Effective filtering on a random slow manifold*

Huijie Qiao¹, Yanjie Zhang² and Jinqiao Duan³,⁴

¹ School of Mathematics, Southeast University, Nanjing, Jiangsu 211189, People’s Republic of China
² Center for Mathematical Sciences & School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, People’s Republic of China
³ Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, United States of America
⁴ Center for Mathematical Sciences, Huazhong University of Science and Technology, Wuhan 430074, People’s Republic of China

E-mail: hjqiaoge@seu.edu.cn, d201477019@hust.edu.cn and duan@iit.edu

Received 15 December 2017, revised 21 June 2018
Accepted for publication 12 July 2018
Published 31 August 2018

Abstract
This work is about a slow-fast data assimilation system under Gaussian noisy fluctuations. First, we obtain its low dimensional reduction via an invariant slow manifold. Second, we prove that the low dimensional filter on the slow manifold approximates the original filter in a suitable metric. Finally, we illustrate this approximate filter numerically in an example.

Keywords: multiscale systems, random slow manifolds, nonlinear filtering, dimension reduction, efficient filtering
Mathematics Subject Classification numbers: 60H10; 37D10, 70K70

(Some figures may appear in colour only in the online journal)

1. Introduction

Stochastic dynamical systems evolving on multiple time scales arise widely in engineering and science. For example, dynamics of chemical reaction networks often take place on notably different time scales, from the order of nanoseconds to the order of several days. The approximation by two time scales is common in various situations. This is especially true for

* This work was supported by NSF of China (No. 11001051, 11371352).
gene regulatory networks [11, 16], as the mRNA synthesis process is significantly faster than the protein dynamics, and this leads to a two-time-scale system [18].

Treating stochastic differential equations with two-time scales, Khasminskii and Yin [17] developed a stochastic averaging principle that enables one to average out the fast-varying variables. The main idea is as follows: under appropriate conditions, with the slow-varying component ‘fixed’, if the fast-varying component has a stationary distribution, it can be shown that the process represented by the slow-changing component converges weakly to a limit averaging system.

For random dynamical systems generated by stochastic differential equations with two-time scales, the theory of invariant manifolds provides another approach for qualitative analysis of dynamical behaviors, as invariant manifolds are geometric structures to describe or reduce stochastic dynamics [6, 19, 21]. Under suitable conditions, Fu–Liu–Duan [6] obtained low dimensional reduction of stochastic evolutionary equations with two-time scales via random slow invariant manifolds.

Filtering is a procedure to extract system state information with the help of observations (see [20]). The state evolution and the observations are usually under noisy fluctuations. The general idea is to achieve the best estimate for the true system state, given only noisy observations for the system. It provides an algorithm for estimating a signal or state of a random dynamical system based on noisy measurements. Stochastic filtering is important in many practical applications, from inertial guidance of aircrafts and spacecrafts to weather and climate prediction. Filtering problems for systems with two-time scales have been studied, with help of stochastic averaging (see [9, 12–15] and references therein).

The goal of this present paper is to investigate filtering for stochastic differential equations with slow and fast time scales. First, we obtain a low-dimensional reduced system on a random slow manifold, as in [6]. Then we prove that the filter of the low dimensional reduced system converges to the original filter in an appropriate sense, and this will be numerically illustrated in an example. Note that Gottwald–Harlim [7] also studied a linear slow-fast system, obtained a reduced system on a random slow manifold and compared two one-dimensional Kalman filters of the origin system and the reduced system. Here our result on nonlinear systems also provides the order of convergence for the filter of the low dimensional reduced system to the original filter.

It is worthwhile to mention that our assumption conditions and method are different from those in available literature on nonlinear filtering problems for stochastic differential equations with slow and fast time scales. On one hand, existence of random slow manifolds need some special conditions. On the other hand, on random slow manifolds these original stochastic differential equations have no Markov property. That means that some techniques, such as the Zakai equations in [13–15] and backward stochastic differential equations in [9], do not work. Therefore, we make use of an exponential martingale technique to deal with these nonlinear filtering problems.

This paper is organized as follows. In section 2, we recall basic concepts in random dynamical systems and random invariant manifolds. The framework for our method for reduced filtering is presented in section 3. In section 4, we present the nonlinear filtering problem and prove the approximation theorem of the filtering. And then a specific example is tested to illustrate our method in section 5. Finally, we summarize our work in section 6.

The following convention will be used throughout the paper: $C$ with or without indices will denote different positive constants (depending on the indices) whose values may change from one place to another.
2. Preliminaries

In the section, we introduce some notations and basic concepts in random dynamical systems.

2.1. Notation and terminology

\( \mathcal{B}(\mathbb{R}^n) \) stands for the Borel \( \sigma \)-algebra on \( \mathbb{R}^n \) and \( \mathcal{B}(\mathbb{R}^n) \) is the set of all real-valued uniformly bounded Borel-measurable functions on \( \mathbb{R}^n \). Let \( C(\mathbb{R}^n) \) denote the set of all real-valued continuous functions on \( \mathbb{R}^n \), and \( C^1_b(\mathbb{R}^n) \) denote the collection of all functions of \( C(\mathbb{R}^n) \) which themselves and their first-order derivatives are uniformly bounded. We introduce the following norm for \( \phi \in C^1_b(\mathbb{R}^n) \):

\[
\| \phi \| = \max_{x \in \mathbb{R}^n} |\phi(x)| + \max_{x \in \mathbb{R}^n} |\nabla \phi(x)|,
\]

where \( \nabla \) stands for the gradient operator. Moreover, \( C^\infty_c(\mathbb{R}^n) \) is the collection of all members of \( C(\mathbb{R}^n) \) with continuous derivatives of all orders and with compact support.

2.2. Random dynamical systems [1]

**Definition 2.1.** Let \( (\Omega, \mathcal{F}, P) \) be a probability space, and \( (\theta_t)_{t \in \mathbb{R}} \) a family of measurable transformations from \( \Omega \) to \( \Omega \). We call \( (\Omega, \mathcal{F}, P; (\theta_t)_{t \in \mathbb{R}}) \) a metric dynamical system if for each \( t \in \mathbb{R} \), \( \theta_t \) preserves the probability measure \( P \), i.e.

\[
\theta_t^* P = P,
\]

and for \( s, t \in \mathbb{R} \),

\[
\theta_0 = 1_{\Omega}, \quad \theta_{t+s} = \theta_t \circ \theta_s.
\]

**Definition 2.2.** Let \( (X, \mathcal{X}) \) be a measurable space. A mapping

\[
\varphi : \mathbb{R} \times \Omega \times X \mapsto X, \quad (t, \omega, x) \mapsto \varphi_t(\omega, x)
\]

with the following properties is called a measurable random dynamical system, or in short, a cocycle:

(i) Measurability: \( \varphi \) is \( \mathcal{B}(\mathbb{R}) \otimes \mathcal{F} \otimes \mathcal{X}/\mathcal{X} \)-measurable.

(ii) Cocycle property: \( \varphi(t, \omega) \) is continuous for \( t \in \mathbb{R} \), and further satisfies the following conditions

\[
\varphi(0, \omega) = \text{id}_X, \tag{1}
\]

\[
\varphi(t+s, \omega) = \varphi(t, \theta_s \omega) \circ \varphi(s, \omega), \tag{2}
\]

for all \( s, t \in \mathbb{R} \) and \( \omega \in \Omega \).

2.3. Random invariant manifolds [4, 21]

Let \( \varphi \) be a random dynamical system on the normed space \( (X, \| \cdot \|_X) \). Then we introduce a random invariant manifold with respect to \( \varphi \).
A family of nonempty closed sets $\mathcal{M} = \{\mathcal{M}(\omega)\}_{\omega \in \Omega}$ is called a random set if for every $y \in \mathbb{X}$, the mapping
\[
\Omega \ni \omega \to \text{dist}(y, \mathcal{M}(\omega)) := \inf_{x \in \mathcal{M}(\omega)} \|x - y\|_{\mathbb{X}}
\]
is measurable. $\mathcal{M}$ is called (positively) invariant with respect to $\varphi$ if
\[
\varphi(t, \omega, \mathcal{M}(\omega)) \subset \mathcal{M}(\theta_t \omega), \quad t \geq 0, \quad \omega \in \Omega.
\]

In the sequel, we consider random sets defined by a Lipschitz continuous graph. Define a function by
\[
\Omega \times \mathbb{R}^n \ni (\omega, x) \to H(\omega, x) \in \mathbb{R}^m
\]
such that for all $\omega \in \Omega$, $H(\omega, x)$ is globally Lipschitzian in $x$ and for any $x \in \mathbb{R}^n$, the mapping $\omega \to H(\omega, x)$ is a random variable. Then
\[
\mathcal{M}(\omega) := \{(x, H(\omega, x)) | x \in \mathbb{R}^n\},
\]
is a random set ([21, lemma 2.1]). The invariant random set $\mathcal{M}(\omega)$ is called a Lipschitz random invariant manifold.

3. Framework

In the section, we present the framework for our reduction method for stochastic filtering and present some results which will be applied in the following sections.

Let $\Omega^1 = C_0(\mathbb{R}, \mathbb{R}^n)$ be the set of continuous functions on $\mathbb{R}$ with values in $\mathbb{R}^n$ that are zero at the origin. This set is equipped with the compact-open topology. Let $\mathcal{P}^1$ be its Borel $\sigma$-algebra and $\mathbb{P}^1$ the Wiener measure on $\Omega^1$. Define
\[
\theta_t \omega_1(\cdot) := \omega_1(\cdot + t) - \omega_1(t), \quad \omega_1 \in \Omega^1, \quad t \in \mathbb{R}.
\]

Then $(\Omega^1, \mathcal{P}^1, \mathbb{P}^1, \theta_t^1)$ is a metric dynamical system. Similarly, we define $\Omega^2 = C_0(\mathbb{R}, \mathbb{R}^m)$ and $\mathcal{P}^2, \mathbb{P}^2, \theta_t^2$. Thus, $(\Omega^2, \mathcal{P}^2, \mathbb{P}^2, \theta_t^2)$ is another metric dynamical system. Introduce
\[
\Omega := \Omega^1 \times \Omega^2, \quad \mathcal{P} := \mathcal{P}^1 \times \mathcal{P}^2, \quad \mathbb{P} := \mathbb{P}^1 \times \mathbb{P}^2, \quad \theta_t := \theta_t^1 \times \theta_t^2,
\]
and then $(\Omega, \mathcal{P}, \mathbb{P}, \theta_t)$ is a metric dynamical system that is used in the sequel.

Consider the following stochastic slow-fast system
\[
\begin{aligned}
\dot{x}^\varepsilon &= A x^\varepsilon + f(x^\varepsilon, y^\varepsilon) + \sigma_1 \dot{V}, \\
\dot{y}^\varepsilon &= \frac{1}{\varepsilon} B y^\varepsilon + \frac{1}{\varepsilon} g(x^\varepsilon, y^\varepsilon) + \frac{\sigma_2}{\varepsilon} \dot{W},
\end{aligned}
\tag{3}
\]
where $A$ and $B$ are $n \times n$ and $m \times m$ matrices respectively, and the interaction functions $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ are Borel measurable. Moreover, $V, W$ are mutually independent standard Brownian motions taking values in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, $\sigma_1$ and $\sigma_2$ are nonzero real noise intensities, and $\varepsilon$ is a small positive parameter representing the ratio of the two time scales. We make the following hypotheses:

\begin{itemize}
\item [(H_1)] There exists a $\gamma_1 \geq 0$ such that
\[
\|A\| \leq \gamma_1,
\]
where $\|A\|$ stands for the norm of the matrix $A$ such that $|Ax| \leq \|A\| |x|$ for every $x \in \mathbb{R}^n$, and $A$ has no eigenvalue on the imaginary axis.
\end{itemize}
(H3) There exists a \( \gamma_2 > 0 \) such that
\[
(By, y) \leq -\gamma_2 |y|^2, \quad y \in \mathbb{R}^m.
\]

(H4) There exists a positive constant \( L \) such that for all \( (x_1, y_1) \in \mathbb{R}^n \times \mathbb{R}^m \)
\[
|f(x_1, y_1) - f(x_2, y_2)| \leq L(|x_1 - x_2| + |y_1 - y_2|),
\]
and
\[
|g(x_1, y_1) - g(x_2, y_2)| \leq L(|x_1 - x_2| + |y_1 - y_2|).
\]

(H5) There exist two positive constants \( C_f, C_g \) such that
\[
\sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} |f(x, y)| = C_f,
\]
\[
\sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} |g(x, y)| = C_g.
\]

Under the assumptions (H3) and (H5), the system (3) has a global unique solution \((x^\varepsilon(t), y^\varepsilon(t))\), with a given initial value \((x_0, y_0)\). Define the solution operator \( \varphi^\varepsilon(x_0, y_0) := (x^\varepsilon(t), y^\varepsilon(t)) \), and then \( \varphi^\varepsilon \) is a random dynamical system. Introduce two auxiliary systems
\[
d\eta = A\eta dt + \sigma_1 dV,
\]
\[
d\xi^\varepsilon = \frac{1}{\varepsilon} B\xi^\varepsilon dt + \frac{\sigma_2}{\sqrt{\varepsilon}} dW.
\]
So, by [21, lemma 3.1], there exist two random variables \( \eta, \xi^\varepsilon \) such that \( \eta(\theta^1_1 \omega_1), \xi^\varepsilon(\theta^2_2 \omega_2) \) solve two above equations, respectively. Set
\[
\bar{x}^\varepsilon := x^\varepsilon - \eta(\theta^1_1 \omega_1),
\]
\[
\bar{y}^\varepsilon := y^\varepsilon - \xi^\varepsilon(\theta^2_2 \omega_2),
\]
and then \((\bar{x}^\varepsilon, \bar{y}^\varepsilon)\) satisfy the following system
\[
\begin{cases}
\dot{x}^\varepsilon = A\bar{x}^\varepsilon + f(x^\varepsilon + \eta(\theta^1_1 \omega_1), y^\varepsilon + \xi^\varepsilon(\theta^2_2 \omega_2)), \\
\dot{y}^\varepsilon = \frac{1}{\varepsilon} B\bar{y}^\varepsilon + \frac{1}{\varepsilon} g(\bar{x}^\varepsilon + \eta(\theta^1_1 \omega_1), \bar{y}^\varepsilon + \xi^\varepsilon(\theta^2_2 \omega_2)).
\end{cases}
\]
Moreover, \((\bar{x}^\varepsilon, \bar{y}^\varepsilon)\) generates a random dynamical system denoted by \( \bar{\varphi}^\varepsilon \). The following theorem comes from [21, theorem 4.2].

**Theorem 3.1 (Random slow manifold).** Suppose that \( \varepsilon > 0 \) is sufficiently small and (H1)–(H5) are satisfied. Then \( \varphi^\varepsilon \) has a random invariant manifold
\[
\bar{\mathcal{M}}^\varepsilon(\omega) = \{(x, H^\varepsilon(\omega, x)), x \in \mathbb{R}^n\},
\]
where for \( \omega \in \Omega \),
\[
\sup_{x_1 \neq x_2 \in \mathbb{R}^n} \frac{|H^\varepsilon(\omega, x_1) - H^\varepsilon(\omega, x_2)|}{|x_1 - x_2|} \leq \frac{2(\gamma_2 - \alpha)}{\gamma_2 - \alpha - L}.
\]
and \( \alpha \) is a positive number satisfying \( \gamma_2 - \alpha > L \).
Based on the relation between $\varphi^\varepsilon$ and $\bar{\varphi}^\varepsilon$, it holds that $\varphi^\varepsilon$ also has a random invariant manifold

$$\mathcal{M}^\varepsilon(\omega) = \{(x + \eta(\omega_1), H^\varepsilon(\omega, x) + \xi^\varepsilon(\omega_2)), x \in \mathbb{R}^n\}.$$ 

By the same deduction as [6, theorem 4.4], we could get a reduction system on $\mathcal{M}^\varepsilon$.

**Theorem 3.2 (Reduced system on the random slow manifold).** Assume that $\varepsilon > 0$ is sufficiently small and $\text{(H1)} - \text{(H5)}$ hold. Then for the system (3), there exists the following reduced low dimensional system on the random slow manifold:

$$\begin{cases}
\dot{x}^\varepsilon = A \bar{x}^\varepsilon + f(\bar{x}^\varepsilon, \bar{y}^\varepsilon) + \sigma \bar{V}, \\
\dot{\bar{y}}^\varepsilon = H^\varepsilon(\theta(\omega), \bar{x}^\varepsilon - \eta(\omega_1)) + \xi^\varepsilon(\omega_2),
\end{cases}$$

(4)

such that for $t > 0$ and almost all $\omega$,

$$|z^\varepsilon(t, \omega) - \bar{z}^\varepsilon(t, \omega)| \leq C_{L, \gamma, \alpha} e^{-\alpha t} |z^\varepsilon(0) - \bar{z}^\varepsilon(0)|,$$

where $\bar{z}^\varepsilon(t) = (\bar{x}^\varepsilon(t), \bar{y}^\varepsilon(t))$ is the solution of the low dimensional system (4) with the initial value $\bar{z}^\varepsilon(0) = (\bar{x}_0, \bar{y}_0)$ and $C_{L, \gamma, \alpha} > 0$ is a constant depending on $L, \gamma, \alpha$.

### 4. An approximate filter on the slow manifold

In the section we introduce nonlinear filtering problems for the system (3) and the reduced system (4) on the random slow manifold, and then study their relation.

#### 4.1. Nonlinear filtering problems

In the subsection we introduce nonlinear filtering problems for the system (3) and the reduced system (4).

For $T > 0$, an observation system is given by

$$r^\varepsilon_t = U_t + \int_0^t h(x^\varepsilon_s, y^\varepsilon_s) \, ds, \quad t \in [0, T],$$

where $U$ is a standard Brownian motion. Note that here $U$ may be either independent $V$ and $W$, or dependent $V$ and $W$ (see [3]). For the observation system $r^\varepsilon$, we make the following additional hypothesis:

$$(\text{H}_6) \quad h \text{ is bounded and Lipschitz continuous in } (x, y) \text{ whose Lipschitz constant is denoted by } \|h\|_{L^p}. $$

Under the assumption $\text{(H}_6)\), $r^\varepsilon$ is well defined. Denote

$$(A^\varepsilon)^{-1} := \exp \left\{ -\int_0^t h(x^\varepsilon_s, y^\varepsilon_s) \, dU_s - \frac{1}{2} \int_0^t |h(x^\varepsilon_s, y^\varepsilon_s)|^2 \, ds \right\},$$

and then $(A^\varepsilon)^{-1}$ is an exponential martingale under $\mathbb{P}$. By use of $(A^\varepsilon)^{-1}$, we can define a probability measure $\mathbb{P}^\varepsilon$ via

$$\frac{d\mathbb{P}^\varepsilon}{d\mathbb{P}} = (A^\varepsilon)^{-1}.$$
By the Girsanov theorem for Brownian motions, we can obtain that under the probability measure $\mathbb{P}^\varepsilon$, $\tilde{r}$ is a standard Brownian motion.

Rewrite $\Lambda_\varepsilon^t$ as

$$
\Lambda_\varepsilon^t = \exp \left\{ \int_0^t h(x_\varepsilon^t, \tilde{y}_\varepsilon^t)dr_\varepsilon^t - \frac{1}{2} \int_0^t |h(x_\varepsilon^t, \tilde{y}_\varepsilon^t)|^2ds \right\},
$$

and define

$$
\rho_\varepsilon^t(\phi) := \mathbb{E}^\varepsilon[\phi(x_\varepsilon^t)\Lambda_\varepsilon^t|\mathcal{R}_\varepsilon^t], \quad \phi \in \mathcal{B}(\mathbb{R}^n),
$$

where $\mathbb{E}^\varepsilon$ stands for the expectation under $\mathbb{P}^\varepsilon$. $\mathcal{R}_\varepsilon^t := \sigma(\tilde{r}_\varepsilon^s : 0 \leq s \leq t) \vee \mathcal{N}$ and $\mathcal{N}$ is the collection of all $\mathbb{P}$-measure zero sets. Here $\rho_\varepsilon^t$ is called nonnormalized filtering of $x_\varepsilon^t$ with respect to $\mathcal{R}_\varepsilon^t$. Introduce the measure-valued process

$$
\pi_\varepsilon^t(\phi) := \mathbb{E}[\phi(x_\varepsilon^t)|\mathcal{R}_\varepsilon^t], \quad \phi \in \mathcal{B}(\mathbb{R}^n),
$$

and then by the Kallianpur–Striebel formula it holds that

$$
\pi_\varepsilon^t(\phi) = \frac{\rho_\varepsilon^t(\phi)}{\rho_\varepsilon^t(1)}.
$$

Moreover, $\pi_\varepsilon^t$ is called normalized filtering of $x_\varepsilon^t$ with respect to $\mathcal{R}_\varepsilon^t$, or the nonlinear filtering problem for $x_\varepsilon^t$ with respect to $\mathcal{R}_\varepsilon^t$.

Besides, we rewrite the reduced system (4) as

$$
\dot{x}_\varepsilon^t = A\tilde{x}_\varepsilon + \tilde{f}(\omega, \tilde{x}_\varepsilon^t) + \sigma_1 \dot{\nu},
$$

where $\tilde{f}(\omega, x) := f(x, H(\theta, \omega, x - \eta(\theta^1\omega_1)) + \xi(\theta^2\omega_2))$, and study the nonlinear filtering problem for $\tilde{x}_\varepsilon$. Set

$$
\tilde{h}(\omega, x) := h(x, H(\theta, \omega, x - \eta(\theta^1\omega_1)) + \xi(\theta^2\omega_2)),
$$

$$
\tilde{\Lambda}_\varepsilon^t := \exp \left\{ \int_0^t \tilde{h}(\omega, \tilde{x}_\varepsilon^t)dr_\varepsilon^t - \frac{1}{2} \int_0^t |\tilde{h}(\omega, \tilde{x}_\varepsilon^t)|^2ds \right\},
$$

and then $\tilde{\Lambda}_\varepsilon^t$ is an exponential martingale under $\mathbb{P}^\varepsilon$. Thus, we define the nonnormalized filtering for $\tilde{x}_\varepsilon$ by

$$
\tilde{\rho}_\varepsilon^t(\phi) := \mathbb{E}^\varepsilon[\phi(\tilde{x}_\varepsilon^t)\tilde{\Lambda}_\varepsilon^t|\mathcal{R}_\varepsilon^t].
$$

And set

$$
\tilde{\pi}_\varepsilon^t(\phi) := \frac{\tilde{\rho}_\varepsilon^t(\phi)}{\tilde{\rho}_\varepsilon^t(1)},
$$

and then we will prove that $\tilde{\pi}_\varepsilon^t$ could be understood as the nonlinear filtering problem for $\tilde{x}_\varepsilon$ with respect to $\mathcal{R}_\varepsilon^t$.

4.2. The relation between $\pi_\varepsilon^t$ and $\tilde{\pi}_\varepsilon^t$

In the subsection we will show that a suitable distance between $\pi_\varepsilon^t$ and $\tilde{\pi}_\varepsilon^t$ converges to zero as $\varepsilon \to 0$. Let us start with two key lemmas.

**Lemma 4.1.** Under $(\text{H}_6)$, there exists a constant $C > 0$ such that

$$
\mathbb{E}|\tilde{\rho}_\varepsilon^t(1)|^{-p} < \exp \left\{ (2p^2 + p + 1)CT/2 \right\}, \quad t \in [0, T], \quad p > 1.
$$
Proof. Let us compute $E|\hat{p}_t^\varepsilon(1)|^{-p}$. By the Hölder inequality, it holds that

$$E|\hat{p}_t^\varepsilon(1)|^{-p} = E^\varepsilon|\hat{p}_t^\varepsilon(1)|^{-p} \Lambda_t^\varepsilon \leq (E^\varepsilon|\hat{p}_t^\varepsilon(1)|^{-2p})^{1/2}(E^\varepsilon(\Lambda_t^\varepsilon)^2)^{1/2}.$$  

For $E^\varepsilon|\hat{p}_t^\varepsilon(1)|^{-2p}$, notice that $\hat{p}_t^\varepsilon(1) = E^\varepsilon[\hat{\Lambda}_t^\varepsilon|\mathcal{R}_t^\varepsilon]$. And then it follows from the Jensen inequality that

$$E^\varepsilon|\hat{p}_t^\varepsilon(1)|^{-2p} = E^\varepsilon\left[E^\varepsilon[\hat{\Lambda}_t^\varepsilon|\mathcal{R}_t^\varepsilon]^{-2p}\right] \leq E^\varepsilon\left[E^\varepsilon[\hat{\Lambda}_t^\varepsilon|\mathcal{R}_t^\varepsilon]^{1-2p}\mathcal{R}_t^\varepsilon\right] = E^\varepsilon[\hat{\Lambda}_t^\varepsilon|\mathcal{R}_t^\varepsilon]^{1-2p}. $$

Thus, the definition of $\hat{\Lambda}_t^\varepsilon$ allows us to obtain that

$$E^\varepsilon[\hat{\Lambda}_t^\varepsilon|\mathcal{R}_t^\varepsilon] = E^\varepsilon\left[\exp\left\{-2p \int_0^t \tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)dr_t^\varepsilon + \frac{2p}{2} \int_0^t |\tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)|^2 ds\right\}\right]$$

$$= E^\varepsilon\left[\exp\left\{-2p \int_0^t \tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)dr_t^\varepsilon - \frac{4p^2}{2} \int_0^t |\tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)|^2 ds\right\}\right]$$

$$\leq \exp\left\{(2p^2 + p)CT\right\} E^\varepsilon\left[\exp\left\{-2p \int_0^t \tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)dr_t^\varepsilon - \frac{4p^2}{2} \int_0^t |\tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)|^2 ds\right\}\right]$$

$$= \exp\left\{(2p^2 + p)CT\right\},$$

where the last step is based on the fact that $\exp\left\{-2p \int_0^t \tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)dr_t^\varepsilon - \frac{4p^2}{2} \int_0^t |\tilde{h}_t^\varepsilon(\omega, x_t^\varepsilon)|^2 ds\right\}$ is an exponential martingale under $P^\varepsilon$.

Similarly, we know that $E^\varepsilon(\hat{\Lambda}_t^\varepsilon)^2 \leq \exp\{CT\}$. So, by simple calculation, it holds that

$$E|\hat{p}_t^\varepsilon(1)|^{-p} \leq \exp\left\{(2p^2 + p + 1)CT/2\right\}. $$

The proof is complete.

Lemma 4.2. Assume that $(H_1)$–$(H_6)$ are satisfied. Then for $\phi \in C^1_b(\mathbb{R}^n)$,

$$E|\rho_t^\varepsilon(\phi) - \hat{\rho}_t^\varepsilon(\phi)|^p \leq C||\phi||_p \varepsilon^p (\varepsilon(0) - \varepsilon(t))^p \mathbb{E}[\varepsilon(t)^{1/4}(\varepsilon(t)^{1/4} + \varepsilon)^{3/4})^{1/4}], \quad t \in [0, T], \quad p > 1,$$

where the constant $C > 0$ is independent of $\varepsilon$.

Proof. For $\phi \in C^1_b(\mathbb{R}^n)$, it follows from the Hölder inequality that

$$E|\rho_t^\varepsilon(\phi) - \hat{\rho}_t^\varepsilon(\phi)|^p = E|\rho_t^\varepsilon(\phi) - \hat{\rho}_t^\varepsilon(\phi)|^p \Lambda_t^\varepsilon \leq (E^\varepsilon|\rho_t^\varepsilon(\phi) - \hat{\rho}_t^\varepsilon(\phi)|^{2p})^{1/2}(E^\varepsilon(\Lambda_t^\varepsilon)^2)^{1/2}$$

$$\leq \exp\left\{CT/2\right\} (E^\varepsilon|\rho_t^\varepsilon(\phi) - \hat{\rho}_t^\varepsilon(\phi)|^{2p})^{1/2}.$$
\[ \mathbb{E}^\varepsilon [\rho^\varepsilon (\phi) - \tilde{\rho}^\varepsilon (\phi)]^{2p} = \mathbb{E}^\varepsilon \left[ \mathbb{E}^\varepsilon [\phi(x^\varepsilon_t)\Lambda_t^\varepsilon | \mathcal{R}_t^\varepsilon ] - \mathbb{E}^\varepsilon [\phi(\tilde{x}^\varepsilon_t)\tilde{\Lambda}_t^\varepsilon | \mathcal{R}_t^\varepsilon ] \right]^{2p} \\
= \mathbb{E}^\varepsilon \left[ \mathbb{E}^\varepsilon [\phi(x^\varepsilon_t)\Lambda_t^\varepsilon - \phi(\tilde{x}^\varepsilon_t)\tilde{\Lambda}_t^\varepsilon | \mathcal{R}_t^\varepsilon ] \right]^{2p} \\
\leq \mathbb{E}^\varepsilon \left[ \mathbb{E}^\varepsilon \left[ \left| \phi(x^\varepsilon_t)\Lambda_t^\varepsilon - \phi(\tilde{x}^\varepsilon_t)\tilde{\Lambda}_t^\varepsilon \right| \right]^{2p} | \mathcal{R}_t^\varepsilon \right] \\
= \mathbb{E}^\varepsilon \left[ \phi(x^\varepsilon_t)\Lambda_t^\varepsilon - \phi(\tilde{x}^\varepsilon_t)\tilde{\Lambda}_t^\varepsilon \right]^{2p} \\
\leq 2^{2p-1} \mathbb{E}^\varepsilon \left[ \phi(x^\varepsilon_t)\Lambda_t^\varepsilon - \phi(\tilde{x}^\varepsilon_t)\tilde{\Lambda}_t^\varepsilon \right]^{2p} \\
+ 2^{2p-1} \mathbb{E}^\varepsilon \left[ \phi(x^\varepsilon_t)\Lambda_t^\varepsilon - \phi(\tilde{x}^\varepsilon_t)\tilde{\Lambda}_t^\varepsilon \right]^{2p} \\
=: I_1 + I_2. \quad (5) \]

First, we deal with \( I_1 \). By the Hölder inequality, it holds that

\[ I_1 \leq 2^{2p-1}(\mathbb{E}^\varepsilon \left[ \phi(x^\varepsilon_t) - \phi(\tilde{x}^\varepsilon_t) \right]^{2p})^{1/2}(\mathbb{E}^\varepsilon \left[ \Lambda_t^\varepsilon \right]^{2p})^{1/2} \]

\[ \leq 2^{2p-1}\|\phi\|^{2p}(\mathbb{E}^\varepsilon \left[ x^\varepsilon_t - \tilde{x}^\varepsilon_t \right]^{4p})^{1/2}\left( \mathbb{E}^\varepsilon \exp \left\{ 4p \int_0^t h(x^\varepsilon_s, y^\varepsilon_s) dr_s^\varepsilon - \frac{4p^2}{2} \int_0^t |h(x^\varepsilon_s, y^\varepsilon_s)|^2 ds \right\} \right)^{1/2} \]

\[ \cdot \exp \left\{ \frac{4p^2}{2} \int_0^t |h(x^\varepsilon_s, y^\varepsilon_s)|^2 ds - 2p \int_0^t |h(x^\varepsilon_s, y^\varepsilon_s)|^2 ds \right\} \]

\[ \leq 2^{2p-1}\|\phi\|^{2p} C_{L^2, \alpha, \beta}^{2p} e^{-\frac{1}{2}p} (\mathbb{E}^\varepsilon \left[ x^\varepsilon_t - \tilde{x}^\varepsilon_t \right]^{4p}) \varepsilon^{-p} e^{-2p} e^{-p}. \quad (6) \]

where the last step is based on theorem \( 3.2 \) and the fact that the process

\[ \exp \left\{ 4p \int_0^t h(x^\varepsilon_s, y^\varepsilon_s) dr_s^\varepsilon - \frac{4p^2}{2} \int_0^t |h(x^\varepsilon_s, y^\varepsilon_s)|^2 ds \right\} \]

is an exponential martingale under \( \mathbb{P}^\varepsilon \).

Next, for \( I_2 \), we know that

\[ I_2 \leq 2^{2p-1}\|\phi\|^{2p} \mathbb{E}^\varepsilon \left[ \Lambda_t^\varepsilon - \tilde{\Lambda}_t^\varepsilon \right]^{2p}. \]

Note that by the Itô formula, \( \Lambda_t^\varepsilon \) and \( \tilde{\Lambda}_t^\varepsilon \) satisfy the following equations, respectively,

\[ \Lambda_t^\varepsilon = 1 + \int_0^t \Lambda_s^\varepsilon h(x_s^\varepsilon, y_s^\varepsilon) dr_s^\varepsilon, \quad \tilde{\Lambda}_t^\varepsilon = 1 + \int_0^t \tilde{\Lambda}_s^\varepsilon h(\omega, x_s^\varepsilon) dr_s^\varepsilon. \]

Thus, by BDG inequality and the Hölder inequality it holds that
This proves the lemma. □
Now, we are ready to state and prove the main result in the paper. First, we give out two concepts used in the proof of theorem 4.5.

**Definition 4.3.** The set \( M \subset C_b^1(\mathbb{R}^n) \) strongly separates points in \( \mathbb{R}^n \) when the convergence \( \lim_{n \to \infty} \phi(x_n) = \phi(x), \forall \phi \in M \), for some \( x, x_n \in \mathbb{R}^n \), implies that \( \lim_{n \to \infty} x_n = x \).

**Definition 4.4.** The set \( N \subset C_b^1(\mathbb{R}^n) \) is convergence determining for the topology of weak convergence of probability measures, if \( \mu_n \) and \( \mu \) are probability measures on \( \mathcal{B}(\mathbb{R}^n) \), such that \( \lim_{n \to \infty} \int_{\mathbb{R}^n} \phi \ d\mu_n = \int_{\mathbb{R}^n} \phi \ d\mu \) for any \( \phi \in N \), then \( \mu_n \) converges weakly to \( \mu \).

**Theorem 4.5 (Approximation by the reduced filter on slow manifold).** Assume the hypotheses (H1)–(H6) hold. Then for \( p > 1, \varepsilon \) sufficiently small, and \( t \in [0, T] \), there exists a positive constant \( C \) such that for \( \phi \in C_b^1(\mathbb{R}^n) \)

\[
\mathbb{E}[|\pi_t^\varepsilon(\phi) - \tilde{\pi}_t^\varepsilon(\phi)|^p] \leq C\|\phi\|^p(\mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{16p})^{1/16}(e^{-\frac{16T}{\alpha\varepsilon}} + \varepsilon)^{1/4}.
\]

Thus, for the distance \( d(\cdot, \cdot) \) in the space of probability measures that induces the weak convergence, the following approximation holds:

\[
\mathbb{E}[d(\pi_t^\varepsilon, \tilde{\pi}_t^\varepsilon)] \leq C(\mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{16p})^{1/16}(e^{-\frac{16T}{\alpha\varepsilon}} + \varepsilon)^{1/4}.
\]

This means the filter for the low dimensional system on the random slow manifold approximates the original filter in this distance \( d(\cdot, \cdot) \).

**Proof.** For \( \phi \in C_b^1(\mathbb{R}^n) \), it follows from lemmas 4.1 and 4.2 that

\[
\mathbb{E}[|\pi_t^\varepsilon(\phi) - \tilde{\pi}_t^\varepsilon(\phi)|^p] = \mathbb{E}\left[\left|\frac{\rho_t^\varepsilon(\phi) - \tilde{\rho}_t^\varepsilon(\phi)}{\tilde{\rho}_t^\varepsilon(1)} - \pi_t^\varepsilon(\phi)\frac{\rho_t^\varepsilon(1) - \tilde{\rho}_t^\varepsilon(1)}{\tilde{\rho}_t^\varepsilon(1)}\right|^p\right]
\]

\[
\leq 2^{p-1}\mathbb{E}\left[\left|\frac{\rho_t^\varepsilon(\phi) - \tilde{\rho}_t^\varepsilon(\phi)}{\tilde{\rho}_t^\varepsilon(1)}\right|^p + 2^{p-1}\mathbb{E}\left[\pi_t^\varepsilon(\phi)\frac{\rho_t^\varepsilon(1) - \tilde{\rho}_t^\varepsilon(1)}{\tilde{\rho}_t^\varepsilon(1)}\right]^p\right]
\]

\[
\leq 2^{p-1}\left(\mathbb{E}|\rho_t^\varepsilon(\phi) - \tilde{\rho}_t^\varepsilon(\phi)|^{2p}\right)^{1/2}\left(\mathbb{E}|\tilde{\rho}_t^\varepsilon(1)|^{-2p}\right)^{1/2}
\]

\[
+ 2^{p-1}\|\phi\|^p\left(\mathbb{E}|\rho_t^\varepsilon(1) - \tilde{\rho}_t^\varepsilon(1)|^{2p}\right)^{1/2}\left(\mathbb{E}|\tilde{\rho}_t^\varepsilon(1)|^{-2p}\right)^{1/2}
\]

\[
\leq C\|\phi\|^p(\mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{8p})^{1/8}(e^{-\frac{16T}{\alpha\varepsilon}} + \varepsilon)^{1/4}.
\]

To complete the proof, we only consider \( \mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{8p} \). By the Hölder inequality, it holds that

\[
\mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{8p} = \mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{8p}(\Lambda_t^\varepsilon)^{-8p} \leq (\mathbb{E}|\varepsilon(0) - \tilde{\varepsilon}(0)|^{16p})^{1/2}(\mathbb{E}(\Lambda_t^\varepsilon)^{-16p})^{1/2}.
\]

By simple calculations, we obtain that
\[ \mathbb{E}(\Lambda_{\delta}^{\epsilon})^{-16\rho} = \mathbb{E}\left( \exp\left\{ -16\rho \int_0^T h(x_{\delta}^\epsilon, y_{\delta}^\epsilon) \, dU_s - \frac{16\rho}{2} \int_0^t |h(x_{\delta}^\epsilon, y_{\delta}^\epsilon)|^2 \, ds \right\} \right) \]

\[ = \mathbb{E}\left[ \exp\left\{ -16\rho \int_0^T h(x_{\delta}^\epsilon, y_{\delta}^\epsilon) \, dU_s - \frac{(16\rho)^2}{2} \int_0^t |h(x_{\delta}^\epsilon, y_{\delta}^\epsilon)|^2 \, ds \right\} \right] \]

\[ \leq \exp\left\{ C \frac{(16\rho)^2 - 16\rho}{2} \right\}, \]

where the last step is based on the fact that \( \exp\left\{ -16\rho \int_0^T h(x_{\delta}^\epsilon, y_{\delta}^\epsilon) \, dU_s - \frac{(16\rho)^2}{2} \int_0^t |h(x_{\delta}^\epsilon, y_{\delta}^\epsilon)|^2 \, ds \right\} \) is an exponential martingale under \( \mathbb{P} \). Thus,

\[ \mathbb{E}[z^\epsilon(0) - \tilde{z}^\epsilon(0)]^{16\rho} \leq C(\mathbb{E}[z^\epsilon(0) - \tilde{z}^\epsilon(0)]^{16\rho})^{1/2}. \]

Next, we know that there exists a countable algebra \( \{\phi_i, i = 1, 2, \cdots\} \) of \( C_b^1(\mathbb{R}^n) \) that strongly separates points in \( \mathbb{R}^n \). Thus, it follows from theorem 3.4.5 in [5] that \( \{\phi_i, i = 1, 2, \cdots\} \) is convergence determining for the topology of weak convergence of probability measures. For two probability measures \( \mu, \tau \) on \( B(\mathbb{R}^n) \), define

\[ d(\mu, \tau) := \sum_{i=1}^{\infty} \left| \int_{\mathbb{R}^n} \phi_i \, d\mu - \int_{\mathbb{R}^n} \phi_i \, d\tau \right| \cdot 2^i. \]

Then \( d \) is a distance in the space of probability measures on \( B(\mathbb{R}^n) \). Since \( \{\phi_i, i = 1, 2, \cdots\} \) is convergence determining for the topology of weak convergence of probability measures, \( d \) induces the weak convergence. The proof is complete. \( \square \)

5. Numerical experiments

In this section, we present an example to illustrate our filtering method on a random slow manifold.

Consider the following slow-fast stochastic system:

\[ \begin{cases} \dot{x}^\epsilon = x^\epsilon + \frac{1}{4} \sin(y^\epsilon) + 0.01V, \\ \dot{y}^\epsilon = -\frac{1}{2}y^\epsilon + \frac{1}{4\epsilon} \cos(x^\epsilon) + \frac{1}{\sqrt{\epsilon}}W, \end{cases} \]  \( \tag{8} \)

where \( A = 1, B = -1, f(x, y) = \frac{1}{2} \sin y \) and \( g(x, y) = \frac{1}{2} \cos x \). It is easy to justify that \( A, B, f, g \) satisfy \( (H_1)-(H_5) \) with \( \gamma_1 = \gamma_2 = 1, L = C_f = C_g = \frac{1}{2} \). Then the system \( (8) \) has a unique solution \( (x^\epsilon, y^\epsilon) \), which generates a random dynamical system \( \varphi^\epsilon \).

Introduce the following two auxiliary systems:

\[ \begin{align*} d\eta &= \eta \, dt + 0.01 \, dV, \\
\, \text{d}^\epsilon &=-\frac{1}{\epsilon^2} \xi^\epsilon \, dt + \frac{1}{\sqrt{\epsilon}} \, dW. \end{align*} \]

Then two equations have the following stationary solutions, respectively,
\[
\begin{align*}
\eta(\omega_1) & = -0.01 \int_0^\infty e^{-s} dV_s(\omega_1), \\
\xi^\omega(\omega_2) & = \frac{1}{\sqrt{\epsilon}} \int_{-\infty}^0 e^{s} dW_s(\omega_2).
\end{align*}
\]

Define
\[
\begin{align*}
\bar{x}_r^\epsilon & := x_r^\epsilon - \eta(\theta_1^1 \omega_1), \\
\bar{y}_r^\epsilon & := y_r^\epsilon - \xi^\omega(\theta_2^2 \omega_2),
\end{align*}
\]
and then \((\bar{x}_r^\epsilon, \bar{y}_r^\epsilon)\) solve the following equation
\[
\begin{align*}
\hat{\bar{x}}^\epsilon & = \bar{x}^\epsilon + \frac{1}{2} \sin \left( \bar{y}^\epsilon + \xi^\omega(\theta_2^2 \omega_2) \right), \quad \bar{x}_0^\epsilon = x \in \mathbb{R}, \\
\hat{\bar{y}}^\epsilon & = -\frac{1}{2} \bar{y}^\epsilon + \frac{1}{\epsilon} \cos \left( \bar{x}^\epsilon + \eta(\theta_1^1 \omega_1) \right), \quad \bar{y}_0^\epsilon = y \in \mathbb{R}.
\end{align*}
\]

Thus, by theorem 3.1, we can get the following random invariant manifold for \((\bar{x}_r^\epsilon, \bar{y}_r^\epsilon)\)
\[
\mathcal{M}^\epsilon(\omega) = \{(x, H^\epsilon(\omega, x)), x \in \mathbb{R}\},
\]
where
\[
H^\epsilon(\omega, x) = \frac{1}{4\epsilon} \int_{-\infty}^0 e^s \cos \left( x_r^\epsilon(\omega, x) + \eta(\theta_1^1 \omega_1) \right) ds.
\]

By (9), it holds that \(\varphi^\epsilon\) has a random invariant manifold
\[
\mathcal{M}^\epsilon(\omega) = \{(x + \eta(\omega_1), H^\epsilon(\omega, x) + \xi^\omega(\omega_2)), x \in \mathbb{R}\}.
\]

Thus, one can obtain the following reduced one dimensional system on \(\mathcal{M}^\epsilon(\omega)\)
\[
\begin{align*}
\hat{x}^\epsilon & = \bar{x}^\epsilon + \frac{1}{2} \sin(\bar{y}^\epsilon) + \sigma_1 \bar{V}, \\
\hat{y}^\epsilon & = H^\epsilon(\theta_1^1 \omega, \bar{x}^\epsilon + \eta(\theta_1^1 \omega_1)) + \xi^\omega(\theta_2^2 \omega_2).
\end{align*}
\]

Next, the observation system is given by
\[
dr_r^\epsilon = \arctan(x_r^\epsilon) dt + dU_r,
\]
where \(h(x, y) = \arctan(x)\). And \(h(x, y)\) satisfies \((H_6)\).

To facilitate numerical simulation, we make some preparations. First, note that \(H^\epsilon(\omega, x)\) has an approximation \(H^0(\omega, x) + H^1(\omega, x)\) (with error \(O(\epsilon^2)\))
\[
\begin{align*}
H^0(\omega, x) & = \int_{-\infty}^0 e^s \left( x + \eta(\theta_1^1 \omega_1), Y_0(s) + \xi^\omega(\theta_2^2 \omega_2) \right) ds \\
& = \frac{1}{4} \int_{-\infty}^0 e^s \cos \left( x + \eta(\theta_1^1 \omega_1) \right) ds,
\end{align*}
\]
and
\[
\begin{align*}
H^1(\omega, x) & = \int_{-\infty}^0 e^s \left( g_1 \cdot \left[ x + \int_{-\infty}^0 f \left( x + \eta(\theta_1^1 \omega_1), Y_0(r) + \xi^\omega(\theta_2^2 \omega_2) \right) dr \right] \\
& + g_1 \left( x + \eta(\theta_1^1 \omega_1), Y_0(s) + \xi^\omega(\theta_2^2 \omega_2) \right) Y_1(s) \right) ds \\
& = -\frac{1}{4} \int_{-\infty}^0 e^s \left\{ \sin \left( x + \eta(\theta_1^1 \omega_1) \right) \cdot \left[ x + \frac{1}{4} \int_{-\infty}^0 \sin \left( Y_0(r) + \xi^\omega(\theta_2^2 \omega_2) \right) dr \right] \right\} ds.
\end{align*}
\]
Here \(Y_0, Y_1\) satisfy the following equations, respectively,
\[ Y_0'(s) = -Y_0(s) + \frac{1}{4} \cos(x + \eta(\theta_1, \omega_1)), \]
\[ Y_0(0) = H^0(\omega, x) \]

and
\[ \begin{cases} Y'_1(s) = -Y_1(s) - \frac{1}{4} \sin(x + \eta(\theta_1, \omega_1)) \cdot \left[ sx + \frac{1}{4} \int_0^s \sin(Y_0(r) + \xi'(\theta_2, \omega_2))dr \right], \\ Y_1(0) = H^1(\omega, x). \end{cases} \]

Second, note that \( W_{\omega_1}(\omega) \) is a Brownian motion. Hence \( \psi : \omega \to \omega \) is defined implicitly by \( W_{\psi(\omega)} = \frac{1}{\sqrt{\varepsilon}} W_{\omega} \). Thus, after a series of simple calculations, we have

\[ \begin{align*}
\text{Figure 1.} & \quad \text{(a) The original filter } \pi_f(x) \text{ (‘+’ curves) versus the reduced filter } \tilde{\pi}_f(x) \text{ (red curves): Initial value } x(0) = 1, y(0) = 1, \dot{x}(0) = 1, \varepsilon = 0.01; \text{ (b) the mean-square error } E[|\pi_f(\phi) - \tilde{\pi}_f(\phi)|]^2. \\
\text{(c) The original filter } \pi_{f_1}(x) \text{ (‘+’ curves) versus the reduced filter } \tilde{\pi}_{f_1}(x) \text{ (red curves): initial value } x(0) = 1, y(0) = 1, \dot{x}(0) = 1, \varepsilon = 0.1; \text{ (d) the mean-square error } E[|\pi_{f_1}(\phi) - \tilde{\pi}_{f_1}(\phi)|]^2. 
\end{align*} \]
\[
\eta(\theta_1 \tau \omega_1) = -0.01 \sqrt{\varepsilon} \int_0^\infty e^{-u} dV_u(\psi_1 \omega_1), \\
\xi^\varepsilon(\hat{\theta}_2 \tau \omega_2) = \int_{-\infty}^0 e^u dW_u(\psi_2 \omega_2).
\]

Set
\[
\eta_1(\psi_1 \omega_1) := -\int_0^\infty e^{-u} dV_u(\psi_1 \omega_1).
\]

Then \(\eta(\hat{\theta}_2 \tau \omega_1)\) is identically distributed as \(0.01 \eta_1(\psi_2 \omega_1)\).

Now we apply a particle filtering method (\cite{2, 13}) to simulate a nonlinear filter, which approximate the stochastic process \(\pi_t\) with discrete random measures of the form

**Figure 3.** (e) The original filter \(\pi_t\) (‘+’ curves) versus the reduced filter \(\hat{\pi}_t\) (red curves): initial value \(x^t(0) = 1, y^t(0) = 1, \hat{x^t}(0) = 0.95, \varepsilon = 0.01\); (f) the mean-square error \(E[\pi_t(\phi) - \hat{\pi}_t(\phi)]^2\).

**Figure 4.** (g) The original filter \(\pi_t\) (‘+’ curves) versus the reduced filter \(\hat{\pi}_t\) (red curves): initial value \(x^t(0) = 1, y^t(0) = 1, \hat{x^t}(0) = 0.95, \varepsilon = 0.1\); (h) the mean-square error \(E[\pi_t(\phi) - \hat{\pi}_t(\phi)]^2\).
\[ \sum_{i=1}^{N} a_i(t) \delta_{x_i^0} \]  

(12)

in other words, with empirical distributions associated with sets of randomly located particles of stochastic mass \( a_1(t), a_2(t), \ldots \), which have stochastic positions \( x_i^0, x_i^0, \ldots \). The particle filtering algorithm can be carried out in the following steps:

1. **Initialization**
   For \( j = 1, 2, \ldots, n \)
   
   - Sample \( x_i^0, y_i^0 \) from \( \pi_0 \).
   - \( a_j(0) = 1 \).
   
   end for

\[ \pi_0 := \frac{1}{n} \sum_{j=1}^{n} \delta_{(x_i^0, y_i^0)} \]
Step 2. Iteration
for l = 0 to m-1
for j = 0 to n
Using Euler method to generate the Gaussian random vector $x^e(t + \frac{\Delta t}{m})$ and $y^e(t + \frac{\Delta t}{m})$.
$$b_j(t + \frac{\Delta t}{m}) := \arctan(x^e(t))(r^e_j - \frac{\Delta t}{m} || \arctan(x^e(t))||^2)$$
$$a_j(t + \frac{\Delta t}{m}) := a_j(t) \exp(b_j(t + \frac{\Delta t}{m}))$$
end for
$t := t + \frac{\Delta t}{m}$
$$\Sigma(t) := \sum_{j=1}^{n} a_j(t)$$
$$\pi^e_t := \frac{1}{\Sigma(t)} \sum_{j=1}^{n} \delta(x^e(t),y^e(t))$$
end for

Step 3. Deterministic resampling
Use the Kitagawa’s deterministic resampling algorithm, as described in [8].
For the particle filtering algorithm, we take: $\phi(x) = \frac{10x_1}{1+\varepsilon}$, $n = 200$, $m = 400$, $\Delta t = 0.02$, $T = 8$.
We will compute the original filter
$$\pi^e_t(\phi) = E[\phi(x^e_t)|R^e_t]$$
the reduced filter
$$\tilde{\pi}^e_t(\phi) = \frac{\tilde{\rho}_t(\phi)}{\tilde{\rho}_t(1)}$$
and the mean-square error
$$E[|\pi^e_t(\phi) - \tilde{\pi}^e_t(\phi)|^2]$$
and plot in the following figures.
As seen in figures 1 and 2, it is clear that if the initial values of the original slow component and the reduced system are the same, the larger $\varepsilon$ is, the larger the fluctuation of the filtering error is. From figures 3–6 it is found that if the difference for the initial values of the original slow component and the reduced system becomes larger, the fluctuation of the filtering error is larger. These figures indicate that the filter error increases initially before decaying. This is due to some relaxation towards the slow manifold. As shown in theorem 4.5, the error depends on the initial values and parameter $\varepsilon$. Moreover, the filter error may likely not change if the filters have some initial spin-up period [10].

6. Conclusion
In this paper, we first obtain the low dimensional reduction of a slow-fast data assimilation system via an invariant slow manifold. Then we show that the low dimensional filter on the slow manifold approximates the original filter. Moreover, by an example we illustrate this approximate filter numerically.

Acknowledgments
The authors would like to thank Dr Xinyong Zhang (Tsinghua University, China) and Dr Jian Ren (Zhengzhou University of Light Industry, China) for helpful discussions and comments.
And they would also like to thank two anonymous referees for giving useful suggestions to improve this paper.

References

[1] Arnold L 1998 Random Dynamical Systems (Berlin: Springer)
[2] Bain A and Crisan D 2009 Fundamentals of Stochastic Filtering (Berlin: Springer)
[3] Cass T, Clarke M and Crisan D 2014 The filtering equations revisited Stochastic Analysis and Applications ed D Crisan et al (Springer Proc. in Mathematics and Statistics vol 100) (Berlin: Springer)
[4] Duan J 2015 An Introduction to Stochastic Dynamics (New York: Cambridge University Press)
[5] Ethier S N and Kurtz T G 1986 Markov Processes: Characterization and Convergence (New York: Wiley)
[6] Fu H, Liu X and Duan J 2013 Slow manifolds for multi-time-scale stochastic evolutionary systems Commun. Math. Sci. 11 141–62
[7] Gottwald G and Harlim J 2013 The role of additive and multiplicative noise in filtering complex dynamical systems Proc. R. Soc. A 469 20130096
[8] Higuchi T 1995 On the resampling scheme in the filtering procedure of the Kitagawa Monte Carlo Filter ISM Research Memo No. 556
[9] Imkeller P, Namachchivaya N S, Perkowski N and Yeong H C 2013 Dimensional reduction in nonlinear filtering: a homogenization approach Ann. Appl. Probab. 23 2290–326
[10] Kalnay E and Yang S-C 2010 Accelerating the spin-up of ensemble Kalman filtering Q. J. R. Meteorol. Soc. 136 1644–51
[11] Kim J, Josic K and Bennett M 2014 The validity of quasi-steady-state approximations in discrete stochastic simulations Biophys. J. 107 783–93
[12] Mitchell L and Gottwald G 2012 Data assimilation in slow-fast systems using homogenized climate models J. Atmos. Sci. 69 1359–77
[13] Park J H, Namachchivaya N S and Yeong H C 2011 Particle filters in a multiscale environment: homogenized hybrid particle filter J. Appl. Mech. 78 1–10
[14] Park J H, Sowers R B and Namachchivaya N S 2010 Dimensional reductionin nonlinear filtering Nonlinearity 23 305–24
[15] Park J H, Rozovskii B and Sowers R B 2011 Efficient nonlinear filtering of a singularly perturbed stochastic hybrid system LMS J. Comput. Math. 14 254–70
[16] Turcotte M, Garcia-Ojalvo J and Stiel G M 2008 A genetic timer through noise-induced stabilization of an unstable state Proc. Natl Acad. Sci. USA 41 15732–7
[17] Khasminskii R Z and Yin G 1996 On transition densities of singularly perturbed diffusions with fast and slow components SIAM J. Appl. Math. 56 1794–819
[18] Wu F, Tian T, Rawling J B and Yin G 2016 Approximate method for stochastic chemical kinetics with two-time scales by chemical Langevin equations J. Chem. Phys. 144 174112
[19] Wang W and Roberts A J 2013 Slow manifold and averaging for slow-fast stochastic differential system J. Math. Anal. Appl. 398 822–39
[20] Evensen G 2009 Data Assimilation: the Ensemble Kalman Filter (Berlin: Springer)
[21] Schmalfüß B and Schneider R 2008 Invariant manifolds for random dynamical systems with slow and fast variables J. Dyn. Differ. Equ. 20 133–64