P2-24 The 10th International Symposium on Data Assimilation

SOKENDAI

Toward optimal state and time-varying parameter estimation using the implicit equal-weights particle filter

Mineto Satoh (SOKENDAI) | Shin'ya Nakano (The Institute of Statistical Mathematics)

1. Introduction

Motivation:

- ✓ Avoid filter degeneracy in higher dimensional systems by the implicit equal-weights particle filter (IEWPF).
- ✓ Determine the scale factor α without prior assumptions.

Challenges:

- ✓ The analytical solution of the Lambert *W* function has two branches, so we have to choose one.
- ✓ In real nonlinear problems, the true posterior pdf is unknown and cannot be pre-tuned.

Contributions:

- ✓ Incorporate a sequential estimation method of the factor α based on the Kullback-Leibler (KL) divergence minimization into IEWPF.
- ✓ Reduce the calculation cost of KL divergence by converting particle distribution into a histogram.
- ✓ Compared with the analysis values of variance using a linear model and confirmed the applicability to the 1000-dimensional nonlinear Lorenz-96 model.

2. Methods

2-1. Implicit Equal-Weights Particle Filter (IEWPF) [1]

- Define the augmented state-space model as follows [2]:

$$\mathrm{z}^n = \begin{pmatrix} x^n \\ \theta^{n-1} \end{pmatrix} = \begin{pmatrix} f(x^{n-1}, \theta^{n-2}) \\ \theta^{n-2} \end{pmatrix} + \tilde{\beta}^n, \tilde{\beta}^n \sim N(0, \tilde{Q}^n). \qquad (1)$$

- The weights of all particles are equal to the target weight.
- Each particle is set to the mode of optimal proposal density plus a scaled random perturbation:

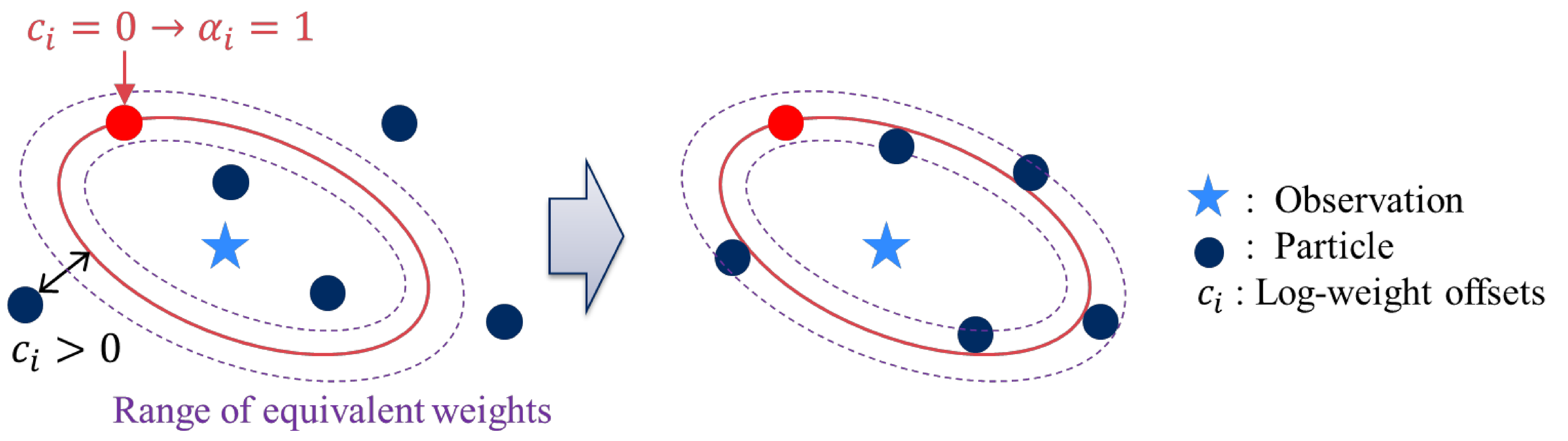
$$z_i^n = \zeta_i^n + \alpha_i^{1/2} P^{1/2} \xi_i^n, \quad \xi \sim N(0, \mathbf{I}), \qquad (2)$$

where P is matrices indicating the width of the pdf and α is the particle-specific scalar factor to be estimated.

- α_i is determined so that all particles are at target weights and is given by the analytical solution of the Lambert *W* function [1].

2-2. Estimation approach of factor α by reverse KL

- Define a distribution $q(z^n|\,\alpha)$ with α as the hyperparameter.
- **Hypothesis:** $\alpha = 1$, i.e., the optimal proposal distribution is assumed to be the target distribution $\hat{p}(z^n)$.



- Denote branches of the Lambert *W* function solutions with $\alpha < 1$ and $\alpha \geq 1$ as $\alpha_{<1}$ and $\alpha_{\geq 1}$.
- Determine α_i for each particle by minimizing the reverse KL:

$$\underset{\alpha_i \in\, \alpha_{<1} \text{ or } \alpha_{\geq 1}}{\arg\min}\ \mathrm{RKL}(\alpha_i) = \underset{\alpha_i \in\, \alpha_{<1} \text{ or } \alpha_{\geq 1}}{\arg\min}\ \mathrm{KL}\big(q(z^n|\,\alpha)\,||\,\hat{p}(z^n)\big) \qquad (3)$$

2-3. Iterative α estimation by approximate KL

- Generate a histogram from the distribution $\hat{p}(z^n)$ and $q(z^n|\,\alpha)$.
- If the number of particles in the i-th bin b_i is $n(b_i)$ and the number of bins is N_b, the probability of i-th bin is expressed as

$$P(b_i) = (n(b_i) + \delta_b)/\textstyle\sum_{i=1}^{N_b} n(b_i) \qquad (4)$$

- The reverse KL in Eq.3 for variable z_j at step n is obtained by

$$\mathrm{RKL}_j^n(\alpha) = \textstyle\sum_{i=1}^{N_b} Q(b_i) \log\left(\frac{Q(b_i)}{P(b_i)}\right) \qquad (5)$$

- Introduce threshold α_{th} $(0 \leq \alpha_{th} \leq 1)$ and select α_i as follows:

$$\alpha_i \in \begin{cases} \alpha_{<1} & \text{if } \ U[0,1] < \alpha_{th} \\ \alpha_{\geq 1} & \text{if } \ U[0,1] \geq \alpha_{th} \end{cases}. \qquad (6)$$

- At iteration k, update α_{th}^k as follows so that RKL in Eq. 5 is minimized:

$$\alpha_{th}^{k+1} \leftarrow \alpha_{th}^k - \lambda_\alpha\, g^k(\alpha), \quad g^k(\alpha) \propto \nabla\, \mathrm{RKL}^n(\alpha) \qquad (7)$$

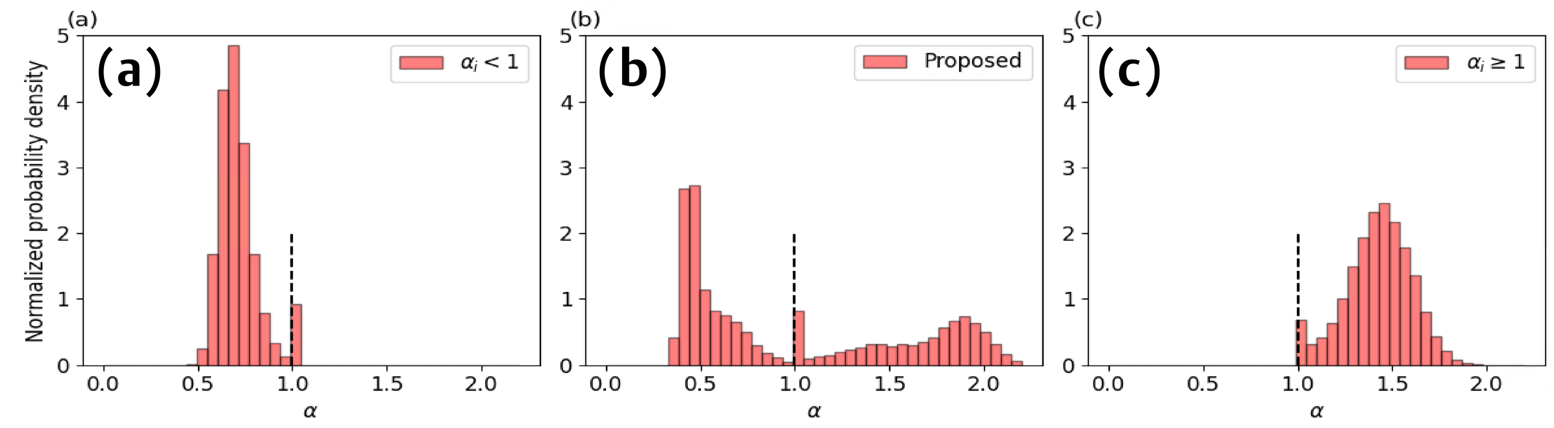
3. Twin experiment 1 : Linear model

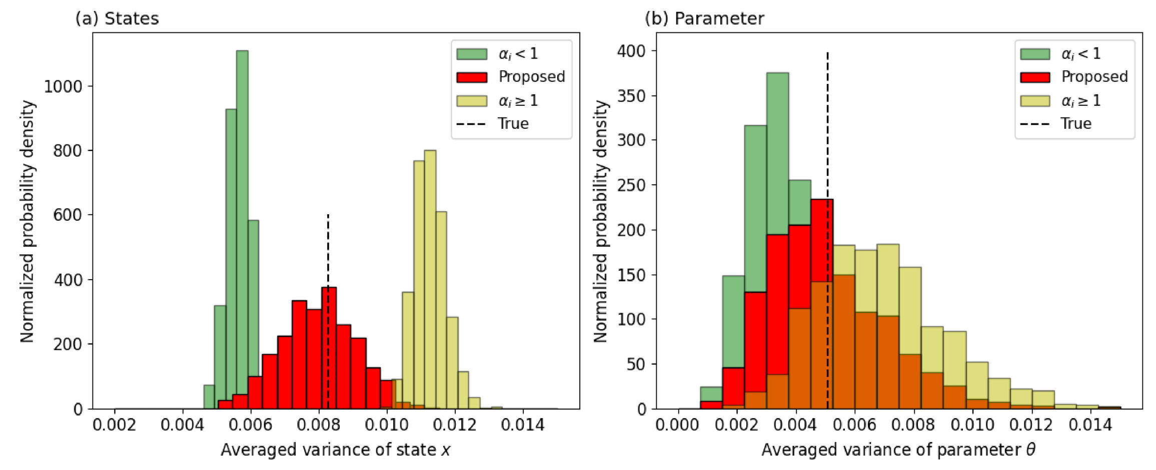
- System model for $x \in \mathbb{R}^{N_x}$ $(N_x = 1000)$ and $\theta \in \mathbb{R}^{N_\theta}$ $(N_\theta = 1)$:

$$\begin{pmatrix} x^n \\ \theta^{n-1} \end{pmatrix} = \begin{pmatrix} F_x & F_{x\theta} \\ O & I \end{pmatrix} \begin{pmatrix} x^{n-1} \\ \theta^{n-2} \end{pmatrix} + \begin{pmatrix} I & F_{x\theta} \\ O & I \end{pmatrix} \begin{pmatrix} \beta^n \\ \eta^{n-1} \end{pmatrix}, \mathrm{F_x} = \mathrm{I}, \mathrm{F}_{x\theta} = (0.1, \dots)^{\mathrm{T}} \qquad (8)$$

- Histogram of α accumulated from the 20th to 1000th steps:

(a) $\alpha < 1$ branch only, (b) proposed method, (c) $\alpha \geq 1$ branch only





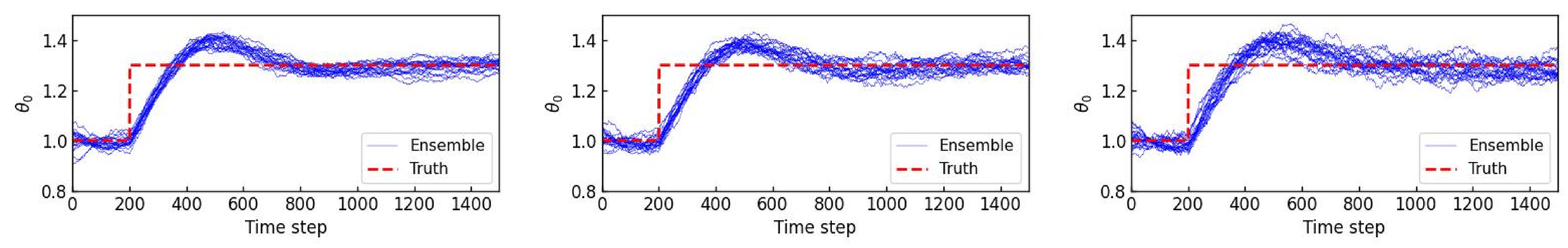
Comparison with the true value of the histogram of the accumulated variance at steps 20-1000.

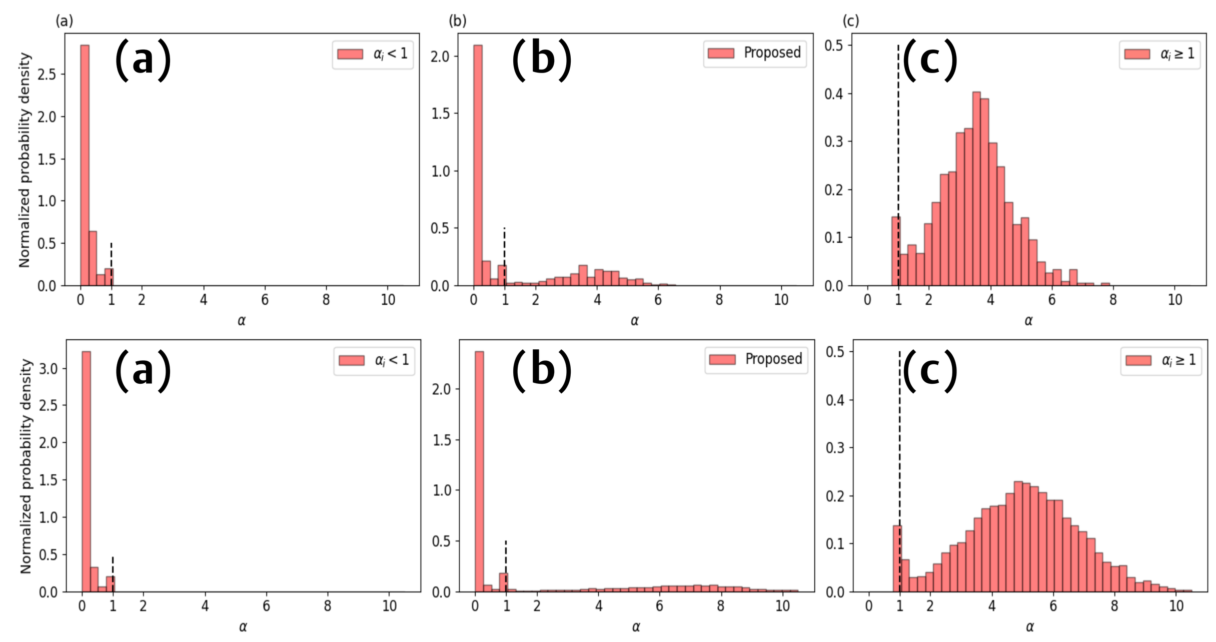
4. Twin experiment 2 : Lorenz-96 model

- System model for $x \in \mathbb{R}^{N_x}$ $(N_x = 1000)$ and $\theta \in \mathbb{R}^{N_\theta}$ $(N_\theta = 3)$:

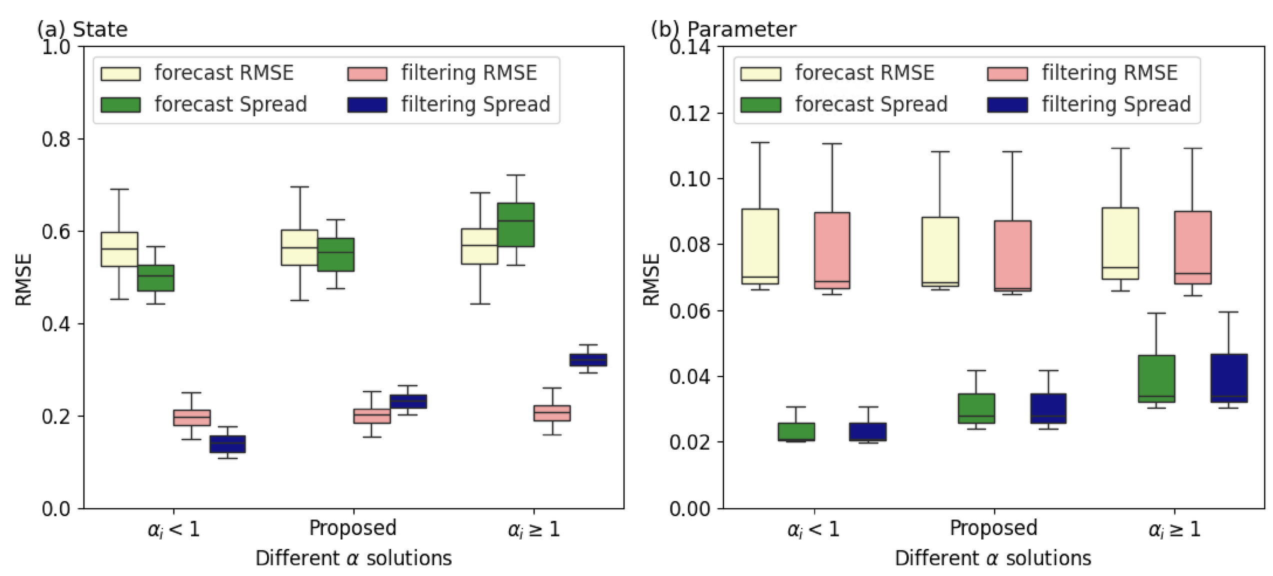
$$\frac{d}{dt} x_j = (x_{j+1} - x_{j-2}) x_{j-1} - x_j + F_j(\theta_0, \theta_1, \theta_2) \qquad (9)$$

- True value and particle trajectory for the time-varying parameter θ_0.





Comparison of α histograms:
- Before parameter change (<200th steps)
- After parameter change (>200th steps)



Comparisons of RMSE and Spread. Each box expresses the dispersion of the state and parameter elements averaged over the forecast and filtering steps in 20-1500.

References

1. Zhu, Mengbin, Peter Jan Van Leeuwen, and Javier Amezcua. "Implicit equal-weights particle filter." Quarterly Journal of the Royal Meteorological Society 142.698 (2016): 1904-1919.
2. Satoh, Mineto, Peter Jan van Leeuwen, and Shin'ya Nakano. "Online state and time-varying parameter estimation using the implicit equal-weights particle filter." Quarterly Journal of the Royal Meteorological Society (2024).