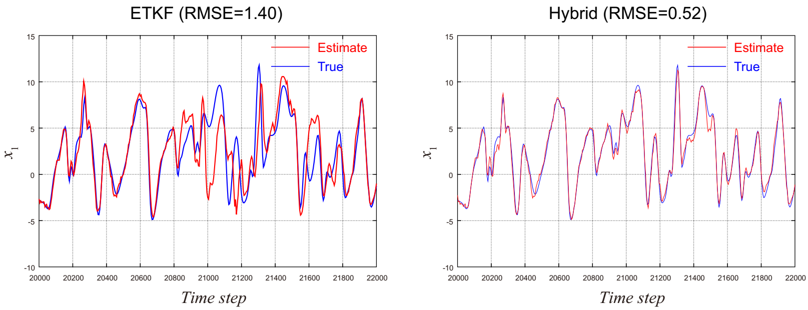
- With 30 ensemble members (and 3840 particles for importance sampling)
- RMSE: 0.52 (with the hybrid algorithm), 1.40 (with the ETKF)



Result

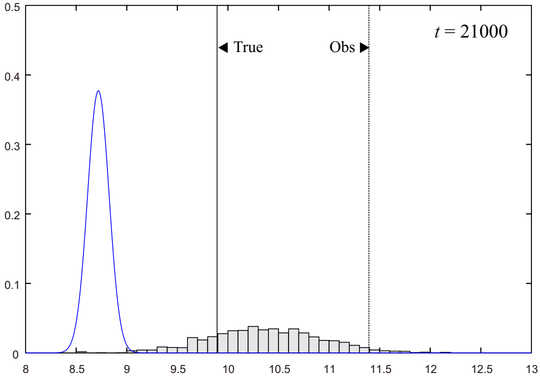
An estimate of an unobserved variable

- With 30 ensemble members (and 3840 particles for importance sampling)
- RMSE: 0.52 (with the hybrid algorithm), 1.40 (with the ETKF)



Result

- Blue line: the result with the ETKF
- Histogram: the result with the hybrid algorithm
- Solid vertical line: the true state
- Dashed vertical line: the observed value



Summary

- A good proposal distribution could improve the computational efficiency of the particle filter.
- We propose a hybrid algorithm which use the ensemble transform Kalman filter (ETKF) to obtain the proposal.
- While the importance sampling method used in the particle filter requires abundant particles, the ETKF is based on a spherical simplex representation which uses less particles than the state dimension. We then make the conversion between a simplex representation and a Monte Carlo representation.
- In our approach, this conversion is performed in the low-dimensional subspace spanned by the forecast ensemble members.
- Even though the uncertainty is considered only in the subspace, the proposed approach seems to well work in the cases with nonlinear, non-Gaussian observation models in which the application of ensemble Kalman filters is not valid.