DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES B
Volume **20**, Number **5**, July **2015**

doi:10.3934/dcdsb.2015.20.1337

pp. **1337–1354**

ERROR ANALYSIS FOR NUMERICAL FORMULATION OF PARTICLE FILTER

Xiaoying Han

Department of Mathematics and Statistics
Auburn University
Auburn, AL 36849, USA

Jinglai Li

Institute of Natural sciences, Department of Mathematics
MOE Key Lab of Scientific and Engineering Computing
Shanghai JiaoTong University, 800 Dongchuan Rd
Minhang 200240, Shanghai, China

Dongbin Xiu

Department of Mathematics
Scientific Computing and Imagining Institute
The University of Utah
Salt Lake City, UT 84112, USA

Abstract. As an approximation of the optimal stochastic filter, particle filter is a widely used tool for numerical prediction of complex systems when observation data are available. In this paper, we conduct an error analysis from a numerical analysis perspective. That is, we investigate the numerical error, which is defined as the difference between the numerical implementation of particle filter and its continuous counterpart, and demonstrate that the error consists of discretization errors for solving the dynamic equations numerically and sampling errors for generating the random particles. We then establish convergence of the numerical particle filter to the continuous optimal filter and provide bounds for the convergence rate. Remarkably, our analysis suggests that more frequent data assimilation may lead to larger numerical errors of the particle filter. Numerical examples are provided to verify the theoretical findings.

1. **Introduction.** Assimilation of data into mathematical models is an essential task in almost all the areas of geophysics and beyond. Simply speaking, data assimilation is to estimate the optimal prediction that combines the output of the mathematical model, which is only an approximation of the real-world system, and the observations with measurement noise. Most of the traditional techniques of data assimilation, such as the Kalman filter [20, 4], are based on linear control theory and optimization, and their applications to highly nonlinear systems are usually challenging (which often require some linearization processes), and sometimes they can even fail [11, 16].

2010 *Mathematics Subject Classification.* Primary: 65C35; Secondary: 65C05.

Key words and phrases. Bayesian filer, data assimilation, particle filter, sequential Monte Carlo.

The work of Li is supported by NSFC grant 11301337. The work of Xiu is partially supported by AFOSR, DOE and NSF.

Particle filter (PF), also known as the sequential Monte Carlo (SMC) method, can deal with problems where strong nonlinearity is present, without any linearization. The basic idea of PF is the following. Suppose that the mathematical model is a nonlinear stochastic dynamical system, and our goal is to estimate the hidden states of the system by combining model predictions and noisy partial observations of the system. One can do this with the so-called Bayes filter (or the optimal filter), where the posterior probability density function (pdf) of the hidden states is estimated by the Bayes' rule recursively [7]. A difficulty here is that, in general, the posterior distribution does not admit an analytical form. Many approximation approaches have been proposed to address this problem, such as the extended Kalman filter [10]. As is mentioned earlier, this kind of approaches often become problematic when the system is highly nonlinear or the posterior distribution is strongly non-Gaussian. The PF method, on the contrast, approximates the posterior distribution with Monte Carlo sampling (hence its name SMC), without making assumptions of linearity on the dynamic model or of Gaussianity on the noise. Specifically, PF employs a number of independent random realizations called particles, sampled directly from the state space, to represent the posterior probability, and update the posterior by involving the new observations. The particle system is properly located, weighted, and propagated recursively by the Bayesian formula. Since its introduction, PF has been found profound applications in many areas, such as signal processing [1], economics [3], robotics [18], geophysics [19, 17], just to name a few. For detailed discussions on particle filter, see [14, 6, 2] and the references therein.

Considerable efforts have been devoted to analyzing the statistical error of PF and its convergence properties in that sense. In particular, weak convergence of the state estimates by particle filter to the estimates by optimal filter was established with a convergence rate of $O(1/\sqrt{M})$, where M is the number of particles ([6, 9]). Moreover, the convergence is uniform if the number of particles increases over time or the kernel of the particle filter is weakly dependent on the past. Such conditions require all particles lie in a compact support subset of the space ([5]) and can not always be satisfied in practice. In such cases, the error due to the inaccuracy of the particle-based approximation may grow quickly ([12]).

To the best of our knowledge, however, little attention has been paid to analyzing the errors of PF due to its numerical implementation. In fact, the state equations in almost all the systems must be solved numerically in practice, which introduce errors into the filtering process inevitably. In what follows we will refer to the implementation of PF with numerical schemes as the numerical particle filter (NPF). To better design and implement PF, it is helpful to not only understand the sampling errors but also the numerical ones.

In this paper, we conduct error analysis of NPF from a numerical analysis perspective. That is, we incorporate the numerical discretization errors, resulted from the numerical solution procedure for the state equations, into the analysis. We first establish the convergence of particle filter to the optimal filter by extending the earlier studies and explicitly incorporating the discretization errors. We then conduct analysis on the error bounds for the convergence rate. The numerical errors are analyzed in two types: local error, which is the error induced by one-step particle filter procedure; and global error, which is the cumulative effects of the local errors. The error bounds clearly establish the convergence of particle filter, in the presence

of both discretization errors and Monte Carlo sampling errors, and reveal an interesting result in that more frequent data assimilation may lead to larger numerical errors. Similar results have been found for the ensemble Kalman filter in [13].

The rest of the paper is arranged as follows. The Bayesian filter and PF are introduced in Section 2. The convergence of PF to the optimal filter in the weak sense is shown in Section 3, and the convergence rate is analyzed in in Section 4. Finally numerical examples are provided in Section 5 to examine the theoretical results.

2. Formulations of PF.

In this section we briefly review the formulation of particle filter for stochastic filtering problems.

2.1. State-space models.

Let us consider the following stochastic filtering problem in a dynamic state-space form:

$$\frac{d\mathbf{u}}{dt} = f(t,\mathbf{u};\mathbf{z}), \qquad \mathbf{u}(0,\mathbf{z}) = \mathbf{u}_0(\mathbf{z}) \tag{1a}$$

$$\mathbf{v} = g(t,\mathbf{u};\boldsymbol{\varepsilon}), \tag{1b}$$

where $\mathbf{u} \in \mathbb{R}^{n_u}$ denotes the state vector, $\mathbf{v} \in \mathbb{R}^{n_v}$ denotes the measurement vector, $\mathbf{z} \in \mathbb{R}^{n_z}$ is a random vector representing the uncertainties in the model, and $\boldsymbol{\varepsilon} \in \mathbb{R}^{n_\varepsilon}$ denotes the measurement error, which is mutually independent with $\mathbf{z}$. Often the problem can be formulated in a discrete manner as follows.

$$\mathbf{u}_{n+1} = f_n(\mathbf{u}_n,\mathbf{z}_n), \qquad \mathbf{u}_0 = \mathbf{u}_0(\mathbf{z}), \tag{2}$$

$$\mathbf{v}_n = g_n(\mathbf{u}_n,\boldsymbol{\varepsilon}_n), \tag{3}$$

where the subscript n denotes the functions evaluated at discrete time levels t_n, $n = 0,1,\ldots$, with $t_0 < t_1 < \cdots$. In many applications, the noises are assumed to be additive and (2) and (3) are often written in the special case

$$\mathbf{u}_{n+1} = f_n(\mathbf{u}_n) + \mathbf{z}_n, \qquad \mathbf{u}_0 = \mathbf{u}_0(\mathbf{z}), \tag{4}$$

$$\mathbf{v}_n = g_n(\mathbf{u}_n) + \boldsymbol{\varepsilon}_n, \tag{5}$$

with $\mathbf{z}_n$ and $\boldsymbol{\varepsilon}_n$ are mutually independent.

In this paper we adopt the general model (1a)-(1b). In data assimilation, the observation $\mathbf{v}$ arrives sequentially in time and the goal is to estimate the true state, denoted as $\mathbf{u}^t$, which is not predicted perfectly by (1a), based on the prediction by (1a) and the measurement (1b).

2.2. Bayesian optimal filter.

Let $\{U_n\}_{n\geq 0}$ be the state process and $\{V_n\}_{n\geq 1}$ be the measurement process and consider two general probabilistic state space models: dynamic model $U_n \sim K(\mathbf{u}_n|\mathbf{u}_{n-1})$ and measurement model $V_n \sim \rho(\mathbf{v}_n|\mathbf{u}_n)$. The dynamical model is Markovian such that any future $\mathbf{u}_n$ is independent of the past given the present $\mathbf{u}_{n-1}$:

$$p(\mathbf{u}_n|\mathbf{u}_{1:n-1},\mathbf{v}_{1:n-1}) = K(\mathbf{u}_n|\mathbf{u}_{n-1}), \tag{6}$$

and the measurements are conditionally independent given $\mathbf{u}_n$

$$p(\mathbf{v}_n|\mathbf{u}_{1:n},\mathbf{v}_{1:n-1}) = \rho(\mathbf{v}_n|\mathbf{u}_n). \tag{7}$$

Given prior distribution $p(\mathbf{u}_0)$ and the data $\mathbf{v}_{1:n}$, Bayesian optimal filter is to construct the distribution $p(\mathbf{u}_n|\mathbf{v}_{1:n})$ recursively in two stages: prediction and update.

Assume that the required pdf $p(\mathbf{u}_{n-1}|\mathbf{v}_{1:n-1})$ of previous time $n-1$ is available. The Chapman-Kolmogorov equation gives the prediction step of

$$p(\mathbf{u}_n|\mathbf{v}_{1:n-1}) = \int_{\mathbb{R}^{n_u}} K(\mathbf{u}_n|\mathbf{u}_{n-1}) \cdot p(\mathbf{u}_{n-1}|\mathbf{v}_{1:n-1}) d\mathbf{u}_{n-1}. \tag{8}$$

At time n, as measurement $\mathbf{v}_n$ becomes available, the prior distribution from (8) can then be updated via Bayes' rule

$$p(\mathbf{u}_n|\mathbf{v}_{1:n}) = \frac{1}{Z_n}\rho(\mathbf{v}_n|\mathbf{u}_n)p(\mathbf{u}_n|\mathbf{v}_{1:n-1}), \tag{9}$$

where the normalization constant $Z_n = p(\mathbf{v}_n|\mathbf{v}_{1:n-1})$ is given by

$$Z_n = \int_{\mathbb{R}^{n_u}} \rho(\mathbf{v}_n|\mathbf{u}_n)p(\mathbf{u}_n|\mathbf{v}_{1:n-1})d\mathbf{u}_n.$$

2.3. **Particle filter.** Particle filter is a numerical approximation to the optimal Bayesian filter and uses an empirical distribution of a set of random samples, called particles, to approximate the conditional distribution $p(\mathbf{u}_n|\mathbf{v}_{1:n})$. Throughout this paper we will use $\mathbf{u}^f$ to denote the solution of (1a), where the superscript indicates this is the so-called *forecast* state variables that are obtained by solving the governing equation (1a), and $\mathbf{u}^a$ to denote the *analyzed* state variables that are obtained by applying the Bayes' rule for update. We will also use $M \geq 1$ to denote the number of particles.

Let $0 = t_0 < t_1 < \cdots < t_n < \cdots$ be a sequence of discrete time levels where observation data are available and Bayesian update are made. Let $\{(\mathbf{u}_n^a)_i\}_{i=1}^M$ denote a set of updated particles at t_n with the empirical distribution $\pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n})$,

$$p(\mathbf{u}_n^a|\mathbf{v}_{1:n}) \approx \pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n}) \triangleq \frac{1}{M}\sum_{i=1}^M \delta(\mathbf{u}_n^a - (\mathbf{u}_n^a)_i), \tag{10}$$

where $\delta(\cdot)$ denotes the Dirac delta function satisfying $\delta(\mathbf{x}-\mathbf{a}) = 1$ at $\mathbf{x} = \mathbf{a}$ and 0 otherwise. The particle filter algorithm is recursive in time and can be summarized as:

- *Initialization:* At t_0, sample $\{(\mathbf{u}_0^a)_i\}_{i=1}^M \sim p(\mathbf{u}_0^a)$ as the initial set of particles.
- *Prediction:* At t_{n-1}, $n \geq 1$, let $\{(\mathbf{u}_{n-1}^a)_i\}_{i=1}^M$ be a set of particles distributed approximately according to $\pi^M(\mathbf{u}_{n-1}^a|\mathbf{v}_{1:n-1})$. We then forward each particle individually from t_{n-1} with initial condition $(\mathbf{u}_{n-1}^a)_i$ till t_n by solving the dynamic model (1a) in time, i.e., drawing independent samples according to

$$(\mathbf{u}_n^f)_i \sim K(\mathbf{u}_n|(\mathbf{u}_{n-1}^a)_i), \quad i = 1, \ldots, M. \tag{11}$$

  Forecast particles $\{(\mathbf{u}_n^f)_i\}_{i=1}^M$ are obtained with an empirical distribution

$$\pi^M(\mathbf{u}_n^f|\mathbf{v}_{1:n-1}) \triangleq \frac{1}{M}\sum_{i=1}^M \delta(\mathbf{u}_n^f - (\mathbf{u}_n^f)_i). \tag{12}$$

- *Update:* At t_n, $n \geq 1$, where new measurements $\mathbf{v}_n$ are available, the Bayes' formula (9) is applied and results in the following approximation of $p(\mathbf{u}_n|\mathbf{v}_{1:n})$

$$\hat{\pi}^M(\mathbf{u}_n^f|\mathbf{v}_{1:n}) \triangleq \frac{\rho(\mathbf{v}_n|\mathbf{u}_n^f)\pi^M(\mathbf{u}_n^f|\mathbf{v}_{1:n-1})}{\int_{\mathbb{R}^{n_u}} \rho(\mathbf{v}_n|\mathbf{u}_n^f)p(\mathbf{u}_n^f|\mathbf{v}_{1:n-1})d\mathbf{u}_n^f} = \sum_{i=1}^M w_n^i \delta(\mathbf{u}_n^f - (\mathbf{u}_n)_i^f),$$

  where

$$w_n^i = \frac{\rho(\mathbf{v}_n|(\mathbf{u}_n^f)_i)}{\sum_{i=1}^M \rho(\mathbf{v}_n|(\mathbf{u}_n^f)_i)}, \qquad 1 \leq i \leq M, \tag{13}$$

are the so-called importance weights. A resampling step is performed to obtain a set of equally weighted particles $\{(\mathbf{u}_n^a)_i\}_{i=1}^M$ from the distribution $\hat{\pi}^M(\mathbf{u}_n^f|\mathbf{v}_{1:n})$ such that

$$p(\mathbf{u}_n^a|\mathbf{v}_{1:n}) \approx \pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n}) = \frac{1}{M}\sum_{i=1}^{M}\delta(\mathbf{u}_n^a - (\mathbf{u}_n)_i^a). \tag{14}$$

3. **Numerical formulation of PF and weak convergence.** As is mentioned in Section 1, the convergence of particle filter to the optimal stochastic filter has been well established in terms of the sampling errors [5]. In practice, however, the governing state equation (1) is often too complicated to solve analytically. In this case, one usually employs numerical methods to solve the equation. Thus, to analyze the convergence of PF more accurately, one must take the errors due to numerical implementation into account as well. In this section, we introduce the numerical formulation of PF and show the week convergence of it to the ideal Bayes filter.

3.1. **Numerical formulation of PF.** Without loss of generality, we assume $\{t_n\}$, the time levels when data assimilation is conducted, are equally distributed with

$$\Delta T = t_{n+1} - t_n, \qquad \forall n \geq 0. \tag{15}$$

We further partition the time interval $[t_n, t_{n+1}]$ into m equal sized sub-intervals,

$$t_{n,j} = t_n + j \cdot \Delta t, \qquad j = 0, \ldots, m, \quad n = 0, 1, \ldots,$$

where $t_{n,0} = t_n$, $t_{n,m} = t_{n+1}$, and $\Delta t > 0$ is the step size with which a stable and accurate numerical scheme is employed to solve the (1a). The numerical scheme is forwarded in time on the stencil $t_{n,j}, 0 \leq j \leq m, n \geq 0$. For simplicity, we assume the scheme is a one-step method in the following form,

$$\widetilde{\mathbf{u}}_{n,j+1}^f = \widetilde{\mathbf{u}}_{n,j}^f + \Delta t \cdot \Phi(t_{n,j}, \widetilde{\mathbf{u}}_{n,j}^f; \Delta t), \qquad 0 \leq j < m, \quad n \geq 0, \tag{16}$$

where $\widetilde{\mathbf{u}}^f$ is the numerical solution of (1a) and $\Phi(\cdot)$ is an *increment function* satisfying the consistency condition

$$\lim_{\Delta t \to 0} \Phi(t_{n,j}, \mathbf{u}_{n,j}^f; \Delta t) = f(t_{n,j}, \mathbf{u}_{n,j}^f). \tag{17}$$

Subsequently, if we define $\xi_{\Delta t} = \max_n |\widetilde{u}_n^f - u_n^f|$, then

$$\lim_{\Delta t \to 0} \xi_{\Delta t} = 0. \tag{18}$$

The numerical scheme is said to have an order of q if $\xi_{\Delta t} \sim O(\Delta t^q)$.

The complete numerical implementation of particle filter, in which the set of the analyzed particles are denoted as $\{(\widetilde{\mathbf{u}}_n^a)_i\}_{i=1}^M$ with empirical distribution $\widetilde{\pi}^M(\widetilde{\mathbf{u}}_n^a|\mathbf{v}_{1:n})$, is obtained via the following recurrent procedure from t_{n-1} to t_n:

- At t_{n-1}, use a set of the particles $\{(\widetilde{\mathbf{u}}_{n-1}^a)_i\}_{i=1}^M$ as initial conditions. When $n = 0$, $\{(\widetilde{\mathbf{u}}_0^a)_i\}_{i=1}^M$ are drawn from $p(\mathbf{u}_0)$. The empirical distribution at t_{n-1} is

$$\widetilde{\pi}^M(\widetilde{\mathbf{u}}_{n-1}^a|\mathbf{v}_{1:n-1}) = \frac{1}{M}\sum_{i=1}^{M}\delta(\widetilde{\mathbf{u}}_{n-1}^a - (\widetilde{\mathbf{u}}_{n-1}^a)_i) \tag{19}$$

- For each particle $i = 1, \ldots, M$, solve the forecast model (1a) via the numerical scheme (16) forward in time till t_n, i.e.,

$$(\widetilde{\mathbf{u}}_n^f)_i = (\widetilde{\mathbf{u}}_{n-1}^f)_i + \Delta t \sum_{j=0}^{m-1} \Phi(t_{n-1,j}, (\widetilde{\mathbf{u}}_{n-1,j}^f)_i; \Delta t), \quad (\widetilde{\mathbf{u}}_{n-1,0}^f)_i = (\widetilde{\mathbf{u}}_{n-1}^a)_i. \tag{20}$$

  Then the weighted empirical distribution at t_n is

$$\hat{\widetilde{\pi}}_M(\widetilde{\mathbf{u}}_n^f|\mathbf{v}_{1:n}) = \sum_{i=1}^{M} w_n^i \delta(\widetilde{\mathbf{u}}_n^f - (\widetilde{\mathbf{u}}_n)_i^f), \tag{21}$$

  in which the importance weights $\{w_n^i\}_{i=1}^M$ are evaluated using (13).
- A new set of equally weighted particles $\{(\widetilde{\mathbf{u}}_n^a)_i\}_{i=1}^M$ are sampled from $\hat{\widetilde{\pi}}_M$, and result in an empirical distribution at t_n

$$\widetilde{\pi}^M(\widetilde{\mathbf{u}}_n^a|\mathbf{v}_{1:n}) = \frac{1}{M}\sum_{i=1}^{M} \delta(\widetilde{\mathbf{u}}_n^a - (\widetilde{\mathbf{u}}_n)_i^a). \tag{22}$$

The procedure is repeated till a desired final time level $T > 0$ is reached. The notation $\hat{\widetilde{\pi}}^M$ is chosen in such a way that the $\widetilde{\ }$ indicates numerical discretization errors for solving the forecast model (1a) via (16) are involved, the $\hat{\ }$ indicates a distribution is sampled by weighted particles.

3.2. **Weak convergence theory.** In general, given a measure μ and a function φ, we define

$$\langle \mu(\cdot), \varphi \rangle = \int \varphi(y)\mu(\cdot)dy. \tag{23}$$

The convergence of $\pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n})$ to $p(\mathbf{u}_n^a|\mathbf{v}_{1:n})$ has been well studied in the literature (e.g. [5]) in the sense that

$$\langle \pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n}), \varphi \rangle \overset{M\to\infty}{\longrightarrow} \langle p(\mathbf{u}_n^a|\mathbf{v}_{1:n}), \varphi \rangle. \tag{24}$$

Throughout the remainder of this article we will study the convergence of $\langle \widetilde{\pi}^M(\widetilde{\mathbf{u}}_n^a|\mathbf{v}_{1:n}), \varphi \rangle$ to $\langle p(\mathbf{u}_n^a|\mathbf{v}_{1:n}), \varphi \rangle$.

First of all, to guarantee that the Bayes' formula in (9) is well defined and can be fulfilled in PF algorithm, we assume the normalization constant Z_n satisfies

(A0) For given $\mathbf{v}_{1:s}$, $s = 1, \cdots, n$,

$$Z_s > \lambda > 0, \quad s = 1, \cdots, n.$$

We shall also assume that the conditional densities K and ρ are continuous, bounded and strictly positive:

(A1) $0 < K(\mathbf{u}_s|\mathbf{u}_{s-1}) < \infty, \quad 0 < \rho(\mathbf{v}_s|\mathbf{u}_s) < \infty, \quad$, for given $\mathbf{v}_{1:s}, s = 1, \cdots, n$.

For general function φ, we assume that

(A2) φ is continuous and satisfies

$$\sup_{\mathbf{u}_s} \left\{ \varphi^2(\mathbf{u}_s) \cdot \rho(\mathbf{v}_s|\mathbf{u}_s) \right\} < C_s,$$

for given $\mathbf{v}_{1:s}$, $s = 1, \cdots, n$, where C_s is a finite constant independent of $\mathbf{u}_{1:s}$.

Note that **(A1)** and **(A2)** imply the conditional second moment of φ is bounded, i.e.

$$\int \varphi^2(\mathbf{u}_n^a) p(\mathbf{u}_n^a|\mathbf{v}_{1:n}) d\mathbf{u}_n^a = \frac{\int \varphi^2(\mathbf{u}_n^f)\rho(\mathbf{v}_n|\mathbf{u}_n^f)p(\mathbf{u}_n^f|\mathbf{v}_{1:n})d\mathbf{u}_n^f}{Z_n} < \infty. \tag{25}$$

Following the work of [9], we denote the class of φ satisfies **(A2)** by $L_n^2(\rho)$ and define

$$\|\varphi\| \triangleq \max_{s=1,\cdots,n} \left\{ 1, \left(\int \varphi^2(\mathbf{u}_s) p(\mathbf{u}_s|\mathbf{v}_{1:s}) d\mathbf{u}_s \right)^{\frac{1}{2}} \right\}.$$

Let $\mathcal{P}(\mathbb{R}^{n_u})$ be the space of all probability measures over the n_u-dimensional Euclidean space $\mathbb{R}^{n_u}$. We define $b_n : \mathcal{P}(\mathbb{R}^{n_u}) \to \mathcal{P}(\mathbb{R}^{n_u})$ to be the mapping

$$b_n(\mu)(\mathbf{u}_n) \triangleq \int_{\mathbb{R}^{\mathbf{n_u}}} K(\mathbf{u}_n|\mathbf{u}_{n-1}) \mu(\mathbf{u}_{n-1}) d\mathbf{u}_{n-1} \tag{26}$$

for any $\mu \in \mathcal{P}(\mathbb{R}^{n_u})$. It is natural to assume that b_n is continuous, since in the context of filtering two realizations of the signal that start from "close" positions will remain "close" at subsequent times. By definition, we have $p(\mathbf{u}_n|\mathbf{v}_{1:n-1}) = b_n(p(\mathbf{u}_{n-1}|\mathbf{v}_{1:n-1}))$ and for $\varphi \in L_n^2(\rho)$

$$\langle b_n(\mu), \varphi \rangle = \int_{\mathbb{R}^{n_u}} \int_{\mathbb{R}^{n_u}} \varphi(\mathbf{u}_n) K(\mathbf{u}_n|\mathbf{u}_{n-1}) \mu(\mathbf{u}_{n-1}) d\mathbf{u}_{n-1} d\mathbf{u}_n. \tag{27}$$

Suppose μ^M is the empirical distribution of a set of particles $\{\mathbf{u}_i\}_{i=1}^M$ and $\widetilde{\mu}^M$ is the empirical distribution of $\{\widetilde{\mathbf{u}}_i\}_{i=1}^M$, which are the particles obtained by forwarding $\{\mathbf{u}_i\}_{i=1}^M$ via the numerical scheme (16). Define $\widetilde{b}_n : \mathcal{P}(\mathbb{R}^{n_u}) \to \mathcal{P}(\mathbb{R}^{n_u})$ to be the mapping

$$\widetilde{b}_n(\mu^M) = \widetilde{\mu}^M, \tag{28}$$

we then have $\widetilde{b}_n(\pi^M(\mathbf{u}_{n-1}|\mathbf{v}_{1:n-1})) = \widetilde{\pi}^M(\widetilde{\mathbf{u}}_n|\mathbf{v}_{1:n-1})$.

Lemma 3.1. *Let μ^M and $\widetilde{\mu}^M$ be the empirical distributions defined as above. Then*

$$\lim_{\Delta t \to 0} \langle \widetilde{b}_n(\mu^M), \varphi \rangle = \langle b_n(\mu^M), \varphi \rangle, \quad \forall \varphi \in L_n^2(\rho). \tag{29}$$

Proof. Let $\{\mathbf{u}_i^a\}_{i=1}^M$ be a set of particles with the empirical distribution μ^M. Let $\{\mathbf{u}_i^f\}_{i=1}^M$ be the particles obtained by forwarding $\{\mathbf{u}_i^a\}_{i=1}^M$ exactly via Eq. (1a) and $\{\widetilde{\mathbf{u}}_i^f\}_{i=1}^M$ be the particles obtained by forwarding $\{\mathbf{u}_i\}_{i=1}^M$ numerically. We then have

$$\langle b_n(\mu^M), \varphi \rangle = \int \varphi(\mathbf{u}^f) \left(\frac{1}{M} \sum_{i=1}^M \delta(\mathbf{u}^f - \mathbf{u}_i^f) \right) d\mathbf{u}^f = \frac{1}{M} \sum_{i=1}^M \varphi(\mathbf{u}_i^f)$$

and

$$\langle \widetilde{b}_n(\mu^M), \varphi \rangle = \int \varphi(\widetilde{\mathbf{u}}^f) \left(\frac{1}{M} \sum_{i=1}^M \delta(\widetilde{\mathbf{u}}^f - \widetilde{\mathbf{u}}_i^f) \right) d\widetilde{\mathbf{u}}^f = \frac{1}{M} \sum_{i=1}^M \varphi(\widetilde{\mathbf{u}}_i^f).$$

Subsequently,

$$\left| \langle b_n(\mu^M), \varphi \rangle - \langle \widetilde{b}_n(\mu^M), \varphi \rangle \right| \leq \frac{1}{M} \sum_{i=1}^M \left| \varphi(\mathbf{u}_i^f) - \varphi(\widetilde{\mathbf{u}}_i^f) \right|. \tag{30}$$

Because of (18), $\lim_{\Delta t \to 0} \widetilde{\mathbf{u}}_i = \mathbf{u}_i$, for $i = 1, \cdots, M$. And since φ is continuous, then given $\forall\ \epsilon > 0$, $\exists\ \Delta t$, such that when $|\mathbf{u}_i - \widetilde{\mathbf{u}}_i|$ is small enough, one has $|\varphi(\mathbf{u}_i) - \varphi(\widetilde{\mathbf{u}}_i)| < \epsilon$, for $i = 1, \cdots, M$. Denote $\epsilon_{\Delta t}$ to be the ϵ which can be reached with numerical time step Δt. Then (30) becomes

$$\left| \langle b_n(\mu^M), \varphi \rangle - \langle \widetilde{b}_n(\mu^M), \varphi \rangle \right| \leq \epsilon_{\Delta t}, \tag{31}$$

which completes the proof. □

Next we define $a_n : \mathcal{P}(\mathbb{R}^{n_u}) \to \mathcal{P}(\mathbb{R}^{n_u})$ to be the mapping

$$a_n(\mu)(\mathbf{u}_n^f) = \frac{\rho(\mathbf{v}_n|\mathbf{u}_n^f)\mu(\mathbf{u}_n^f)}{\int \rho(\mathbf{v}_n|\mathbf{u}_n^f)\mu(\mathbf{u}_n^f)d\mathbf{u}_n^f}, \quad \mu \in \mathcal{P}(\mathbb{R}^{n_u}). \tag{32}$$

It is also natural to assume that a_n is continuous, which means that a slight variation in two distributions will not result in a large variation in the distributions when the observations are taken into account. Note that for $\varphi \in L^2_n(\rho)$, we have $p(\mathbf{u}_n|\mathbf{v}_{1:n}) = a_n(p(\mathbf{u}_n|\mathbf{v}_{1:n-1}))$ and

$$\langle a_n(\mu), \varphi\rangle = \langle \mu, \rho\rangle^{-1}\langle \mu, \varphi\rho\rangle. \tag{33}$$

At last, we define c^M to a the resampling operator. The output of $c^M(\mu)$ is the empirical distribution of a sample of size M from a distribution μ. Let $c^{M,\omega}$, $M > 0$, $\omega \in \Omega$ be a realization of such a sampling such that

$$c^{M,\omega}(\mu) = \frac{1}{M}\sum_{i=1}^{M} \delta_{\{\Gamma_i(\omega)\}} \qquad \forall\ \mu \in \mathcal{P}(\mathbb{R}^{n_u}), \tag{34}$$

where $\Gamma_i : \Omega \to \mathbb{R}^{n_u}$ are i.i.d random variables with a common distribution μ. By Lemma 2 in [5], for almost all $\omega \in \Omega$, $c^{M,\omega}$ converges uniformly to the identity function.

Let us now consider $\pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n})$ and $\widetilde{\pi}^M(\mathbf{u}_n^a|\mathbf{v}_{1:n})$. It is easy to see that after the resampling step of PF, one has

$$\pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n}) = c^M \circ a_n \circ b_n \left(\pi^M(\mathbf{u}_{n-1}^a|\mathbf{v}_{1:n-1})\right)$$

and

$$\widetilde{\pi}^M(\widetilde{\mathbf{u}}_n^a|\mathbf{v}_{1:n}) = c^M \circ a_n \circ \widetilde{b}_n \left(\pi^M(\mathbf{u}_{n-1}^a|\mathbf{v}_{1:n-1})\right).$$

Therefore by Lemma 3.1, Theorem 1 of [5], and properties of composition of continuous functions, we obtain

$$\lim_{M\to\infty,\Delta t\to 0} \widetilde{\pi}^M(\widetilde{\mathbf{u}}_n^a|\mathbf{v}_{1:n}) = \lim_{M\to\infty} \pi^M(\mathbf{u}_n^a|\mathbf{v}_{1:n}) = p(\mathbf{u}_n^a|\mathbf{v}_{1:n}). \tag{35}$$

This implies that the convergence of numerical PF to optimal filter can be realized by increasing the number of particles and decreasing the size of the time step of the numerical scheme, as expected.

4. Numerical convergence rate.

In this section we estimate the convergence rate of the NPF to the optimal filter. Similar to traditional numerical analysis on ordinary differential equations, we break down the errors into two parts: *local error* and *global error*. Also instead of using weak convergence, we use the following convergence criterion:

Definition 4.1. let $\{\mu_M^\omega\}_{M=1}^\infty$ be a sequence of random probability measures, we say μ_M^ω converges to $\mu \in \mathcal{P}(\mathbb{R}^n)$, if

$$\lim_{M\to\infty} \mathbb{E}[(\langle \mu_M^\omega, \varphi\rangle - \langle \mu, \varphi\rangle)^2] = 0, \quad \forall \varphi \in \mathcal{P}(\mathbb{R}^n), \tag{36}$$

where the expectation is taken over all the realizations of the random sampling.

For notation convenience, we denote the compositions of a_n, b_n, $\widetilde{b}_n$ and c^M as the following three operators

$$r_n \triangleq a_n \circ b_n, \quad r_n^M \triangleq c^M \circ a_n \circ b_n, \quad \text{and} \quad \widetilde{r}_n^M \triangleq c^M \circ a_n \circ \widetilde{b}_n, \tag{37}$$

for $n = 0, 1, 2, \ldots$, and $a_0 \equiv b_0 \equiv \widetilde{b}_0$ are set to be the identity mapping.

4.1. **Local error.** Let us consider the time interval $[t_{n-1}, t_n], n \geq 1$. Let $p_{n-1|n-1} \triangleq p(\mathbf{u}_{n-1}|\mathbf{v}_{1:n-1})$ be the exact distribution of the state vector at t_{n-1}, produced by the optimal Bayesian filter. Using the definitions in (37), the exact distribution of state at t_n is

$$p_{n|n} = a_n \circ b_n(p_{n-1|n-1}) = r_n(p_{n-1|n-1}). \tag{38}$$

Let $\{(\mathbf{u}^a_{n-1})_i\}_{i=1}^M$ denote the particles drawn from the exact distribution of the state $p_{n-1|n-1}$ with empirical distribution $\pi^M_{n-1|n-1} \triangleq \pi^M(\mathbf{u}^a_{n-1}|\mathbf{v}_{1:n-1}) = c^M(p_{n-1|n-1})$. In one-step PF, they are used as initial condition of (1a).

Forwarding exactly, we obtain forecast particles $\{(\mathbf{u}^f_n)_i\}_{i=1}^M$ with empirical distribution

$$b^M_{n|n-1} \triangleq \pi^M(\mathbf{u}^f_n|\mathbf{v}_{1:n-1}) = b_n(\pi^M_{n-1|n-1}), \tag{39}$$

and then after Bayes' formula the new set of particle $\{(\mathbf{u}^a_n)_i\}_{i=1}^M$ are represented as

$$\nu^M_{n|n} = r^M_n(\pi^M_{n-1|n-1}).$$

Forwarding $\{(\mathbf{u}^a_{n-1})_i\}_{i=1}^M$ numerically by the numerical scheme (20) to obtain empirical distribution of $\{(\widetilde{\mathbf{u}}^f_n)_i\}_{i=1}^M$

$$\widetilde{b}^M_{n|n-1} \triangleq \widetilde{b}_n(\pi^M_{n-1|n-1})$$

and assimilating the measurement by the Bayes' rule, the empirical distribution of the state vectors at t_n, denoted as $\widetilde{\nu}^M_{n|n}$, can be represented as

$$\widetilde{\nu}^M_{n|n} = \widetilde{r}^M_n(\pi^M_{n-1|n-1}). \tag{40}$$

Definition 4.2. The local error of theoretical PF is the L^2-norm of the difference between $\nu^M_{n|n}$ and $p_{n|n}$

$$\bar{e}_n = (\mathbb{E}[(\langle \nu^M_{n|n}, \varphi\rangle - \langle p_{n|n}, \varphi\rangle)^2])^{\frac{1}{2}}, \quad \varphi \in L^2_n(\rho). \tag{41}$$

Define the local error of numerical PF to be the L^2-norm of the difference between $\widetilde{\nu}^M_{n|n}$ and $p_{n|n}$

$$e_n = (\mathbb{E}[(\langle \widetilde{\nu}^M_{n|n}, \varphi\rangle - \langle p_{n|n}, \varphi\rangle)^2])^{\frac{1}{2}}, \quad \varphi \in L^2_n(\rho). \tag{42}$$

Denote by d_n the local difference between $\nu^M_{n|n}$ and $\widetilde{\nu}^M_{n|n}$, i.e.,

$$d_n = (\mathbb{E}[(\langle \widetilde{\nu}^M_{n|n}, \varphi\rangle - \langle \nu^M_{n|n}, \varphi\rangle)^2])^{\frac{1}{2}}, \quad \varphi \in L^2_n(\rho). \tag{43}$$

Lemma 4.3. *[Local difference] Under assumptions* (A0) - (A2)*, there exists a constant C independent of M such that the local difference defined in* (43) *is bounded by*

$$d_n \leq \left(\frac{\|\varphi\| + 1}{\lambda}\right)\epsilon_{\Delta t} + \frac{2C^{1/2}\|\varphi\|}{\sqrt{M}}. \tag{44}$$

Proof. For later use, define the following intermediate distributions:

- $\hat{r}_{n|n} = a_n \circ b_n(\pi^M_{n-1|n-1})$ is the empirical distribution of $\{(\mathbf{u}^f_n)_i\}_{i=1}^M$ associated with $\{(w^i_n)_i\}_{i=1}^M$;
- $\hat{\widetilde{r}}_{n|n} = a_n \circ \widetilde{b}_n(\pi^M_{n-1|n-1})$ is the empirical distribution of $\{(\widetilde{\mathbf{u}}^f_n)_i\}_{i=1}^M$ associated with $\{(w^i_n)_i\}_{i=1}^M$;
- $b^M_{n|n-1} = b_n(\pi^M_{n-1|n-1})$ is the empirical distribution of $\{(\mathbf{u}^f_n)_i\}_{i=1}^M$ before update;

- $\widetilde{b}^M_{n|n-1} = \widetilde{b}_n(\pi^M_{n-1|n-1})$ is the empirical distribution of $\{(\widetilde{\mathbf{u}}^f_n)_i\}_{i=1}^M$ before update.

By splitting the local error, we have

$$\langle\widetilde{\nu}^M_{n|n},\varphi\rangle - \langle\nu^M_{n|n},\varphi\rangle = \langle\widetilde{\nu}^M_{n|n},\varphi\rangle - \langle\hat{\widetilde{r}}_{n|n},\varphi\rangle + \langle\hat{\widetilde{r}}_{n|n},\varphi\rangle - \langle\hat{r}_{n|n},\varphi\rangle + \langle\hat{r}_{n|n},\varphi\rangle - \langle\nu^M_{n|n},\varphi\rangle. \tag{45}$$

Then the Minkowski's inequality gives

$$\begin{aligned} d_n \leq & (\mathbb{E}[(\langle\widetilde{\nu}^M_{n|n},\varphi\rangle - \langle\hat{\widetilde{r}}_{n|n},\varphi\rangle)^2])^{\frac{1}{2}} + (\mathbb{E}[(\langle\hat{\widetilde{r}}_{n|n},\varphi\rangle - \langle\hat{r}_{n|n},\varphi\rangle)^2])^{\frac{1}{2}} \\ & + (\mathbb{E}[(\langle\hat{r}_{n|n},\varphi\rangle - \langle\nu^M_{n|n},\varphi\rangle)^2])^{\frac{1}{2}}. \end{aligned}$$

Note that the resampling procedure $\hat{r}_{n|n} \to \nu^M_{n|n}$ and $\hat{\widetilde{r}}_{n|n} \to \widetilde{\nu}^M_{n|n}$ satisfy

$$\mathbb{E}[\langle\nu^M_{n|n},\varphi\rangle] = \langle\hat{r}_{n|n},\varphi\rangle \quad \text{and} \quad \mathbb{E}[\langle\widetilde{\nu}^M_{n|n},\varphi\rangle] = \langle\hat{\widetilde{r}}_{n|n},\varphi\rangle. \tag{46}$$

therefore there exists a constant C such that

$$\left(\mathbb{E}[(\langle\nu^M_{n|n},\varphi\rangle - \langle\hat{r}^M_{n|n},\varphi\rangle)^2]\right)^{\frac{1}{2}} \leq \frac{C^{1/2}\|\varphi\|}{\sqrt{M}}, \tag{47}$$

and

$$\left(\mathbb{E}[(\langle\widetilde{b}^M_{n|n},\varphi\rangle - \langle\hat{\widetilde{r}}^M_{n|n},\varphi\rangle)^2]\right)^{\frac{1}{2}} \leq \frac{C^{1/2}\|\varphi\|}{\sqrt{M}}, \tag{48}$$

It then remains to consider $\langle\hat{\widetilde{r}}_{n|n},\varphi\rangle - \langle\hat{r}_{n|n},\varphi\rangle$. Recall definition of a_n,

$$\langle\hat{\widetilde{r}}_{n|n},\varphi\rangle - \langle\hat{r}_{n|n},\varphi\rangle = \frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle\widetilde{b}^M_{n|n-1},\rho\rangle} - \frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle} + \frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle} - \frac{\langle b^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle}$$

where

$$\begin{aligned} \left|\frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle\widetilde{b}^M_{n|n-1},\rho\rangle} - \frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle}\right| &= \frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle\left|\langle b^M_{n|n-1},\rho\rangle - \langle\widetilde{b}^M_{n|n-1},\rho\rangle\right|}{\langle\widetilde{b}^M_{n|n-1},\rho\rangle\langle b^M_{n|n-1},\rho\rangle} \\ &\leq \frac{\|\varphi\|}{\langle b^M_{n|n-1},\rho\rangle}\left|\langle b^M_{n|n-1},\rho\rangle - \langle\widetilde{b}^M_{n|n-1},\rho\rangle\right|. \end{aligned}$$

By using Minkowski's inequality again, we obtain

$$\begin{aligned} & \left(\mathbb{E}[(\langle\hat{\widetilde{b}}^M_{n|n},\varphi\rangle - \langle b^M_{n|n},\varphi\rangle)^2]\right)^{\frac{1}{2}} \\ \leq & \left(\mathbb{E}[(\frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle\widetilde{b}^M_{n|n-1},\rho\rangle} - \frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle})^2]\right)^{\frac{1}{2}} + \left(\mathbb{E}[(\frac{\langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle} - \frac{\langle b^M_{n|n-1},\rho\varphi\rangle}{\langle b^M_{n|n-1},\rho\rangle})^2]\right)^{\frac{1}{2}} \\ \leq & \frac{\|\varphi\|}{\lambda}\left(\mathbb{E}[(\langle b^M_{n|n-1},\rho\rangle - \langle\widetilde{b}^M_{n|n-1},\rho\rangle)^2]\right)^{\frac{1}{2}} + \frac{1}{\lambda}\left(\mathbb{E}[(\langle b^M_{n|n-1},\rho\varphi\rangle - \langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle)^2]\right)^{\frac{1}{2}} \end{aligned}$$

From Lemma 3.1, we have

$$\begin{aligned} \left|\langle b^M_{n|n-1},\rho\rangle - \langle\widetilde{b}^M_{n|n-1},\rho\rangle\right| &\leq \epsilon_{\Delta t}, \\ \left|\langle b^M_{n|n-1},\rho\varphi\rangle - \langle\widetilde{b}^M_{n|n-1},\rho\varphi\rangle\right| &\leq \epsilon_{\Delta t}. \end{aligned}$$

So

$$(\mathbb{E}[(\langle\hat{\widetilde{r}}_{n|n},\varphi\rangle - \langle\hat{r}_{n|n},\varphi\rangle)^2])^{\frac{1}{2}} \leq \left(\frac{\|\varphi\|+1}{\lambda}\right)\epsilon_{\Delta t}. \tag{49}$$

The proof is then established combining (47) , (48) and (49). □

It is obvious that the local error consists of two contributions: the numerical discretization error and the resampling error. If one further assumes the convergence of the numerical scheme (16) has order $q \geq 1$, and the resampling procedure is of order $O(M^{-1/2})$, then

$$e_n \sim O(\Delta t^q, M^{-1/2}), \quad \Delta t \to 0, \quad M \to \infty. \tag{50}$$

Theorem 4.4. *[Local error] Under assumptions* (A0) - (A2)*, there exists a constant* C^L *independent of* M *such that the local error defined in* (41) *is bounded by*

$$e_n \leq \left(\frac{\|\varphi\| + 1}{\lambda} \right) \epsilon_{\Delta t} + \frac{C^L \|\varphi\|}{\sqrt{M}}. \tag{51}$$

Proof. By definitions of e_n, $\bar{e}_n$ and d_n, it is obvious that

$$e_n \leq \bar{e}_n + d_n.$$

According to Theorem 2 of [5], there exists $c_{n|n}$ independent of M such that

$$\bar{e}_n \leq \sqrt{c_{n|n}} \frac{\|\varphi\|}{\sqrt{M}}.$$

Let $C^L = \sqrt{c_{n|n}} + 2C^{1/2}$, it then follows immediately from Lemma 4.3 that

$$e_n \leq \left(\frac{\|\varphi\| + 1}{\lambda} \right) \epsilon_{\Delta t} + \frac{C^L \|\varphi\|}{\sqrt{M}}.$$

□

4.2. **Global error.** Consider the time interval $[t_0, t_n]$, $n \geq 0$. Let $p_{n|n}$ be the exact distribution at time t_n produced by optimal Bayesian filter and denote by $\widetilde{\pi}^M_{n|n} \triangleq \widetilde{\pi}^M(\widetilde{\mathbf{u}}^a_n | \mathbf{v}_{1:n})$ the empirical distribution at t_n approximated by NPF. The global error at $t = t_n$ is defined as

$$E_n = \left(\mathbb{E}[(\langle \widetilde{\pi}^M_{n|n}, \varphi \rangle - \langle p_{n|n}, \varphi \rangle)^2] \right)^{\frac{1}{2}}, \quad \varphi \in L^2_n(\rho). \tag{52}$$

To study the global error, extra assumptions are required for a_n, b_n and $\widetilde{b}_n$.

(A3) Assume that a_n, b_n and $\widetilde{b}_n$ are uniformly Lipschitz continuous with Lipschitz constants Λ^a_n, Λ^b_n and $\widetilde{\Lambda}^b_n$ respectively, in the sense that for any μ_1, $\mu_2 \in \mathcal{P}(\mathbb{R}^{n_u})$,

$$\begin{aligned}
|\langle a_n(\mu_1), \varphi \rangle - \langle a_n(\mu_2), \varphi \rangle| &\leq \Lambda^a_n |\langle \mu_1, \varphi \rangle - \langle \mu_2, \varphi \rangle|; \\
|\langle b_n(\mu_1), \varphi \rangle - \langle b_n(\mu_2), \varphi \rangle| &\leq \Lambda^b_n |\langle \mu_1, \varphi \rangle - \langle \mu_2, \varphi \rangle|; \\
\left| \langle \widetilde{b}_n(\mu_1), \varphi \rangle - \langle \widetilde{b}_n(\mu_2), \varphi \rangle \right| &\leq \widetilde{\Lambda}^b_n |\langle \mu_1, \varphi \rangle - \langle \mu_2, \varphi \rangle|.
\end{aligned}$$

It is straightforward to see that under assumption (A3), $r_n = a_n \circ b_n$ is uniformly Lipschitz continuous with constant $\Lambda_n = \Lambda^a_n \cdot \Lambda^b_n$, and $\widetilde{r}_n$ is uniformly Lipschitz continuous with constant $\widetilde{\Lambda}_n = \Lambda^a_n \cdot \widetilde{\Lambda}^b_n$, i.e. $\forall \mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^{n_u})$, r_n and $\widetilde{r}_n$ satisfy respectively

$$|\langle r_n(\mu_1), \varphi \rangle - \langle r_n \langle \mu_2), \varphi \rangle| \leq \Lambda_n \cdot |\langle \mu_1, \varphi \rangle - \langle \mu_2, \varphi \rangle| \tag{53}$$

and

$$|\langle \widetilde{r}_n(\mu_1), \varphi \rangle - \langle \widetilde{r}_n(\mu_2), \varphi \rangle| \leq \widetilde{\Lambda}_n \cdot |\langle \mu_1, \varphi \rangle - \langle \mu_2, \varphi \rangle| \tag{54}$$

The global error is the cumulative effect of all the local errors from the initial time to the final time t_n and is expected to grow as time increases. In what follows we will derive error bounds of the global error in two similar and yet different forms.

Theorem 4.5. *[Global error: first bound] The global error defined in (52) can be bounded by*

$$E_n \leq \frac{\Lambda^n - 1}{\Lambda - 1} e, \tag{55}$$

where $\Lambda = \max_{1\leq s\leq n} \Lambda_s$ *and* $e = \max_{1\leq s\leq n} e_s$.

Proof. We first split the global error.

$$\begin{aligned}&\langle \widetilde{\pi}^M_{n|n}, \varphi\rangle - \langle p_{n|n}, \varphi\rangle = \langle \widetilde{\pi}^M_{n|n}, \varphi\rangle - \langle r_n(p_{n-1|n-1}), \varphi\rangle \\ =&\langle \widetilde{r}^M_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle - \langle r_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle + \langle r_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle - \langle r_n(p_{n-1|n-1}), \varphi\rangle.\end{aligned}$$

By using the Minkowski's inequality, we have

$$\begin{aligned} E_n \quad \leq \quad & \left(\mathbb{E}[(\langle \widetilde{r}^M_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle - \langle r_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle)^2]\right)^{\frac{1}{2}} \\ + \quad & \left(\mathbb{E}[(\langle r_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle - \langle r_n(p_{n-1|n-1}), \varphi\rangle)^2]\right)^{\frac{1}{2}}. \end{aligned} \tag{56}$$

The first term in (56) is the local error with the measure $\widetilde{\pi}^M_{n-1|n-1}$. Since Theorem 4.3 demonstrates that the local error is independent of the measure, we obtain

$$\left(\mathbb{E}[(\langle \widetilde{r}^M_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle - \langle r_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle)^2]\right)^{\frac{1}{2}} \leq e_n$$

and

$$E_n \leq e_n + \left(\mathbb{E}[(\langle r_n(\widetilde{\pi}^M_{n-1|n-1}), \varphi\rangle - \langle r_n(p_{n-1|n-1}), \varphi\rangle)^2]\right)^{\frac{1}{2}}.$$

Using the Lipschitz continuity of r_n recursively, we obtain

$$\begin{aligned} E_n \quad \leq \quad & e_n + \Lambda_n \left(\mathbb{E}[(\langle \widetilde{\pi}^M_{n-1|n-1}, \varphi\rangle - \langle p_{n-1|n-1}, \varphi\rangle)]\right)^{\frac{1}{2}} \\ \leq \quad & e_n + \Lambda_n \cdot e_{n-1} + \Lambda_n \cdot \Lambda_{n-1} \left(\mathbb{E}[(\langle \widetilde{\pi}^M_{n-2|n-2}, \varphi\rangle - \langle p_{n-2|n-2}, \varphi\rangle)^2]\right)^{\frac{1}{2}} \\ \leq \quad & \cdots \leq \sum_{s=1}^{n} \Lambda^{n-s} e_s \leq \left(\sum_{s=1}^{n} \Lambda^{n-s}\right) e = \frac{\Lambda^n - 1}{\Lambda - 1} e. \end{aligned}$$

□

Alternatively, we define a one-step *global increment function* Ψ for the particle filter via

$$\Psi(\mu)(t_{n-1}, \varphi, \rho, K; \Delta t, \Delta T) = \frac{1}{\Delta T}\left(\langle \widetilde{r}^M_n(\mu), \varphi\rangle - \langle \mu, \varphi\rangle\right), \quad n = 1, 2, \cdots. \tag{57}$$

Lemma 4.6. *Assume (A0) - (A3) hold, then for any* $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^{n_u})$*, the global increment function* Ψ *satisfies*

$$\left(\mathbb{E}\left[(\Psi(\mu_1) - \Psi(\mu_2))^2\right]\right)^{\frac{1}{2}} \leq \frac{\widetilde{\Lambda}_n + 1}{\Delta T} \cdot \left(\mathbb{E}[((\mu_1, \varphi) - (\mu_2, \varphi))^2]\right)^{\frac{1}{2}} + \frac{2\sqrt{C}}{\Delta T} \cdot \frac{\|\varphi\|}{\sqrt{M}}. \tag{58}$$

Proof. From the definition in (57), we have

$$\begin{aligned}\Psi(\mu_1)-\Psi(\mu_2)=&\frac{1}{\Delta T}\left(\langle\widetilde{r}_n^M(\mu_1),\varphi\rangle-\langle\widetilde{r}_n(\mu_1),\varphi\rangle\right)+\frac{1}{\Delta T}\left(\langle\widetilde{r}_n(\mu_1),\varphi\rangle-\langle\widetilde{r}_n(\mu_2),\varphi\rangle\right)\\&+\frac{1}{\Delta T}\left(\langle\widetilde{r}_n(\mu_2),\varphi\rangle-\langle\widetilde{r}_n^M(\mu_2),\varphi\rangle\right)+\frac{1}{\Delta T}\left(\langle\mu_2,\varphi\rangle-\langle\mu_1,\varphi\rangle\right),\end{aligned}$$

and therefore

$$\begin{aligned}\mathbb{E}[(\Psi(\mu_1)-\Psi(\mu_2)^2)]^{\frac{1}{2}}\quad\leq\quad&\frac{1}{\Delta T}\mathbb{E}[(\langle\widetilde{r}_n^M(\mu_1),\varphi\rangle-\langle\widetilde{r}_n(\mu_1),\varphi\rangle)^2]^{\frac{1}{2}}\\&+\frac{1}{\Delta T}\mathbb{E}[(\langle\widetilde{r}_n(\mu_1),\varphi\rangle-\langle\widetilde{r}_n(\mu_2),\varphi\rangle)^2]^{\frac{1}{2}}\\&+\frac{1}{\Delta T}\mathbb{E}[(\langle\widetilde{r}_n(\mu_2),\varphi\rangle-\langle\widetilde{r}_n^M(\mu_2),\varphi\rangle)^2]^{\frac{1}{2}}\\&+\frac{1}{\Delta T}\mathbb{E}[(\langle\mu_2,\varphi\rangle-\langle\mu_1,\varphi\rangle)^2]^{\frac{1}{2}}.\end{aligned}\tag{59}$$

By convergence of $\widetilde{r}_n^M$ to $\widetilde{r}_n$, we have that

$$\mathbb{E}[(\langle\widetilde{r}_n^M(\mu_1),\varphi\rangle-\langle\widetilde{r}_n(\mu_1),\varphi\rangle)^2]^{\frac{1}{2}}\leq\sqrt{C}\frac{\|\varphi\|}{M},\tag{60}$$

and

$$\mathbb{E}[(\langle\widetilde{r}_n(\mu_2),\varphi\rangle-\langle\widetilde{r}_n^M(\mu_2),\varphi\rangle)^2]^{\frac{1}{2}}\leq\sqrt{C}\frac{\|\varphi\|}{M}.\tag{61}$$

By (54) we have that

$$\mathbb{E}[(\langle\widetilde{r}_n(\mu_1),\varphi\rangle-\langle\widetilde{r}_n(\mu_2),\varphi\rangle)^2]^{\frac{1}{2}}\leq\widetilde{\Lambda}_n\mathbb{E}[(\langle\mu_2,\varphi\rangle-\langle\mu_1,\varphi\rangle)^2]^{\frac{1}{2}}.\tag{62}$$

The statement then follows immediately by collecting (59) through (62). □

Theorem 4.7. *[Global error: second bound] Assume (A0) - (A3) hold, then the global error can be bounded by*

$$E_n\leq\left(E_0+\sum_{s=1}^{n}e'_s\right)\cdot exp\left\{C^G\cdot t_n\right\}\tag{63}$$

where $e'_s=e_s+\sqrt{C}\cdot\|\varphi\|/\sqrt{M}$ *for* $s=1,\ldots,n$, *and* $C^G=\max_{1\leq s\leq n}\left\{(\widetilde{\Lambda}_s+1)/\Delta T\right\}$.

Proof. By (57), on any interval $[t_{n-1},t_n]$, the numerical particle filter satisfies

$$\langle\widetilde{\pi}_{n|n}^M,\varphi\rangle=\langle\widetilde{\pi}_{n-1|n-1}^M,\varphi\rangle+\Delta T\cdot\Psi(\widetilde{\pi}_{n-1|n-1}^M),\quad n\geq 1.\tag{64}$$

On the other hand, the exact solution of particle filter satisfies

$$\langle p_{n|n},\varphi\rangle=\langle p_{n-1|n-1},\varphi\rangle+\Delta T\cdot\Psi(p_{n-1|n-1})+e_n,\quad n\geq 0.\tag{65}$$

By subtracting (64) from (65), we obtain

$$\begin{aligned}&\langle p_{n|n},\varphi\rangle-\langle\widetilde{\pi}_{n|n}^M,\varphi\rangle\\=&(\langle p_{n-1|n-1},\varphi\rangle-\langle\widetilde{\pi}_{n-1|n-1}^M,\varphi\rangle)+\Delta T\cdot[\Psi(p_{n-1|n-1})-\Psi(\widetilde{\pi}_{n-1|n-1}^M)]+e_n\end{aligned}$$

Again by using the Minkowski's inequality, we obtain

$$\begin{aligned}E_n\quad\leq\quad&(\mathbb{E}[(\langle p_{n-1|n-1},\varphi\rangle-\langle\widetilde{\pi}_{n-1|n-1}^M,\varphi\rangle)^2])^{\frac{1}{2}}\\&+\Delta T\cdot(\mathbb{E}[(\Psi(p_{n-1|n-1})-\Psi(\widetilde{\pi}_{n-1|n-1}^M))^2])^{\frac{1}{2}}+e_n\\=\quad&E_{n-1}+\Delta T\cdot\left(\mathbb{E}[(\Psi(p_{n-1|n-1})-\Psi(\widetilde{\pi}_{n-1|n-1}^M))^2]\right)^{\frac{1}{2}}+e_n.\end{aligned}\tag{66}$$

By applying the formula recursively, we obtain

$$\begin{aligned} E_n &\leq E_0 + \sum_{s=1}^{n} e_s + \Delta T \cdot \sum_{s=1}^{n} \left(\mathbb{E}[(\Psi(p_{s-1|s-1}) - \Psi(\widetilde{\pi}^M_{s-1|s-1}))^2] \right)^{\frac{1}{2}} \\ &\leq E_0 + \sum_{s=1}^{n} e'_s + \Delta T \cdot C^G \sum_{s=1}^{n} \left(\mathbb{E}[(\langle p_{s-1|s-1}, \varphi \rangle - \langle \widetilde{\pi}^M_{s-1|s-1}, \varphi \rangle)^2] \right)^{\frac{1}{2}} \\ &\leq E_0 + \sum_{s=1}^{n} e'_s + \Delta T \cdot C^G \sum_{s=1}^{n} E_s. \end{aligned}$$

After applying the discrete Gronwall (see [15]), the statement (63) follows. □

Note that the constant C^G in Th. 4.5 is inversely proportional to ΔT, which, after being substituted into Eq. (63), implies that the bound depends exponentially on n. Although the two bounds are in different form, they both indicate that as the number of grid points n increases, the global numerical error grows. In other words, the more frequent data assimilation via smaller ΔT will result in larger global error. It is intuitively desirable to use more frequent data assimilation, provided that sufficient observation data are available, in the hope that the assimilated results will be closer to the "true state". However, the numerical analysis here suggests that such an action will incur larger numerical errors. Therefore, the proper choice of the size of assimilation step should be a balanced choice.

5. Numerical examples.

In this section, we present numerical tests to examine the theoretical results. Since the purpose is to verify the error analysis, we employ two simple benchmark problems whose "exact" optimal filtering solutions can be obtained. We then compare the numerical particle filter solutions against the exact solutions and examine the error convergence with respect to various parameter settings. In both examples we use the first four moments of the solutions and examine the numerical errors in them. Note that whether particle filter is the "best" filter for these benchmark problems is a different topic that is not a concern of this paper. Also, particle filter has been widely applied to many complex systems, where its performance and applicability have been examined extensively. We refer the interested readers to the large amount of available literature for applications of complex systems. Finally, since numerical error associated with solving the dynamic equations is usually well understood, we solve all the dynamic equations with sufficient accuracy so that the time discretization error is subdominant.

5.1. Linear Gaussian model.

We first consider a linear dynamic equation with initial values following a Gaussian distribution:

$$\frac{du^f}{dt} = au^f + b, \quad u^f(0) \sim \mathcal{N}(0, \sigma^2). \tag{67}$$

For simplicity, we fix $a = 0$, $b = 1$, and $\sigma = 1$. We construct a true state u^t, which is unavailable to the simulation, by adding an error following $\mathcal{N}(0, 1)$ to the mean solution of (67). Measurements are then made on u^t at every $\Delta T = 0.1$ time unit with error following $\mathcal{N}(0, 1)$. For this simple linear system with Gaussian noise, analytical solution of the optimal stochastic filter can be obtained. (In fact the optimal filter in this case becomes the well known Kalman filter, which can also be computed analytically.) We define error as the difference between the results obtained by the numerical particle filter and the exact solutions of the optimal

filter. (Note this is not the difference between the numerical particle filter solutions and the true states.)

In Fig. 1 the time evolution of the errors in the first four moments of the assimilated solutions are shown. It is clear that the numerical errors accumulate in time, consist with the error analysis on global error.

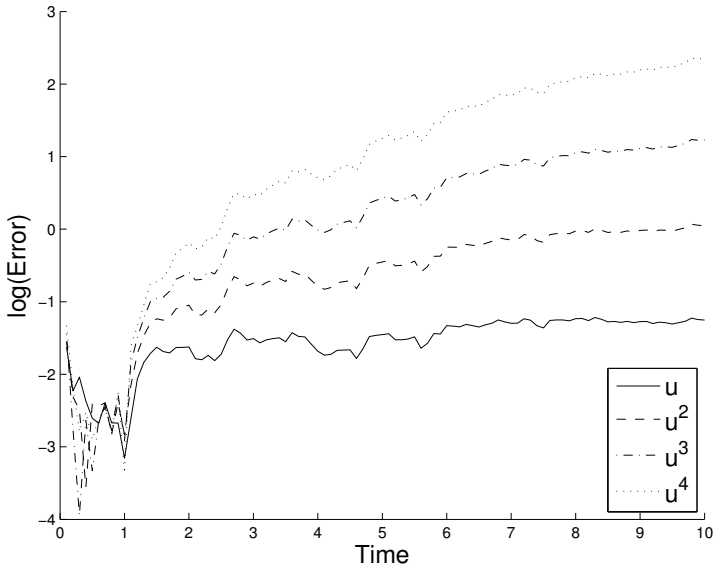


FIGURE 1. Time evolution of numerical errors in the first four moments.

Next we examine the error dependence on various parameters. In Fig. 2, the error convergence of the first four moments of the analyzed solution is shown against the size of the particle ensemble (M). The rate of convergence is approximately $1/2$, which is the rate of convergence of the resampling procedure and consistent with the analysis. In Fig. 3, we examine the error dependence against the size of the assimilation step (ΔT). Despite some oscillations due to the random sampling nature of the method, it is clear that the errors in the first four moments depend on the size of ΔT inversely. That is, smaller assimilation step ΔT results in larger numerical errors. This is consistent with the estimate in Theorem 4.7, where the Lipschitz constant becomes larger with smaller ΔT according to Lemma 4.6.

5.2. **Nonlinear population equation.** Here we consider the following population equation

$$\frac{du^f}{dt} = -r(1 - \frac{u^f}{A})u^f, \qquad u^f(0) \sim \mathcal{N}(a, 1), \tag{68}$$

where r and A are positive real parameters. The solution of (68) is sensitive to the initial condition. If $u^f(0) > A$, the solution will grow exponentially; if $0 < u^f(0) < A$, the solution will converge to 0.

We fix $r = 1$, $A = 2$, $a = 2.1$, and use the deterministic solution in time interval $[0, 1]$ of (68) with initial condition $u^f(0) = a$ as the "true state" (but unknown to the simulations). Measurements are made at every $\Delta T = 0.1$ time unit by adding

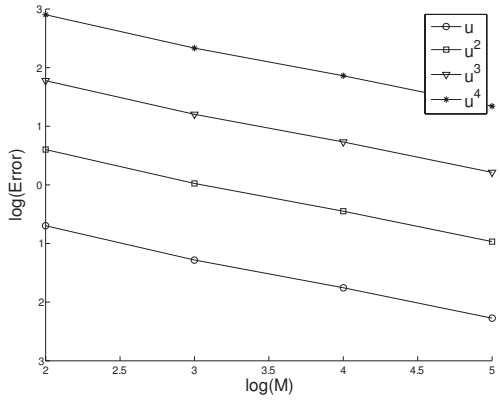


FIGURE 2. Error convergence of particle filter solutions with increased particle size M.

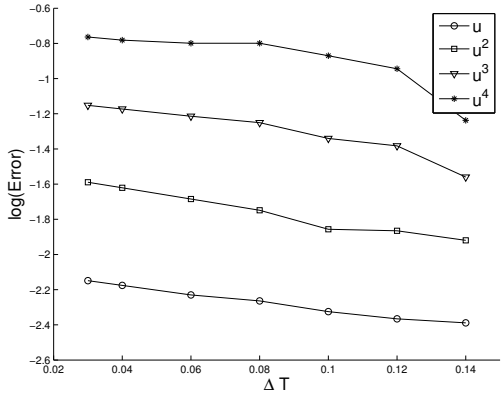
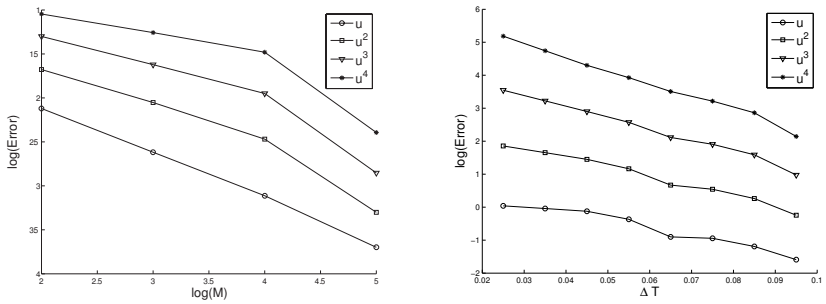


FIGURE 3. Error dependence on the assimilation time step ΔT ($M = 10^5$).

measurement errors following $\mathcal{N}(0, 0.3^2)$ on top of u^t. For this relatively simple but nonlinear problem, no explicit formulas exist for the optimal filter. We thus employ the numerical estimates by particle filter with very large ensemble size of $M = 10^6$ as the (numerically) "exact" solution and compare the errors obtained by particle filter at smaller sample size.

Again we examine the errors in the first four moments of the numerical particle filter solutions. Fig. 4 (a) illustrates the convergence property of the errors as the size of particles increases (at $T = 1$), where the $1/2$ convergence rate is visible. Fig. 4 (b) shows the dependence of the errors on the size of the assimilation step ΔT at $T = 3$. Again the inverse error dependence on ΔT is obvious.

6. **Conclusion.** In this paper, we have conducted a rigorous analysis on the numerical errors of PF. After establishing convergence of the NPF to the optimal stochastic filter, we provided estimate on the convergence rate. The results indicate



(a) (b)

FIGURE 4. (a) Error convergence of particle filter with respect to increased number of particles M, (b)Error dependence on the assimilation time step ΔT ($M = 10^4$).

that the errors grow over time and consist of the discretization errors of solving the state equations the (re)sampling errors for particle generation. Though more accurate numerical procedure for the state equations and larger sample size can reduce the errors, more frequent data assimilation will, however, result in larger numerical errors. This finding implies that in practical simulations even though it is preferred to use more frequent data assimilation whenever (reliable) measurements are available to track the true states more closely, one needs to be mindful of the accumulation of numerical errors in this case. Therefore, choice of assimilation step size should be a balanced issue.

REFERENCES

[1] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking, *IEEE Tran. Signal Process.,* **50** (2002), 174–188.

[2] O. Cappé, S. J. Godsill and E. Moulines, *An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo,* Proc. IEEE 95, May 2007.

[3] S. Chib, F. Nardari and N. Shephard, Markov chain Monte Carlo methods for stochastic volatility models, *J. Econometr.,* **108** (2002), 281–316.

[4] S. E. Cohn, An introduction to estimation theory, *J. Meteor. Soc. Jpn.,* **75** (1997), 257–288.

[5] D. Crisan and A. Doucet, A survey of convergence results on particle filtering for practitioners, *IEEE Trans. Signal Process.,* **50** (2002), 736–746.

[6] A. Doucet, N. Defreitas and N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[7] A Doucet, S Godsill and C Andrieu, On sequential Monte Carlo sampling methods for Bayesian filtering, *Stat. & Comput.,* **10** (2000), 197–208.

[8] Y. Ho and R. Lee, A Bayesian approach to problems in stochastic estimation and control, *IEEE Tran. Auto. Control*, **9** (1964), 333–339.

[9] X. Hu, T. B. Schön and L. Ljung, A basic convergence result for particle filtering, *IEEE Tran. Signal Process.,* **56** (2008), 1337–1348.

[10] A. H. Jazwinski, *Stochastic Processes and Filtering Theory,* Academic Press, 1970.

[11] L. Kuznetsov, K. Ide and C. K. R. T. Jones, A method for assimilation of Lagrangian data, *Mon. Wea. Rev.,* **131** (2003), 2247–2260.

[12] C. Lemieux, D. Ormoneit and D. J. Fleet, *Lattice Particle Filters,* Proc. 17th Ann. Conf. UAI, 2002.

[13] J. Li and D. Xiu, On numerical properties of the ensemble kalman filter for data assimilation, *Comput. Methods Appl. Mech. Engrg.,* **197** (2008), 3574–3583.

[14] J. Liu and R. Chen, Sequential Monte Carlo methods for dynamic systems, *J. Am. Stat. Assoc.,* **93** (1998), 1032–1044.
[15] A. Quarteroni, R. Sacco and F. Saleri, *Numerical Mathematics*, Springer-Verlag, New York, 2000.
[16] H. Salman, L. Kuznetsov, C. K. R. T. Jones and K. Ide, A method for assimilating Lagrangian data into a shallow-water equation ocean model, *Mon. Wea. Rev.,* **134** (2006), 1081–1101.
[17] E. T. Spiller, A. Budhirajab, K. Ide and C. K. R. T. Jones, Modified particle filter methods for assimilating Lagrangian data into a point-vortex model, *Phys. D,* **237** (2008), 1498–1506.
[18] S. Thrun, *Particle Filters in Robotics,* Proc. 17th Ann. Conf. UAI, 2002.
[19] P. Van Leeuwen, Particle filtering in geophysical systems, *Mon. Wea. Rev.,* **137** (2009), 4089–4114.
[20] G. Welch and G. Bishop, *An introduction to the Kalman filter,* Tech. Rep. TR95-041.

Received December 2013; revised January 2015.

E-mail address: xzh0003@auburn.edu
E-mail address: jinglaili@sjtu.edu.cn
E-mail address: dongbin.xiu@utah.edu