2000

Convergence of Particle Filtering Method for Nonlinear Estimation of Vortex Dynamics

Sritharan, Sivaguru S.

Monterey, California. Naval Postgraduate School

CONVERGENCE OF PARTICLE FILTERING METHOD FOR NONLINEAR ESTIMATION OF VORTEX DYNAMICS

SIVAGURU S. SRITHARAN* AND MENG XU

Abstract. In this paper we formulate a numerical approximation method for the nonlinear filtering of vortex dynamics subject to noise using particle filter method. We prove the convergence of this scheme allowing the observation vector to be unbounded.

1. Introduction

Nonlinear estimation of turbulence and vortical structures has many applications in engineering and in geophysical sciences. In [24], [49], [50] and [51], mathematical foundation of nonlinear filtering methods was developed for viscous flow and for reacting and diffusing systems. The current work is in part an effort to contribute towards concrete computational methods to solve the nonlinear filtering equations derived in the above papers. We will however focus our attention on much simpler fluid dynamic models in terms of point vortices, which nevertheless contain significant physical attributes of fluid mechanics.

The particle filter method is a generalization of the traditional Monte-Carlo method and is often called the sequential Monte-Carlo method. The difference with Monte-Carlo method is the presence of an additional correction procedure applied at regular time intervals to the system of particles. At the correction time, each particle is replaced by a random number of particles. This amounts to particles branching into a random number of offsprings. The general principle is that particles with small weights have no offspring, and particles with large weights are replaced by several offsprings.

As a numerical method for nonlinear filtering problem, particle filter can be used to approximate general stochastic differential equations. In recent years, different variations of it have been studied, such as particle filter with occasional sampling [12], particle filter with variance reduction [13], branching particle filter [14], [15] and regularized particle filter [16], [34], most of which are applicable in discrete time setting and have been implemented computationally. In this paper, we will work in the continuous time setting and study the continuous time particle filter.

2000 *Mathematics Subject Classification.* Primary 60H15; Secondary 65C35.

Key words and phrases. nonlinear filtering, vortex method, stochastic vortex model, Zakai equation, particle filter .

* This research is supported by the Army Research Probability and Statistics Program through the grant DODARMY41712.

Our focus will be on the convergence of particle filter method applied to nonlinear filtering problem for stochastic diffusion. A. Bain and D. Crisan [4] proved convergence of this method for uniformly bounded observation process h. We extend their result by allowing h to have linear growth and h as a function of the signal to have an upper bound $f(t)$ as a L^2 function.

To understand how particle filter method can be formulated in our problem, we first introduce some background on nonlinear filtering.

We begin with a complete probability space $(\Omega, \mathcal{F}, P)$ on which our stochastic process will be defined. Consider the stochastic differential equation for the signal process X_t

$$d\mathrm{X}_t = f(\mathrm{X}_t)dt + \sigma(\mathrm{X}_t)d\mathrm{W}_t, \tag{1.1}$$

where $f : \mathrm{R}^n \to \mathrm{R}^n$ and $\sigma : \mathrm{R}^n \to \mathrm{R}^{n\times n}$, called the drift coefficient and diffusion coefficient respectively. $\mathrm{W} = (\mathrm{W}^j)_{j=1}^n$ is the n-dimensional Brownian motion. X_t which solves equation (1.1) is the n-dimensional signal process.

Denoting the filtration generated by $\{\mathrm{X}_s, s \leq t\}$ as $\mathcal{F}_t$, we can define the filtered probability space $(\Omega, \mathcal{F}_t, \mathcal{F}, P)$.

The observation process Y satisfies

$$d\mathrm{Y}_t = h(\mathrm{X}_t)dt + d\mathrm{B}_t, \tag{1.2}$$

where $\mathrm{Y} = (\mathrm{Y}_i)_{i=1}^m$ and $h = (h_i)_{i=1}^m : \mathrm{R}^n \to \mathrm{R}^m \quad \text{with} \quad m < n$. B is a standard m-dimensional Brownian motion independent of W.

The nonlinear filtering problem is to calculate the following conditional expectation

$$\pi_t(\varphi) = E[\varphi(\mathrm{X}_t)|\mathcal{Y}_t], \tag{1.3}$$

where $\mathcal{Y}_t$ is the σ-algebra generated by the back measurements $Y_s, 0 \leq s \leq t$. In fact one can prove that (1.3) is the least square estimate for $\varphi(\mathrm{X}_t)$ given $\mathcal{Y}_t$. $\pi_t(\varphi)$ satisfies a nonlinear stochastic differential equation, called the Fujisaki-Kallianpur-Kunita [FKK] equation [19]. The idea then is to use Girsanov theorem to analyze $\rho_t(\varphi)$, the unnormalized conditional density, which is related to $\pi_t(\varphi)$ by Kallianpur-Striebel formula and satisfies a linear stochastic differential equation.

Theorem 1.1 (Girsanov Theorem[38]). *Assume that $\psi(\cdot)$ is a R^m-valued $\mathcal{F}_t$-predictable process such that*

$$E\bigg(\int_0^T |\psi(s)|^2 ds\bigg) < \infty, \tag{1.4}$$

and

$$E\bigg(exp\bigg(\int_0^T \psi(s)^T dW(s) - \frac{1}{2}\int_0^T |\psi(s)|^2 ds\bigg)\bigg) = 1. \tag{1.5}$$

Then the process

$$\tilde{W}(t) = W(t) - \int_0^t \psi(s)ds, \quad t \in [0,T] \tag{1.6}$$

is a m-dimensional Brownian motion with respect to $\{\mathcal{F}_t\}_{t\geq 0}$ *on the probability space* $(\Omega, \mathcal{F}, \tilde{P})$ *where*

$$d\tilde{P}(\omega) = exp\bigg(\int_0^T \psi(s)^T dW(s) - \frac{1}{2}\int_0^T |\psi(s)|^2 ds\bigg)dP(\omega). \tag{1.7}$$

Here $|\cdot|$ *denotes the standard Euclidean norm.*

Let us also recall a result on moment estimate for X_t from I.I. Gihman and A.V. Skorohod [20].

Lemma 1.2. *Assume Lipschitz and growth conditions for the coefficients* f *and* σ. *Let the initial data satisfy* $\mathrm{E}|X_0|^{2d} < \infty$. *Then for any* $0 < t < T$, *the solution* X_t *of (1) will possess finite moments up to and including* $2n$*-order, i.e.*

$$E[|X_t|^{2d}] < \infty, \qquad for \quad any \quad d = 1, 2, \cdots, n. \tag{1.8}$$

Assume h is globally Lipschitz, non-negative and satisfies linear growth condition:

$$|h(x)|^2 \leq C(1 + |x|^2), \qquad \forall x \in \mathrm{R}^n, \quad C > 0. \tag{1.9}$$

Lemma 1.2 and growth rate (1.9) imply:

$$E\int_0^t |h(\mathrm{X}(s))|^2 ds \leq CE\int_0^t |\mathrm{X}(s)|^2 ds + Ct < \infty, \quad \text{for} \quad 0 < t < \infty. \tag{1.10}$$

Define

$$Z_t = \exp\bigg(\int_0^t h(\mathrm{X}_s)^T d\mathrm{W}_s - \frac{1}{2}\int_0^t |h(\mathrm{X}_s)|^2 ds\bigg). \tag{1.11}$$

Girsanov theorem holds if one can also show that $E[Z_t] = 1$ for all $t > 0$ and a well-known sufficient condition is the Novikov condition:

$$E[\exp(\frac{1}{2}\int_0^t |h(X_s)|^2 ds)] < \infty. \tag{1.12}$$

However, the Novikov condition is usually difficult to check unless function h is bounded.
For unbounded h of the type in (1.9), we use a truncation function approach by B. Ferrairo [18] to obtain $E[Z_t] = 1$.
In that paper the truncation function χ^N was introduced as follows:

$$\chi_t^N(v) = \begin{cases} 1 & \text{if } \int_0^t |h(v(s))|^2 ds \leq N \\ 0 & \text{otherwise.} \end{cases} \tag{1.13}$$

Novikov condition

$$E[\exp(\frac{1}{2}\int_0^t |\chi_s^N(X_s)h(\mathrm{X}_s)|^2 ds)] < \infty \tag{1.14}$$

is obviously satisfied. Hence, for any $N = 1, 2, \cdots$,

$$E[Z_t^N] = 1, \tag{1.15}$$

where

$$Z_t^N = \exp\bigg(\int_0^t \chi_s^N(\mathrm{X}_s)h(\mathrm{X}_s)^T dW_s - \frac{1}{2}\int_0^t \chi_s^N(X_s)|h(\mathrm{X}_s)|^2 ds\bigg). \tag{1.16}$$

To prove $E[Z_t] = 1$, we consider

$$\begin{aligned} E[Z_t^N] &= E[\chi_t^N(\mathrm{X}_t)Z_t^N] + E[(1-\chi_t^N(\mathrm{X}_t))Z_t^N] \\ &= E[\chi_t^N(\mathrm{X}_t)Z_t] + P\{\chi_t^N(\mathrm{X}_t) = 0\}. \end{aligned} \tag{1.17}$$

Hence by monotone convergence theorem,

$$\lim_{N\to\infty} E[\chi_t^N(\mathrm{X}_t)Z_t] = E[Z_t]. \tag{1.18}$$

On the other hand,

$$\lim_{N\to\infty} P\{\chi_t^N(\mathrm{X}_t) = 0\} = \lim_{N\to\infty} P\{\int_0^t |h(\mathrm{X}_s)|^2 ds > N\} = 0, \tag{1.19}$$

by (1.10) and Chebyshev inequality. Thus $E[Z_t] = 1$ and Z_t is an exponential martingale. By Girsanov theorem, there exists a new probability $\tilde{P}$ such that

$$\frac{d\tilde{P}}{dP} = Z_t. \tag{1.20}$$

It can be shown that under $\tilde{P}$, Y is a Brownian motion independent of X.
Denote by $\tilde{E}$ the expectation under the new probability measure $\tilde{P}$ and define
Define

$$\tilde{Z}_t = \exp\bigg(\int_0^t h(\mathrm{X}_s)^T dY_s - \frac{1}{2}\int_0^t |h(\mathrm{X}_s)|^2 ds\bigg). \tag{1.21}$$

The Kallianpur-Striebel formula [27] gives

$$\pi_t(\varphi) = \frac{\rho_t(\varphi)}{\rho_t(1)}, \tag{1.22}$$

where $\rho_t(\varphi) = \tilde{E}[\varphi(\mathrm{X}_t)\tilde{Z}_t|\mathcal{Y}_t]$ is called the unnormalized conditional distribution of X. One can prove that ρ_t satisfies the following evolution equation, called the Zakai equation:

$$\rho_t(\varphi) = \pi_0(\varphi) + \int_0^t \rho_s(A\varphi)ds + \int_0^t \rho_s(h^T\varphi)d\mathrm{Y}_s \quad \tilde{P} - a.s. \tag{1.23}$$

Here $\varphi \in D(A)$ and h^T denotes the transpose of h.

The structure of this paper is as follows: In section 2 we introduce the vortex model in terms of certain regularized kernels. In section 3 we consider the associated nonlinear filtering problem and prove uniqueness of measure valued solution to the Zakai equation. In section 4, an exposition of particle filtering is given. We prove the main convergence result in section 5, allowing for unboundedness in the observation process h.

2. Vortex method and stochastic vortex model

The equation of motion for interacting point vortices was first introduced by H. Helmholtz in a seminal paper published in 1858, in which he elucidated many properties. The standard vortex method in two dimensions was developed by L. Rosenhead [45], who approximated the motion of a two-dimensional vortex sheet by evolving in time the positions of point vortices(see for example: R. Krasny [30] and [31]). In this section, we will introduce the point vortex method and formulate the stochastic vortex model.

The Euler equations for vorticity-velocity field in two dimensions are as follows:

$$\begin{cases} \dfrac{\partial \omega}{\partial t} + \mathbf{u}\cdot\nabla\omega = 0, \qquad \text{in} \quad \mathrm{R}^2 \times \mathrm{R}_+, \\ \omega(x,0) = \omega_0(x), \qquad x \in \mathrm{R}^2, \\ \nabla\cdot\mathbf{u}(x,t) = 0, \qquad \text{in} \quad \mathrm{R}^2 \times \mathrm{R}_+, \\ \nabla\times\mathbf{u}(x,t) = \omega(x,t), \qquad \text{in} \quad \mathrm{R}^2 \times \mathrm{R}_+, \\ |\mathbf{u}(x,t)| \to \mathbf{u}_\infty(t) \quad \text{as} \quad |x| \to \infty, \quad t \in \mathrm{R}_+. \end{cases} \tag{2.1}$$

Let us formulate the point vortex approximation.

The velocity $\mathbf{u}$ is coupled through relations $\nabla\cdot\mathbf{u} = 0$ and $\nabla\times\mathbf{u} = \omega$, which imply

$$\triangle\mathbf{u} = -\nabla\times\omega. \tag{2.2}$$

Let G be the Green's function for the Laplacian operator in two dimensions and by K the rotational counterpart $\nabla\times G$, so that

$$G(x) = -\frac{1}{2\pi}\log(|x|), \quad K(x) = (2\pi|x|^2)^{-1}(-x_2, x_1). \tag{2.3}$$

The Boit-Savart law can be written as

$$\mathbf{u} = \mathbf{u}_\infty + K * \omega. \tag{2.4}$$

An approximation of Biot-Savart law obtained by removing the singularity of K, which makes the equation (2.3) to have very large value when two point vortices approach each other. In R. Krasny's calculation, a small positive constant is added to prevent the denominator in K from vanishing. Another approach is to replace K by a mollification K_ϵ in the following way.

Introduce a smooth cutoff function ζ such that $\int \zeta(x)dx = 1$. Define

$$\zeta_\epsilon(x) = \epsilon^{-2}\zeta(\frac{x}{\epsilon}), \quad \text{for} \quad \epsilon > 0. \tag{2.5}$$

Set $K_\epsilon = K * \zeta_\epsilon$, then

$$\frac{dx_i}{dt} = \mathbf{u}(x_i, t), \tag{2.6}$$

$$\mathbf{u} = K_\epsilon * \omega. \tag{2.7}$$

Above equations of motion for point vortices are given by A.J. Chorin [9].

First, We will talk about regularization of the singular kernel using cutoff functions.

Denote $x = (x_1, x_2)$, $r = |x| = \sqrt{x_1^2 + x_2^2}$ and consider a cutoff

$$\zeta(x) = \bar{\zeta}(|x|). \tag{2.8}$$

The kernel $G_\epsilon(x) = \bar{G}_\epsilon(|x|)$ satisfies

$$-\frac{1}{r}\frac{\partial}{\partial r}(r\frac{\partial \bar{G}_\epsilon}{\partial r}) = \bar{\zeta}_\epsilon, \tag{2.9}$$

from which we can derive

$$\frac{\partial \bar{G}_\epsilon}{\partial r} = \frac{1}{r}(x_2, -x_1)\frac{\partial \bar{G}_\epsilon}{\partial r} = -\frac{1}{r^2}(x_2, -x_1)\int_0^r s\bar{\zeta}_\epsilon(s)ds. \tag{2.10}$$

In the case of a Gaussian function

$$\bar{\zeta}(r) = \frac{1}{\pi}\exp(-r^2), \tag{2.11}$$

one obtains

$$K_\epsilon(x) = \frac{1}{2\pi r^2}(-x_2, x_1)[1 - \exp(-r^2/\epsilon^2)]. \tag{2.12}$$

A.J. Chorin [8] introduced an unbounded cutoff function with continuous kernel:

$$\zeta(x) = \begin{cases} \frac{1}{2\pi r} & \text{if } r \leq 1 \\ 0 & \text{if } r > 1, \end{cases} \tag{2.13}$$

and

$$K_\epsilon = \begin{cases} \frac{(-x_2, x_1)}{2\pi\epsilon} & \text{if } r \leq \epsilon \\ \frac{(-x_2, x_1)}{2\pi r^2} & \text{if } r > \epsilon. \end{cases} \tag{2.14}$$

which is derived from (2.5) and $K_\epsilon = K * \zeta_\epsilon$.
Let us consider this cutoff function and analyze the boundedness and differentiability of the mollified kernel.

For each fixed $\epsilon > 0$

$$\begin{aligned} |K_\epsilon(x)|^2 &= (|K_\epsilon(x)|^2)_{r\leq\epsilon} + (|K_\epsilon(x)|^2)_{r>\epsilon} \\ &= (\frac{1}{2\pi\epsilon})^2 r^2\big|_{r\leq\epsilon} + (\frac{1}{2\pi r^2})^2 r^2\big|_{r>\epsilon} \\ &\leq \frac{1}{4\pi^2} + \frac{1}{4\pi^2 r^2}\big|_{r>\epsilon} \\ &\leq \frac{1}{4\pi^2}(1 + \frac{1}{\epsilon^2}) < +\infty. \end{aligned} \tag{2.15}$$

In the following, we check the differentiability of K_ϵ by studying its gradient norm $|\nabla K_\epsilon|^2$.

$$\nabla K_\epsilon = \begin{cases} \begin{pmatrix} 0 & \frac{1}{2\pi\epsilon} \\ -\frac{1}{2\pi\epsilon} & 0 \end{pmatrix} & \text{if } r \leq \epsilon \\ \begin{pmatrix} \frac{x_1x_2}{\pi r^4} & \frac{x_2^2 - x_1^2}{2\pi r^4} \\ \frac{x_2^2 - x_1^2}{2\pi r^4} & \frac{-x_1x_2}{\pi r^4} \end{pmatrix} & \text{if } r > \epsilon. \end{cases} \tag{2.16}$$

$$\begin{aligned} |\nabla K_\epsilon|^2 &= (|\nabla K_\epsilon|^2)_{r\leq\epsilon} + (|\nabla K_\epsilon|^2)_{r>\epsilon} \\ &= 2(\frac{1}{2\pi\epsilon})^2 + (\frac{1}{2\pi})^2[2(\frac{2x_1x_2}{r^4})^2 + 2(\frac{x_2^2-x_1^2}{r^4})^2]\big|_{r>\epsilon} \\ &\leq \frac{1}{2\pi^2\epsilon^2} + \frac{1}{2\pi^2\epsilon^4} \\ &= \frac{1}{2\pi^2\epsilon^2}(1+\frac{1}{\epsilon^2}) < \infty. \end{aligned} \tag{2.17}$$

Hence K_ϵ has bounded first derivative. It can be shown by mean value theorem that K_ϵ is globally Lipschitz.

We introduce a stochstic counterpart of the deterministic vortex model (2.6) given by A.J. Chorin [9] in the following way. Denote $X_i(t)$ as the position for the i-th point vortice with initial data ξ_i, then

$$X_i(t) = \xi_i + \int_0^t \mathbf{u}_{\epsilon,s}(X_i(s))ds + \int_0^t \sigma(X_i(s))dW_s, \quad \text{for} \quad i = 1, \cdots, N \tag{2.18}$$

with

$$\mathbf{u}_{\epsilon,t}(x) = \sum_{j=1}^N \alpha_j K_\epsilon(x - x_j(t)), \qquad \forall x \in \mathrm{R}^2. \tag{2.19}$$

Equation (2.18) defines the signal process of the nonlinear filtering problem we will study in next section.

3. Nonlinear filtering problem for vortex model

Recall the nonlinear filtering problem defined in first section

$$\text{signal}: \qquad dX_t = f(X_t)dt + \sigma(X_t)dW_t \tag{3.1}$$

$$\text{observation}: \qquad dY_t = h(X_t)dt + dB_t \tag{3.2}$$

$X_t = (X_1(t), X_2(t), \cdots, X_N(t))$, N is the number of point vortices. $X_i(t) \in \mathrm{R}^2$ represents the position for i-th point vortice at time t. Y_t is the m-dimensional observation process. Note that in this paper we focus on the uncorrelated noises from signal and observation processes. For unique solvability for correlated case see B.L. Rozovskii [48] and S.S. Sritharan [49], [51].

For the stochastic vortex equations (2.18) and (2.19), we have:

$f(x_t) = (f(x_1), \cdots, f(x_N))$, $\quad f(x_i) = \sum_{k=1}^N a_k K_\epsilon(x_i - x_k)$, $\qquad i = 1, \cdots, N.$

a_i is the associated vorticity intensity for i-th point vortice.

K_ϵ is the regularized Biot-Savart kernel (2.14).

$\sigma(\cdot) : \mathrm{R}^{2N} \to \mathrm{R}^{2N} \times \mathrm{R}^{2N}$ is bounded and globally Lipschitz, i.e.

$$\|\sigma(x) - \sigma(y)\| \leq C_1|x - y| \qquad \text{with} \quad \|\sigma(x)\| := \sqrt{\sum_{i=1}^{2N}\sum_{j=1}^{2N}\sigma_{ij}(x)}, \tag{3.3}$$

for some constant C_1.

$a(x) := \frac{1}{2}\sigma(x)\sigma^T(x)$ is locally Lipschitz and position definite.

W is $2N$-dimensional Brownian motion.

$h(\cdot) : \mathrm{R}^{2N} \to \mathrm{R}^m$ is globally Lipschitz and satisfies the linear growth condition (1.9).

B is an m-dimensional Brownian motion independent of W.

Since K_ϵ is globally Lipschitz, using triangle inequality and Jensen's inequality, we have

$$
\begin{aligned}
|f(x) - f(y)|^2 &= |f(x_1) - f(y_1)|^2 + \cdots |f(x_N) - f(y_N)|^2 \\
&= |\sum_{k=1}^{N} a_k K_\epsilon(x_1 - x_k) - \sum_{k=1}^{N} a_k K_\epsilon(y_1 - y_k)|^2 + \\
&\cdots + |\sum_{k=1}^{N} a_k K_\epsilon(x_N - x_k) - \sum_{k=1}^{N} a_k K_\epsilon(y_N - y_k)|^2 \\
&= |\sum_{k=1}^{N} a_k (K_\epsilon(x_1 - x_k) - K_\epsilon(y_1 - y_k))|^2 + \\
&\cdots + |\sum_{k=1}^{N} a_k (K_\epsilon(x_N - x_k) - K_\epsilon(y_N - y_k))|^2 \\
&\leq |\sum_{k=1}^{N} a_k C_{k,1}(|x_1 - y_1| + |x_k - y_k|)|^2 + \\
&\cdots + |\sum_{k=1}^{N} a_k C_{k,N}(|x_N - y_N| + |x_k - y_k|)|^2 \\
&\leq 2N \sum_{k=1}^{N} a_k^2 C_{k,1}^2 (|x_1 - y_1|^2 + |x_k - y_k|^2) + \\
&\cdots + 2N \sum_{k=1}^{N} a_k^2 C_{k,N}^2 (|x_N - y_N|^2 + |x_k - y_k|^2) \\
&\leq \max_{i,j}\{C_{i,j}^2\} 2N \sum_{k=1}^{N} a_k^2 (|x_1 - y_1|^2 + \cdots + |x_N - y_N|^2) + \\
&\max_{i,j}\{C_{i,j}^2\} 2N^2 \sum_{k=1}^{N} a_k^2 |x_k - y_k|^2 \\
&\leq 4N^2 \max_{i,j}\{C_{i,j}^2\} \max_k \{a_k^2\} |x - y|^2.
\end{aligned}
\tag{3.4}
$$

Here $C_{i,j}$ are the C^1-coefficients when applying mean-value theorem for K_ϵ.

Now we have the global Lipschitz condition for both f and σ.
Suppose $||\sigma(x(0))|| < \infty$ and $|f(x(0))| < \infty$, then we can prove, there exist C_3 and C_4, such that

$$\|\sigma(x)\|^2 \leq C_3(1+|x|)^2, \tag{3.5}$$

and

$$|f(x)| \leq C_4(1+|x|). \tag{3.6}$$

Also, there exists C_5, such that

$$\|a(x)\| \leq C_5(1+|x|^2). \tag{3.7}$$

Under these conditions, equation (3.1) has a unique strong solution by I.I. Gihman and A.V. Skorokhod [20].

Let $\varphi : R^{2N} \to R^1$ be a function in C_b^2, which is twice differentiable with all its derivatives and the function itself be bounded. Denote by $\pi_t(\varphi) := E[\varphi(X(t))|Y_t]$, it satisfies the well-known Fujisaki-Kallianpur-Kunita(FKK) equation [19]:

$$d\pi_t(\varphi) = \pi_t(L\varphi)dt + \Big(\pi_t(M\varphi) - \pi_t(\varphi)\pi_t(h)\Big)\Big(dY_t - \pi_t(h)dt\Big), \tag{3.8}$$

where

$$L\varphi = \sum_{i,j=1}^{2N} a^{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}\varphi(x) + \sum_{i=1}^{2N} f^i(t,x)\frac{\partial}{\partial x_i}\varphi(x), \tag{3.9}$$

$$M\varphi = h(t,x)\varphi(x), \tag{3.10}$$

$$a(x) := \frac{1}{2}\sigma(t,x)\sigma^T(t,x). \tag{3.11}$$

One can further show that the smooth density $\pi_t(x)$ satisfies

$$\begin{aligned} d\pi_t(x) = L^*\pi_t(x)dt + \Big(&M^*\pi_t(x) \\ &- \pi_t(x)\int_{R^{2N}} h(t,x)\pi_t(x)dx\Big)\Big(dY_t - \int_{R^{2N}} h(t,x)\pi_t(x)dxdt\Big). \end{aligned} \tag{3.12}$$

Here

$$dY_t - \int_{R^{2N}} h(t,x)\pi_t(x)dxdt \tag{3.13}$$

is called the innovation process, L^* and M^* denote the adjoint operators of L and M respectively.
The above equation is referred to as the Kushner equation [33] and is difficult to analyze because of its nonlinear structure. Using Girsanov theorem, Zakai [52] showed that if the transition probability is absolutely continuous, the density $\pi_t(x)$ can be represented as

$$\pi_t(x) = \rho_t(x)/\int_{R^{2N}} \rho_t(x)dx \tag{3.14}$$

with $\rho_t(x)$ satisfying a linear stochastic partial differential equation

$$d\rho_t(x) = L^*\rho_t(x)dt + M^*\rho_t(x)dY_t. \tag{3.15}$$

The function ρ is usually referred to as the unnormalized filtering density and equation (3.15) is called the Zakai equation.
In the following we will study the uniqueness of the measure valued solution to the Zakai equation, with the help of unique solvability theorem of backward Kolmogorov equation. This method is called the PDE approach.
The uniqueness of measure valued solution to the Zakai equation is given by B.L. Rozovskii [48], we extend his result here by allowing h to be an unbounded function in the sense described in (1.9).
Denote $\mathcal{M}(\mathrm{R}^{2N})$ the set of totally finite, countably additive signed measure with the topology of weak convergence. If $\mu \in \mathcal{M}(\mathrm{R}^{2N})$, we denote

$$\langle \mu, f \rangle := \int_{\mathrm{R}^{2N}} f(x)\mu(dx). \tag{3.16}$$

Definition 3.1. An Y_t-adapted stochastic process μ_t taking value in $\mathcal{M}(\mathrm{R}^{2N})$ is said to be a measure valued solution to the Zakai equation corresponding to the initial condition $\mu_0(dx) = P(x_0 \in dx|\mathcal{Y}_0)$, if $\langle |\mu_\cdot|, 1\rangle \in \mathrm{L}_2([0,T] \times \Omega; dt \times dP)$, for every $t \leq T$ and $T < \infty$, $\langle |\mu_t|, 1\rangle \in \mathrm{L}_2(\Omega, dP)$ and for any $\psi \in \mathrm{C}_b^2(\mathrm{R}^{2N})$ the following equality holds P-almost surely.

$$\begin{aligned}\langle \mu_t, \psi\rangle = \langle \mu_0, \psi\rangle + &\int_0^t \langle \mu_s, \mathrm{L}\psi\rangle ds \\ &+ \int_0^t \langle \mu_s, \mathrm{M}\psi\rangle dY_s, \quad \forall t \in [0,T].\end{aligned} \tag{3.17}$$

Let $b \in C([0,T], \mathrm{R}^{2N})$ and be non-negative, define

$$\mathrm{L}_b\psi(x) := \mathrm{L}\psi(x) + b(t)\mathrm{M}\psi(x). \tag{3.18}$$

Consider the backward Cauchy problem

$$-\frac{\partial \eta^b(t,x)}{\partial t} = \mathrm{L}_b\eta^b(t,x), \qquad t < T_0 \quad x \in \mathrm{R}^m, \tag{3.19}$$

$$\eta^b(T_0, x) = \beta(x). \tag{3.20}$$

The coefficients of the operators L and M are continuous in t and we can show that they are locally Lipschitz. Problem (3.19) and (3.20) have a solution in $\mathrm{C}_b^{1,2}([0,T_0] \times \mathrm{R}^{2N})$ for every $T_0 \leq T$ by S.D. Eidel'man [17].

Theorem 3.2. *Assume $a(x)$, $f(x)$ and $h(x)$ are locally Lipschitz in x, h satisfies*

$$E\Big(\int_0^t |h(X(s))|^2 ds\Big) < \infty, \quad t \in [0,T], \tag{3.21}$$

where X is solution to the signal process (3.1), then the measure valued solution to the Zakai equation (3.17) is unique.

Proof. Assume μ_t is a measure valued solution to the Zakai equation, such that for each $\eta \in \mathrm{C}_b^{1,2}([0,T] \times \mathrm{R}^{2N})$,

$$\begin{aligned}\langle \mu_t, \eta(t) \rangle = \langle \mu_0, \eta(0) \rangle + \int_0^t \langle \mu_s, \frac{\partial}{\partial s}\eta(s) + \\ \mathrm{L}\eta(s) \rangle ds + \int_0^t \langle \mu_s, \mathrm{M}\eta(s) \rangle dY_s.\end{aligned} \tag{3.22}$$

Now fix $b \in \mathrm{C}([0,T], \mathrm{R}^m)$. Let $\eta(t) = \eta^b(t)$ be a solution to the Cauchy problem (3.19),(3.20) for this b. Define

$$q_t := \exp\bigg(\int_0^t b(s) dY_s - \frac{1}{2} \int_0^t |b(s)|^2 ds \bigg), \tag{3.23}$$

$$p_t^{-1} := \exp\bigg(- \int_0^t h(X(s))^T dY_s + \frac{1}{2} \int_0^t |h(X(s))|^2 ds \bigg), \tag{3.24}$$

$$\gamma_t := q_t p_t^{-1}. \tag{3.25}$$

Applying Ito formula for q_t, p_t^{-1} and γ_t respectively,

$$dq_t = q_t b(t) h(X(t)) dt + q_t b(t) dB_t, \tag{3.26}$$

$$dp_t^{-1} = -h(X(t)) p_t^{-1} dB_t, \tag{3.27}$$

$$d\gamma_t = \gamma_t b(t) dB_t - \gamma_t h(X(t)) dB_t. \tag{3.28}$$

Thus,

$$\begin{aligned}\langle \mu_t, \eta^b(t) \rangle \gamma_t = \langle \mu_0, \eta^b(0) \rangle + \int_0^t \langle \mu_s, \frac{\partial}{\partial s}\eta^b(s) + \mathrm{L}_b \eta^b(s) \rangle \gamma_s ds + \\ \int_0^t \langle \mu_s, \eta^b(s) \rangle b(s) \gamma_s dB_s.\end{aligned} \tag{3.29}$$

The second term on the right hand side is zero because $\eta(s)$ is a solution to the Cauchy problem (3.19),(3.20). The third term on the right hand side is a martingale, which can be shown by the truncation function technique.
Now take expectation on both sides, we get

$$E(\langle \mu_{T_0}, \beta \rangle \gamma_{T_0}) = E\eta(0, X_0). \tag{3.30}$$

By Feymann-Kac formula,

$$E\eta(0, X) = E[\beta(X(T_0)) \exp \int_0^{T_0} h(X(s)) b(s) ds], \tag{3.31}$$

where X is the solution to the signal process

$$X(t) = x_0 + \int_0^t f(X(s)) ds + \int_0^t \sigma(X(s)) dW_s. \tag{3.32}$$

By Girsanov's theorem

$$\begin{aligned}E[\beta(X(T_0)) \exp[\int_0^{T_0} h(X(s)) b(s) ds]] &= E[\beta(X(T_0)) q_{T_0}] \\ &= \tilde{E}[\beta(X(T_0)) p_{T_0} q_{T_0}] \\ &= \tilde{E}[\tilde{E}[\beta(X(T_0)) p_{T_0} | \mathcal{Y}_{T_0}] q_{T_0}].\end{aligned} \tag{3.33}$$

Here we used the fact that p_t is an exponential martingale satisfying $E[p_t] = 1$ as we showed in the introduction. The first equality holds because

$$
\begin{aligned}
&E[\beta(X(T_0))q_{T_0}] \\
&= E[\beta(X(T_0))\exp\Big(\int_0^{T_0} b(s)dY_s - \frac{1}{2}\int_0^{T_0} |b(s)|^2 ds\Big)] \\
&= E[\beta(X(T_0))\exp\Big(\int_0^{T_0} h(X(s))b(s)ds\Big)\exp\Big(\int_0^{T_0} b(s)dB_s - \frac{1}{2}\int_0^{T_0} |b(s)|^2 ds\Big)] \\
&= E[\beta(X(T_0))\exp\Big(\int_0^{T_0} h(X(s))b(s)ds\Big)]E[\exp\Big(\int_0^{T_0} b(s)dB_s - \frac{1}{2}\int_0^{T_0} |b(s)|^2 ds\Big)] \\
&= E[\beta(X(T_0))\exp\Big(\int_0^{T_0} h(X(s))b(s)ds\Big)],
\end{aligned}
\tag{3.34}
$$

since X_t and Y_t are independent and $\exp\Big(\int_0^{T_0} b(s)dB_s - \frac{1}{2}\int_0^{T_0} |b(s)|^2 ds\Big)$ is an exponential martingale. On the other hand,

$$E(\langle \mu_{T_0}, \beta\rangle \gamma_{T_0}) = \tilde{E}(\langle \mu_T, \beta\rangle q_{T_0}), \tag{3.35}$$

hence

$$\tilde{E}[\tilde{E}[\beta(x(T_0))p_{T_0}|\mathcal{Y}_{T_0}]q_{T_0}] = \tilde{E}(\langle \mu_{T_0}, \beta\rangle q_{T_0}). \tag{3.36}$$

Note that $Y(t)$ is a Wiener martingale on $(\Omega, \mathcal{F}, \tilde{P})$. Furthermore, N. Wiener pointed out that $\{q_{T_0} := q_{T_0}(b), \quad b \in \mathrm{C}([0,T], \mathrm{R}^d)\}$ is total in $\mathrm{L}_2(\Omega, Y_{T_0}, \tilde{P})$, which means that if $\beta \in \mathrm{L}_2(\Omega, Y_{T_0}, P)$ and $\tilde{E}[\beta q_{T_0}(b)] = 0$ for all $b \in \mathrm{C}([0,T]; \mathrm{R}^m)$, then $\beta = 0$ P-a.s.[38]. Therefore

$$\langle \mu_{T_0}, \beta\rangle = \tilde{E}[\beta(x(T_0))p_{T_0}|\mathcal{Y}_{T_0}] \quad P-a.s. \tag{3.37}$$

The proof is complete. □

Remark: The idea of this proof is to show that all the measure valued solution to the Zakai equation can be represented as the the conditional expectation in the proof above. Since it is proved that the conditional expectation given observation $\mathcal{Y}_t$ satisfies (3.22), it guarantees existence and uniqueness of the measure valued solution to the Zakai equation. Our proof improves the result of B. Rozovskii's [48] in the sense that here we consider h as an unbounded function satisfying (3.22). By B. Ferrairo's truncation function approach, we are able to show p_t is an exponential martingale and used Girsanov's theorem in the proof.

4. Particle filter method

In this section, we describe the basic idea of particle filter method and some of its properties. We will focus on the approximate solution π_t^n to the FKK equation. Some details of algorithm are explained here and interested readers can refer to [4].

At the initial time, n particles have equal weights $\frac{1}{n}$ and positions ξ_j^n for $j =$

$1, \cdots, n$. ξ_j^n are independent, identically distributed random variables with common distribution π_0. The approximating measure at $t = 0$ is

$$\pi_0^n = \frac{1}{n} \sum_{j=1}^{n} \delta_{\xi_j^n}. \tag{4.1}$$

Now partition the time interval $[0, \infty)$ to be sub-intervals with same length ε. For $t \in [i\varepsilon, (i+1)\varepsilon)$, $i = 0, 1, \cdots$,

$$X_j^n(t) = X_j^n(i\varepsilon) + \int_{i\varepsilon}^{t} f(X_j^n(s))ds + \int_{i\varepsilon}^{t} \sigma(X_j^n(s))dW_s^{(j)}, \qquad j = 1, \cdots, n, \tag{4.2}$$

meaning the particles all move with the same law as the signal X_t. The weight for particle j at time t is

$$\bar{a}_j^n(t) := \frac{a_j^n(t)}{\sum_{k=1}^{n} a_k^n(t)}, \tag{4.3}$$

where

$$a_j^n(t) = 1 + \sum_{k=1}^{m} \int_{i\varepsilon}^{t} a_j^n(s) h^k(X_j^n(s)) dY_s^k. \tag{4.4}$$

Hence also:

$$a_j^n(t) = \exp(\int_{i\varepsilon}^{t} h(X_j^n(s))^T dY_s - \frac{1}{2} \int_{i\varepsilon}^{t} |h(X_j^n(s))|^2 ds). \tag{4.5}$$

Define

$$\pi_t^n := \sum_{j=1}^{n} \bar{a}_j^n(t) \delta_{X_j^n(t)}, \quad t \geq 0, \tag{4.6}$$

and

$$\rho_t^n := \frac{1}{n} \sum_{j=1}^{n} a_j^n(t) \delta_{X_j(t)}, \quad t \geq 0. \tag{4.7}$$

Here π_t^n approximates solution of the FKK equation and ρ_t^n approximates solution of the Zakai equation.

At the end of the interval, each particle branches into a random number of particles. Each offspring particle initially inherits the spatial position of its parent. After branching all the particles are reindexed (from 1 to n) and all the unnormalized weights are reinitialized back to 1. Denote

- $j' = 1, 2, \cdots, n$ as the particle index prior to the branching event.
- $j = 1, 2, \cdots, n$ as the particle index after the branching event.

Define

$$\mathcal{F}_{(i+1)\varepsilon} = \sigma\{\mathcal{F}_s, \quad s \leq (i+1)\varepsilon\}. \tag{4.8}$$

Let $\lambda_{j'}^{n,(i+1)\varepsilon}$ be the number of offsprings produced by the j'th particle at time $(i+1)\varepsilon$, then $\lambda_{j'}^{n,(i+1)\varepsilon}$ is $\mathcal{F}_{(i+1)\varepsilon}$-adapted.
Define

$$\lambda_{j'}^{n,(i+1)\varepsilon} = \begin{cases} [n\bar{a}_{j'}^{n,(i+1)\varepsilon}] & \text{with probability } 1 - \{n\bar{a}_{j'}^{n,(i+1)\varepsilon}\} \\ [n\bar{a}_{j'}^{n,(i+1)\varepsilon}] + 1 & \text{with probability } \{n\bar{a}_{j'}^{n,(i+1)\varepsilon}\}, \end{cases} \tag{4.9}$$

where $[x]$ is the largest integer smaller than x and $\{x\}$ is the fractional part of x; which is, $\{x\} = x - [x]$. $\bar{a}_{j'}^{n,(i+1)\varepsilon}$ is the value of the particle's weight immediately prior to the branching. That is $\bar{a}_{j'}^{n,(i+1)\varepsilon} = \lim_{t\uparrow(i+1)\varepsilon} \bar{a}_{j'}^{n}(t)$.
Define

$$\mathcal{F}_{(i+1)\varepsilon-} = \sigma\{\mathcal{F}_s, \quad s < (i+1)\varepsilon\}. \tag{4.10}$$

By the definition,

$$E[\lambda_{j'}^{n,(i+1)\varepsilon}|\mathcal{F}_{(i+1)\varepsilon-}] = n\bar{a}_{j'}^{n,(i+1)\varepsilon}. \tag{4.11}$$

The conditional variance is

$$\begin{aligned}
&E[(\lambda_{j'}^{n,(i+1)\varepsilon})^2|\mathcal{F}_{(i+1)\varepsilon-}] - (E[\lambda_{j'}^{n,(i+1)\varepsilon}|\mathcal{F}_{(i+1)\varepsilon-}])^2 \\
&= ([n\bar{a}_{j'}^{n,(i+1)\varepsilon}])^2(1 - \{n\bar{a}_{j'}^{n,(i+1)\varepsilon}\}) + ([n\bar{a}_{j'}^{n,(i+1)\varepsilon}] + 1)^2(\{n\bar{a}_{j'}^{n,(i+1)\varepsilon}\}) - (n\bar{a}_{j'}^{n,(i+1)\varepsilon})^2 \\
&= \{n\bar{a}_{j'}^{n,(i+1)\varepsilon}\}(1 - \{n\bar{a}_{j'}^{n,(i+1)\varepsilon}\}).
\end{aligned} \tag{4.12}$$

It can be shown that $\lambda_j^{n,(i+1)\varepsilon}$ has conditional minimal variance in the set of all integer valued random variables ξ such that $E[\xi|\mathcal{F}_{(i+1)\varepsilon-}] = n\bar{a}_{j'}^{n,(i+1)\varepsilon}, \quad j = 1, \cdots, n$. See [4].
We wish to control the branching process so that the number of particles in the system remains constant n:

$$\sum_{j'=1}^{n} \lambda_{j'}^{n,(i+1)\varepsilon} = n. \tag{4.13}$$

Thus $\lambda_{j'}^{n,(i+1)\varepsilon}, \quad j' = 1, \cdots, n$ are correlated.

Proposition 9.3 in A. Bain and D. Crisan [4] shows that $\lambda_j^n, \quad j = 1, \cdots, n$ have the following properties:

- $\sum_{j=1}^{n} \lambda_j^n = n$.
- For any $j = 1, \cdots, n$, we have $E[\lambda_j^n] = n\bar{a}_j^n$.
- For any $j = 1, \cdots, n$, λ_j^n has minimal variance, specifically

$$E[(\lambda_j^n - n\bar{a}_j^n)^2] = \{n\bar{a}_j^n\}(1 - \{n\bar{a}_j^n\}). \tag{4.14}$$

- For any $k = 1, \cdots, n-1$, the random variables $\lambda_{1:k}^n = \sum_{j=1}^{k} \lambda_j^n$ and $\lambda_{k+1:n}^n = \sum_{j=k+1}^{n} \lambda_j^n$ have variance

$$\begin{aligned}
E[(\lambda_{1:k}^n - n\bar{a}_{1:k}^n)^2] &= \{n\bar{a}_{1:k}^n\}(1 - \{n\bar{a}_{1:k}^n\}), \\
E[(\lambda_{k+1:n}^n - n\bar{a}_{k+1:n}^n)^2] &= \{n\bar{a}_{k+1:n}^n\}(1 - \{n\bar{a}_{k+1:n}^n\}).
\end{aligned} \tag{4.15}$$

  where $\bar{a}_{1:k}^n = \sum_{j=1}^{k} \bar{a}_j^n$ and $\bar{a}_{k+1:n}^n = \sum_{j=k+1}^{n} \bar{a}_j^n$.
- For $1 \le i < j \le n$, λ_i^n and λ_j^n are non-positively correlated, that is

$$E[(\lambda_i^n - n\bar{a}_i^n)(\lambda_j^n - n\bar{a}_j^n)] \le 0. \tag{4.16}$$

Remark As boundedness of function h is not used in the proof of above properties, they hold for both bounded and unbounded h.

If the system does not undergo any corrections, that is $\varepsilon = \infty$, then the above method is simply the Monte-Carlo method. The convergence of the Monte-Carlo approximation for nonlinear filtering problem has been studied by G.N. Milstein and M.V. Tretyakov in [37]. It has the drawback that particles wander away from the signal's trajectory, which force the un-normalized weights to become infinitesimally small. In particle filter, the branching correction procedure is introduced to remove the unlikely particles and multiply those situated in the right areas.

However, the branching procedure introduces randomness into the system as it replaces each weight with a random number of offsprings. The distribution of the number of offsprings has to be chosen with great care to minimize the variance. That is, as the mean number of offsprings is pre-determined, we should choose the $\lambda_{j'}^n$'s to have the smallest possible variance amongst all integer valued random variables with the given mean $n\bar{a}_{j'}^n$. The way we defined $\lambda_{j'}^n$ above has the minimal variance.[4]

5. Convergence result of numerical methods

In this section, we will prove the convergence for the approximation of the solution to Zakai equation using Monte Carlo method and particle filter method. The latter has convergence rate $\frac{1}{n}$.

Definition 5.1. An adapted, càdlàg process X is a local martingale if there exists a sequence of increasing stopping time $\{T_n\}$, with $\lim_{n\to\infty} T_n = \infty$ a.s. such that $X_{t\wedge T_n}$ is a uniformly integrable martingale for each n.

Definition 5.2. A stopping time T reduces a process X if $X_{t\wedge T}$ is a uniformly integrable martingale.

We need the following lemma [43] and theorem [38].

Lemma 5.3. *Let X be a càdlàg process and let T_N be a sequence of stopping times increasing to ∞ a.s. such that $X_{t\wedge T_N}$ is a local martingale for each N, then X is a local martingale.*

Theorem 5.4 (Doob's Martingale Convergence Theorem)**.** *A process X_t is uniformly integrable if and only if there exists an integrable random variable $\bar{X}$ such that*

$$X_t \to \bar{X} \quad as \quad t \to \infty, \quad P - a.s \quad and \quad \mathrm{L}^1. \tag{5.1}$$

Doob's Martingale convergence theorem tells us that if X is a uniformly integrable martingale, then X_t converges to $X_\infty = Y$ in L^1 as well as almost surely.

The following property is important in proving convergence for Monte Carlo method and is an improvement of the estimate from A. Bain and D. Crisan [4] by allowing h to be unbounded.

Lemma 5.5. *For any $t \geq 0$ and $p \geq 2$, if $|h(X(s,\omega)))| \leq f(s)$ for all $(s,\omega) \in \Omega \times [0,\infty)$ with*

$$\int_0^t |f(s)|^2 ds < \infty, \quad for \quad each \quad t > 0. \tag{5.2}$$

Then there exists a constant $C_p(t)$ such that

$$\tilde{E}[(\tilde{Z}_t)^p] \leq C_p(t). \tag{5.3}$$

Proof. Recall that

$$\tilde{Z}_t = \exp\bigg(\int_0^t h(X_s)^T dY_s - \frac{1}{2}\int_0^t |h(X_s)|^2 ds\bigg), \tag{5.4}$$

which can be written as

$$\tilde{Z}_t = 1 + \int_0^t \tilde{Z}_s h(X_s)^T dY_s. \tag{5.5}$$

By Ito formula,

$$\begin{aligned}(\tilde{Z}_t)^p = 1 &+ p\int_0^t (\tilde{Z}_s)^{p-1}\tilde{Z}_s h(X_s)^T dY_s \\ &+ \frac{p(p-1)}{2}\int_0^t (\tilde{Z}_s)^{p-2}(\tilde{Z}_s)^2 h^T(X_s)h(X_s)ds.\end{aligned} \tag{5.6}$$

Applying Lemma 5.3 to equations (5.5) and (5.6), one can show that

$$(\tilde{Z}_t)^p - \frac{p(p-1)}{2}\int_0^t (\tilde{Z}_s)^p h^T(X_s)h(X_s)ds \tag{5.7}$$

and $\tilde{Z}_t$ are local martingales. Let $\{T_N^1\}$ be a sequence of stopping times reducing the local martingale

$$(\tilde{Z}_t)^p - \frac{p(p-1)}{2}\int_0^t (\tilde{Z}_s)^p h^T(X_s)h(X_s)ds. \tag{5.8}$$

let $\{T_N^2\}$ be a sequence of stopping times reducing the local martingale $\tilde{Z}_t$. Then $\{T_N := T_N^1 \wedge T_N^2\}$ defines a sequence of stopping times that reduces both $\tilde{Z}_t$ and

$$(\tilde{Z}_t)^p - \frac{p(p-1)}{2}\int_0^t (\tilde{Z}_s)^p h^T(X_s)h(X_s)ds, \tag{5.9}$$

which means $\tilde{Z}_{t\wedge T_N}$ and

$$(\tilde{Z}_{t\wedge T_N})^p - \frac{p(p-1)}{2}\int_0^{t\wedge T_N} (\tilde{Z}_s)^p h^T(X_s)h(X_s)ds \tag{5.10}$$

are uniformly integrable martingales.

Now take expectation on both sides of (5.6):

$$\begin{aligned}\tilde{E}[(\tilde{Z}_{t\wedge T_N})^p] &= 1 + \frac{p(p-1)}{2}\tilde{E}\int_0^{t\wedge T_N}(\tilde{Z}_s)^p|h(X_s)|^2 ds \\ &\le 1 + \frac{p(p-1)}{2}\int_0^{t\wedge T_N}|f(s)|^2\tilde{E}[(\tilde{Z}_s)^p]ds \\ &\le 1 + \frac{p(p-1)}{2}\int_0^{t}|f(s)|^2\tilde{E}[(\tilde{Z}_{s\wedge T_N})^p]ds.\end{aligned} \tag{5.11}$$

By Gronwall's inequality we have

$$\tilde{E}[(\tilde{Z}_{t\wedge T_N})^p] \le \exp\biggl(\frac{p(p-1)}{2}\int_0^t |f(s)|^2 ds\biggr). \tag{5.12}$$

Denote bound on the right hand side of the inequality by $C_p(t)$, we have:

$$\tilde{E}[(\tilde{Z}_{t\wedge T_N})^p] \le C_p(t) < \infty. \tag{5.13}$$

Applying Doob's L^p maximal inequality to the process $(\tilde{Z}_{t\wedge T_N})^p$ yields

$$\tilde{E}[\sup_{0\le s\le t\wedge T_N}(\tilde{Z}_s)^p] \le 4\tilde{E}[(\tilde{Z}_{t\wedge T_N})^p] \le 4C_p(t). \tag{5.14}$$

It follows from (5.14) that $\{(\tilde{Z}_{s\wedge T_N})^p : s \in [0,T]\}$ is uniformly integrable. Hence, by Doob's martingale convergence theorem, $(\tilde{Z}_{t\wedge T_N})^p$ converges both a.s. and in L^1 to $(\tilde{Z}_t)^p$ as $N \to \infty$. Hence by (5.13) and Fatou's lemma

$$\tilde{E}[(\tilde{Z}_t)^p] \le C_p(t). \tag{5.15}$$

The conclusion is proved. $\square$

We also need the following lemma to prove the main result. Define $\|\phi\|_\infty = \sum \sup_{x\in R^{2n}} |\phi(x)|$.

Lemma 5.6. *Let us assume the conditions in Lemma 5.5 and that $\phi \in C_b(R^{2n})$. Define $\mathcal{Y}_t$-adapted random variable $C_\phi(t)$ as*

$$C_\phi(t) = \tilde{E}\Bigl[(\phi(X_t)\tilde{Z}_t - \rho_t(\phi))^2|\mathcal{Y}_t\Bigr], \tag{5.16}$$

then

$$\tilde{E}[C_\phi(t)] < 4\|\phi\|_\infty^2 C_2(t). \tag{5.17}$$

Proof. We have by Jensen's inequality

$$\begin{aligned}\tilde{E}[C_\phi(t)] &= \tilde{E}\Bigl[\tilde{E}\Bigl(\phi(X_t)\tilde{Z}_t - \rho_t(\phi)\Bigr)^2|\mathcal{Y}_t\Bigr] \\ &= \tilde{E}\Bigl[\Bigl(\phi(X_t)\tilde{Z}_t - \rho_t(\phi)\Bigr)^2\Bigr] \\ &\le 2\tilde{E}\Bigl[\Bigl(\phi(X_t)\tilde{Z}_t\Bigr)^2 + \Bigl(\rho_t(\phi)\Bigr)^2\Bigr] \\ &\le 2||\phi||_\infty^2\Bigl(\tilde{E}[\tilde{Z}_t^2] + \tilde{E}[(\rho_t(1))^2]\Bigr).\end{aligned} \tag{5.18}$$

Here we used the fact that ϕ is uniformly bounded. For second term $\tilde{E}[(\rho_t(1))^2]$, we use Jensen's inequality for conditional expectation and Lemma 5.5:

$$\begin{aligned}\tilde{E}[(\rho_t(1))^2] &= \tilde{E}\bigg[(\tilde{E}[\tilde{Z}_t|\mathcal{Y}_t])^2\bigg] \\ &\le \tilde{E}\bigg[\tilde{E}[\tilde{Z}_t^2|\mathcal{Y}_t]\bigg] \\ &= \tilde{E}[\tilde{Z}_t^2] < \infty.\end{aligned} \tag{5.19}$$

Therefore

$$\tilde{E}[C_\phi(t)] < 4\|\phi\|_\infty^2\tilde{E}[(\tilde{Z}_t)^2] \le 4\|\phi\|_\infty^2 C_2(t) \tag{5.20}$$

by choosing $p = 2$ in Lemma 5.5. □

Now, we state the convergence result for Monte Carlo method about $\rho_t^n(\phi)$ to $\rho_t(\phi)$ for any $\phi \in \mathrm{C}_b(\mathbb{R}^{2n})$. This would imply that ρ_t^n converges to ρ_t as measure-valued random variables.

Theorem 5.7. *Let the coefficients σ and f be globally Lipschitz, with finite initial data $\sigma(X_0)$ and $f(X_0)$. h satisfies the condition in Lemma 5.5, then for any $T > 0$ and $\phi \in C_b(\mathrm{R}^{2n})$,*

$$\tilde{E}[(\rho_t^n(\phi) - \rho_t(\phi))^2] \le \frac{4\|\phi\|_\infty^2 C_2(t)}{n}, \qquad t \in [0, T]. \tag{5.21}$$

In particular, ρ_t^n converges in expectation to ρ_t.

Proof. Let $a_j, j = 1, \cdots, n$ be the following exponential martingale

$$a_j(t) = 1 + \int_0^t a_j(s)h(X_j(s))^T dY_s, \quad t \ge 0, \tag{5.22}$$

also given as

$$a_j(t) = \exp\bigg(\int_0^t h(X_j(s))^T dY_s - \frac{1}{2}\int_0^t |h(X_j(s))|^2 ds\bigg), \quad t \ge 0. \tag{5.23}$$

Here $X_j, j = 1, \cdots, n$ are n mutually independent stochastic processes and independent of Y, each X_j is a solution to the SDE satisfying the signal process (3.1). Hence, the triples $(X_j, a_j, Y), j = 1, \cdots, n$ are identically distributed and have the same distribution as the triple $(X, \tilde{Z}, Y)$ under probability measure $\tilde{P}$. Exercise 8.1.2 in [4] shows that the pairs $(X_j(t), a_j(t)), j = 1, \cdots, n$ are mutually independent conditional upon the observation σ-algebra $\mathcal{Y}_t$, we have for $j = 1, \cdots, n$,

$$\tilde{E}[\phi(X_j(t))a_j(t)|\mathcal{Y}_t] = \tilde{E}[\phi(X_t)\tilde{Z}_t|\mathcal{Y}_t] = \rho_t(\phi). \tag{5.24}$$

Recall the approximation of the Zakai equation ρ_t^n

$$\rho_t^n = \frac{1}{n}\sum_{j=1}^n a_j(t)\delta_{X_j(t)}, \quad t \ge 0. \tag{5.25}$$

Thus

$$\tilde{E}[\rho_t^n(\phi)|\mathcal{Y}_t] = \rho_t(\phi) \tag{5.26}$$

Define random variables $\theta_j^\phi, j = 1, \cdots, n$ as

$$\theta_j^\phi := \phi(X_j(t))a_j(t) - \rho_t(\phi), \quad j = 1, \cdots, n, \quad t < T \tag{5.27}$$

with zero mean and the same distribution as $\phi(X_t)\tilde{Z}_t - \rho_t(\phi)$. Since

$$\begin{aligned} E[\phi(X_j(t))a_j(t) - \rho_t(\phi)] &= E[\phi(X_t)\tilde{Z}_t - \rho_t(\phi)] \\ &= E\Big[E[\phi(X_t)\tilde{Z}_t - \rho_t(\phi)]\big|\mathcal{Y}_t\Big] \\ &= E\Big[E[\phi(X_t)\tilde{Z}_t|\mathcal{Y}_t] - \rho_t(\phi)\Big] \\ &= 0. \end{aligned} \tag{5.28}$$

It then follows that

$$\frac{1}{n}\sum_{j=1}^n \theta_j^\phi = \rho_t^n(\phi) - \rho_t(\phi). \tag{5.29}$$

Since the pairs $(X_i(t), a_i(t))$ and $(X_j(t), a_j(t))$ for $i \neq j$, conditional upon $\mathcal{Y}_t$ are independent, it follows that the random variables $\theta_j^\phi, j = 1, \cdots, n$ are mutually independent conditional upon $\mathcal{Y}_t$. Thus

$$\begin{aligned} \tilde{E}\Big[(\rho_t^n(\phi) - \rho_t(\phi))^2|\mathcal{Y}_t\Big] &= \frac{1}{n^2}\tilde{E}\bigg[\Big(\sum_{j=1}^n \theta_j^\phi\Big)^2\bigg|\mathcal{Y}_t\bigg] \\ &= \frac{1}{n^2}\sum_{j=1}^n \tilde{E}[(\theta_j^\phi)^2|\mathcal{Y}_t] \\ &= \frac{1}{n^2}\sum_{j=1}^n \tilde{E}[(\phi(X_j(t))a_j(t) - \rho_t(\phi))^2|\mathcal{Y}_t] \\ &= \frac{C_\phi(t)}{n}. \end{aligned} \tag{5.30}$$

Taking expectation $\tilde{E}$ on both sides, by Lemma 5.6,

$$\tilde{E}[(\rho_t^n(\phi) - \rho_t(\phi))^2] = \frac{\tilde{E}[C_\phi(t)]}{n} \leq \frac{4\|\phi\|_\infty^2 C_2(t)}{n}, \qquad t \in [0, T]. \tag{5.31}$$

Theorem 5.7 is proved. □

Now we state the convergence result for particle filter method. As we mentioned in the remarks, the difference between particle filter method and tranditional Monte Carlo method is that correction step is introduced after each time interval. Before we prove the theorem, we first define the solution to the dual Zakai equation and give one useful lemma similar to Lemma 5.5.

Recall the Zakai equation (3.15)

$$d\rho(t, x) = \mathrm{L}^*\rho(t, x)dt + h^T\rho(t, x)dY_t \tag{5.32}$$

with initial density

$$\rho(0, x) = P_0. \tag{5.33}$$

To define its dual, we follow E. Pardoux [41] to first introduce the backward Ito integral.
Let $\mathcal{Y}_t^s = \sigma\{Y_r - Y_t, s \le r \le t\}$. $\tilde{Y}_s := Y_s - Y_t$ is a '$\mathcal{Y}_t^s$ backward Wiener process': i.e. $0 \le r \le s$, $\tilde{Y}_r - \tilde{Y}_s$ is a Gaussian distributed operator of covariance $(s-r)I$, independent of $\mathcal{Y}_t^s$. If $\{\varsigma_s, s \in [0,t]\}$ is a process with values in R^m, $\mathcal{Y}_t^s$-adapted continuous path and bounded, we can define the backward Ito integral:

$$\int_s^t \varsigma_r \circ dY_r := P - \lim_{\varepsilon_n \downarrow 0} \sum_{i=1}^n \varsigma_{t_{i+1}} \cdot (Y_{t_{i+1}} - Y_{t_i}), \tag{5.34}$$

where $s = t_0 < t_1 < \cdots < t_n = t$, and $\varepsilon_n = \sup_{k \le n}(t_k - t_{k-1})$.
If ς_s is measurable and $\mathcal{Y}_t^s$-adapted, with

$$E\bigg(\int_0^t |\varsigma_s|^2 ds\bigg) < \infty, \tag{5.35}$$

then

$$\{\int_s^t \varsigma_r \circ dY_r, 0 \le s \le t\} \tag{5.36}$$

is a backward martingale, it is also a backward Ito integral.
Consider a $\mathcal{Y}_t^s$-adapted random variable $v(t,x)$, which satisfies the backward stochastic PDE

$$dv(t,x) + \mathrm{L}v(t,x)dt + hv(t,x) \circ dY_t = 0, \quad 0 \le t \le T \tag{5.37}$$

with final time condition

$$v(T,x) = \phi(x). \tag{5.38}$$

It turns out that (5.37) is the adjoint equation to the Zakai equation. In addition, E. Pardoux [41] proved the following interesting result.

Lemma 5.8. *The process $\{(v(s,\cdot), \rho(s,\cdot)), 0 \le s \le t\}$ is a constant a.s. Here $(\cdot,\cdot)$ denotes the scalar product in $\mathrm{L}^2(\mathbb{R}^{2n})$.*

We say $v(s,x)$ is the dual solution to Zakai equation in the sense of Lemma 5.8. Under assumptions we made at the beginning of section 3, E. Pardoux [41] showed that there exists a unique solution $v(t,x)$ to (5.37).

$$v \in \mathrm{L}^2(\Omega \times (0,t); H^1(\mathbb{R}^{2n})) \cap \mathrm{L}^2(\Omega; C([0,t]; \mathrm{L}^2(\mathbb{R}^{2n}))) \tag{5.39}$$

Using Feynman-Kac formula the solution of (5.37) can be expressed for $s \in [0,t], \phi \in C_b(\mathbb{R}^{2n})$ as

$$v(t, X_s) = \tilde{E}[\phi(X_t) a_s^t(X) | \mathcal{F}_s \vee \mathcal{Y}_t]. \tag{5.40}$$

Here

$$a_s^t(X) = \exp\bigg(\int_s^t h(X_r) dY_r - \frac{1}{2}\int_s^t |h(X_r)|^2 dr\bigg). \tag{5.41}$$

Define the $\mathcal{F}_t$-adapted random variable $\psi^n = \{\psi_t^n, \quad t \ge 0\}$ by

$$\psi_t^n := (\prod_{i=1}^{[t/\varepsilon]} \frac{1}{n} \sum_{j=1}^n a_j^{n,i\varepsilon})(\frac{1}{n} \sum_{j=1}^n a_j^n(t)). \tag{5.42}$$

The following property is important in proving of convergence later and is a improvement of the estimate given by Bain and Crisan [**?**] by allowing h to have nonuniform upper bound.

Lemma 5.9. *For any $t \geq 0$ and $p \geq 2$, if $|h(X_j^n(t))| \leq g(t)$ with*

$$\int_0^t |g(s)|^2 ds < \infty \quad for \quad each \quad t > 0. \tag{5.43}$$

Then there exist two constants $c_1^{t,p}$ and $c_2^{t,p}$ such that

$$\tilde{E}[(a_j^n(t))^p] \leq c_1^{t,p}, \quad j = 1, \cdots, n, \tag{5.44}$$

and

$$\tilde{E}[(\psi_t^n)^p] \leq c_2^{t,p}. \tag{5.45}$$

Proof. Inequality (5.44) can be proved the same as Lemma 5.5, thus we only prove inequality (5.45) and it is obvious from (5.44) that

$$\tilde{E}[(a_j^n(t))^p | \mathcal{F}_{k\varepsilon}] \leq c_1^{t,p}, \quad k = 1, \cdots, n. \tag{5.46}$$

Hence also

$$\tilde{E}[(\frac{1}{n}\sum_{j=1}^n a_j^n(t))^p | \mathcal{F}_{k\varepsilon}] \leq c_1^{t,p}. \tag{5.47}$$

Therefore

$$\begin{aligned} E[(\psi_t^n)^p | \mathcal{F}_{[t/\varepsilon]\varepsilon}] &= (\psi_{[t/\varepsilon]\varepsilon}^n)^p E[(\frac{1}{n}\sum_{j=1}^n a_j^n(t))^p | \mathcal{F}_{[t/\varepsilon]\varepsilon}] \\ &\leq (\psi_{[t/\varepsilon]\varepsilon}^n)^p c_1^{t,p}, \end{aligned} \tag{5.48}$$

since $(\psi_{[t/\varepsilon]\varepsilon}^n)^p$ in an integrable random variable by (5.42) and (5.44).

Also, we have

$$E[(\psi_{k\varepsilon}^n)^p | \mathcal{F}_{(k-1)\varepsilon}] \leq (\psi_{(k-1)\varepsilon}^n)^p c_1^{t,p}. \tag{5.49}$$

Similarly,

$$E[(\psi_{(k-1)\varepsilon}^n)^p | \mathcal{F}_{(k-2)\varepsilon}] \leq (\psi_{(k-2)\varepsilon}^n)^p c_1^{t,p}, \tag{5.50}$$

$$\vdots$$

$$E[(\psi_{\varepsilon}^n)^p | \mathcal{F}_0] \leq (\psi_0^n)^p c_1^{t,p}. \tag{5.51}$$

Hence,

$$E[(\psi_{k\varepsilon}^n)^p] \leq (c_1^{t,p})^k = c_2^{t,p}. \tag{5.52}$$

We proved the conclusion (5.45). □

Let $\rho^n = \{\rho^n_t, \quad t \geq 0\}$ be the measure-valued process defined by

$$
\begin{aligned}
\rho^n_t := \psi^n_t \pi^n_t &= \Big(\prod_{i=1}^{[t/\varepsilon]} \frac{1}{n}\sum_{j=1}^n a_j^{n,i\varepsilon}\Big)\Big(\frac{1}{n}\sum_{j=1}^n a_j^n(t)\Big)\Big(\sum_{j=1}^n \bar{a}_j^n(t)\delta_{X_j^n(t)}\Big)\\
&= \Big(\prod_{i=1}^{[t/\varepsilon]} \frac{1}{n}\sum_{j=1}^n a_j^{n,i\varepsilon}\Big)\Big(\frac{1}{n}\sum_{j=1}^n a_j^n(t)\Big)\Big(\sum_{j=1}^n \frac{a_j^n(t)}{\sum_{k=1}^n a_k^n(t)}\delta_{X_j^n(t)}\Big)\\
&= \Big(\prod_{i=1}^{[t/\varepsilon]} \frac{1}{n}\sum_{j=1}^n a_j^{n,i\varepsilon}\Big)\Big(\frac{1}{n}\sum_{j=1}^n a_j^n(t)\delta_{X_j^n(t)}\Big)\\
&= \frac{\psi^n_{[t/\varepsilon]\varepsilon}}{n}\sum_{j=1}^n a_j^n(t)\delta_{X_j^n(t)}.
\end{aligned}
\tag{5.53}
$$

Here we used definition (4.6) and (5.42). ρ^n_t approximates the solution to the Zakai equation ρ_t and formula (5.53) is the approximation of Kallianpur-Striebel formula in [27]. Before we give the main convergence result, let us mention another property of ρ^n given by A. Bain and D. Crisan [4].

Proposition 5.10. *$\rho^n = \{\rho^n_t, \quad t \geq 0\}$ is a measure-valued process which satisfies the following evolution equation*

$$
\rho^n_t(\phi) = \pi^n_0(\phi) + \int_0^t \rho^n_s(\mathrm{A}\phi)ds + \bar{S}^{n,\phi}_t + \bar{M}^{n,\phi}_{[t/\delta]} + \sum_{k=1}^m \int_0^t \rho^n_s(h_k\phi)dY^k_s \tag{5.54}
$$

for any $\phi \in C^2_b(\mathbb{R}^{2N})$. $\bar{S}^{n,\phi} = \{\bar{S}^{n,\phi}_t, \quad t \geq 0\}$ is an $\mathcal{F}_t$-adapted martingale

$$
\bar{S}^{n,\phi}_t = \frac{1}{n}\sum_{i=0}^\infty \sum_{j=1}^n \int_{i\varepsilon\wedge t}^{(i+1)\varepsilon\wedge t} \psi^n_{i\varepsilon} a^n_j(s)((\nabla\phi)^T\sigma)(X^n_j(s))dW^j_s, \tag{5.55}
$$

and $\bar{M}^{n,\phi} = \{\bar{M}^{n,\phi}_k, \quad k > 0\}$ is the stochastic process

$$
\bar{M}^{n,\phi}_k = \frac{1}{n}\sum_{i=1}^k \psi^n_{i\varepsilon} \sum_{j'=1}^n (\lambda^n_{j'}(i\varepsilon) - n\bar{a}^{n,i\varepsilon}_{j'})\phi(X^n_{j'}(i\varepsilon)), \quad k > 0. \tag{5.56}
$$

Now, we state the main result about $\rho^n_t(\phi)$ which converges to $\rho_t(\phi)$ for any $\phi \in \mathrm{C}_b(\mathbb{R}^{2n})$, which implies that ρ^n_t converges to ρ_t as measure-valued random variables. Our proof extends A. Bain and D. Crisan's result [4] to unbounded observation vector h with the help of Lemma 5.9.

Theorem 5.11. *If the coefficients σ and f are globally Lipschitz and have finite initial data. h satisfies the condition in Lemma 5.5. Then for any $T \geq 0$, there exists a constant c^T_3 independent of n such that for any positive $\phi \in C_b(\mathbb{R}^{2n})$, we have*

$$
\tilde{E}[(\rho^n_t(\phi) - \rho_t(\phi))^2] \leq \frac{c^T_3}{n}\|\phi\|^2_\infty, \qquad t \in [0, T]. \tag{5.57}
$$

In particular, for all $t \geq 0$, ρ^n_t converges in expectation to ρ_t.

Proof. For any $\phi \in \mathrm{C}_b(\mathbb{R}^{2n})$, by Proposition 5.10, we have $\pi_0^n(v(t,X_0)) = \rho_0^n(v(t,X_0))$, use the fact that $(X_j^n(s), a_j^n(s))$ have the same law as $(\mathrm{X}, \tilde{\mathrm{Z}})$ we can show

$$\begin{aligned}
\rho_t(\phi) &= \tilde{E}[\phi(\mathrm{X}_t)\tilde{\mathrm{Z}}_t|\mathcal{Y}_t] \\
&= \tilde{E}[\phi(X_j^n(t))a_j^n(t)|\mathcal{Y}_t] \\
&= \tilde{E}[\tilde{E}[\phi(X_j^n(t))a_j^n(t)|\mathcal{F}_s \vee \mathcal{Y}_t]|\mathcal{Y}_t] \\
&= \tilde{E}[v(t,X_s)a_j^n(s)|\mathcal{Y}_t] \\
&= \tilde{E}[\tilde{\mathrm{Z}}_s v(t,X_s)|\mathcal{Y}_t] \\
&= \tilde{E}[\tilde{\mathrm{Z}}_s v(t,X_s)|\mathcal{Y}_s] \\
&= \rho_s(v(t,X_s)).
\end{aligned} \tag{5.58}$$

for any $s \in [0,t]$, so that $\pi_0(v(t,X_0)) = \rho_0(v(t,X_0)) = \rho_t(\phi)$.
Thus $\rho_t^n(\phi) - \rho_t(\phi)$ can be split as

$$\begin{aligned}
\rho_t^n(\phi) - \rho_t(\phi) = (\rho_t^n(\phi) - \rho_{[t/\varepsilon]\varepsilon}^n(v(t,X_{[t/\varepsilon]\varepsilon}))) + \sum_{k=1}^{[t/\varepsilon]}(\rho_{k\varepsilon}^n(v(t,X_{k\varepsilon})) - \rho_{k\varepsilon-}^n(v(t,X_{k\varepsilon-}))) \\
+ \sum_{k=1}^{[t/\varepsilon]}(\rho_{k\varepsilon-}^n(v(t,X_{k\varepsilon-})) - \rho_{(k-1)\varepsilon}^n(v(t,X_{(k-1)\varepsilon}))) \\
+ (\pi_0^n(v(t,X_0)) - \pi_0(v(t,X_0))).
\end{aligned} \tag{5.59}$$

Here $X_{k\varepsilon-}$ is the position of particles right before the branching time $k\varepsilon$.
We must bound each term on the right hand side individually. We first derive the following relation from (5.40),

$$\begin{aligned}
&\rho_{[t/\varepsilon]\varepsilon}^n(v(t,X_{[t/\varepsilon]\varepsilon})) \\
&= \frac{\psi_{[t/\varepsilon]\varepsilon}^n}{n}\sum_{j=1}^n a_j^n([t/\varepsilon]\varepsilon)v(t,X_j^n([t/\varepsilon]\varepsilon)) \\
&= \frac{\psi_{[t/\varepsilon]\varepsilon}^n}{n}\sum_{j=1}^n a_j^n([t/\varepsilon]\varepsilon)\tilde{E}[\phi(X_j^n(t))a_{[t/\varepsilon]\varepsilon}^t(X)|\mathcal{F}_{[t/\varepsilon]\varepsilon} \wedge \mathcal{Y}_t] \\
&= \frac{\psi_{[t/\varepsilon]\varepsilon}^n}{n}\sum_{j=1}^n \tilde{E}[\phi(X_j^n(t))a_j^n(t)|\mathcal{F}_{[t/\varepsilon]\varepsilon} \vee \mathcal{Y}_t] \\
&= \tilde{E}[\rho_t^n(\phi)|\mathcal{F}_{[t/\varepsilon]\varepsilon} \vee \mathcal{Y}_t].
\end{aligned} \tag{5.60}$$

For the first term, using the fact that random variables $X_j^n(t)$ for $j = 1, 2, \cdots, n$ are mutually independent conditional upon $\mathcal{F}_{[t/\varepsilon]\varepsilon} \vee \mathcal{Y}_t$, because the generating

Brownian motions $W^{(j)}$, for $j=1,2,\cdots,n$ are mutually independent. We have

$$\begin{aligned}
&\tilde{E}[(\rho_t^n(\phi)-\rho^n_{[t/\varepsilon]\varepsilon}(v(t,X_{[t/\varepsilon]\varepsilon})))^2|\mathcal{F}_{[t/\varepsilon]\varepsilon}\vee\mathcal{Y}_t]\\
&=\tilde{E}[(\rho_t^n(\phi)-\tilde{E}[\rho_t^n(\phi)|\mathcal{F}_{[t/\varepsilon]\varepsilon}\vee\mathcal{Y}_t])^2|\mathcal{F}_{[t/\varepsilon]\varepsilon}\vee\mathcal{Y}_t]\\
&=\frac{(\psi^n_{[t/\varepsilon]\varepsilon})^2}{n^2}\tilde{E}[(\sum_{j=1}^n\phi(X_j^n(t))a_j^n(t))^2|\mathcal{F}_{[t/\varepsilon]\varepsilon}\vee\mathcal{Y}_t]\\
&-\frac{(\psi^n_{[t/\varepsilon]\varepsilon})^2}{n^2}(\sum_{j=1}^n\tilde{E}[\phi(X_j^n(t))a_j^n(t)|\mathcal{F}_{[t/\varepsilon]\varepsilon}\vee\mathcal{Y}_t])^2\\
&\le\frac{(\psi^n_{[t/\varepsilon]\varepsilon})^2}{n^2}\|\phi\|_\infty^2\sum_{j=1}^n\tilde{E}[a_j^n(t)^2|\mathcal{F}_{[t/\varepsilon]\varepsilon}\vee\mathcal{Y}_t].
\end{aligned}\tag{5.61}$$

Taking expectation on both sides of (5.61), then using Cauchy-Schwartz inequality and Lemma 5.9 for $p=4$, we obtain

$$\begin{aligned}
\tilde{E}[(\rho_t^n(\phi)-\rho^n_{[t/\varepsilon]\varepsilon}(v(t,X_{[t/\varepsilon]\varepsilon})))^2]&\le\frac{\|\phi\|_\infty^2}{n^2}\sum_{j=1}^n\left(\tilde{E}[(\psi^n_{[t/\varepsilon]\varepsilon})^4]\right)^{1/2}\left(\tilde{E}[a_j^n(t)^4]\right)^{1/2}\\
&\le\frac{\sqrt{c_1^{t,4}c_2^{t,4}}}{n}\|\phi\|_\infty^2.
\end{aligned}\tag{5.62}$$

Similarly,

$$\begin{aligned}
&\tilde{E}[(\rho^n_{k\varepsilon-}(v(t,X_{k\varepsilon-}))-\rho^n_{(k-1)\varepsilon}(v(t,X_{(k-1)\varepsilon})))^2]\\
&\qquad\le\frac{1}{n^2}\sum_{j'=1}^n\tilde{E}[(\psi^n_{(k-1)\varepsilon}a_{j'}^{n,k\varepsilon})^2v(t,X_{j'}^n(k\varepsilon))^2].
\end{aligned}\tag{5.63}$$

From (5.40) we deduce that

$$v(t,X_{j'}^n(k\varepsilon))=\tilde{E}[\phi(X_j^n(t))a^t_{k\varepsilon}(X_j^n)|\mathcal{F}_{k\varepsilon}\vee\mathcal{Y}_t].\tag{5.64}$$

Hence by Jensen's inequality,

$$\begin{aligned}
\tilde{E}[(v(t,X_{j'}^n(k\varepsilon)))^p]&\le\tilde{E}[\tilde{E}[\phi(X_j^n(t))a^t_{k\varepsilon}(X_j^n)|\mathcal{F}_{k\varepsilon}\vee\mathcal{Y}_t]^p]\\
&=\tilde{E}[(\phi(X_j^n(t))a^t_{k\varepsilon}(X_j^n))^p].
\end{aligned}\tag{5.65}$$

Using $p=4,8$ in lemma 5.9 and Cauchy-Schwarz inequality twice, (5.63) becomes,

$$\begin{aligned}
&\tilde{E}[(\rho^n_{k\varepsilon-}(v(t,X_{k\varepsilon-}))-\rho^n_{(k-1)\varepsilon}(v(t,X_{(k-1)\varepsilon})))^2]\\
&\qquad\le(c_1^{t,4})^{1/2}(c_1^{t,8})^{1/4}(c_2^{t,8})^{1/4}\frac{\|\phi\|_\infty^2}{n}.
\end{aligned}\tag{5.66}$$

For the second term on the right hand side of (5.59), observe that

$$\begin{aligned}
&\tilde{E}[(\rho^n_{k\varepsilon}(v(t,X_{k\varepsilon}))-\rho^n_{k\varepsilon-}(v(t,X_{k\varepsilon-})))^2|\mathcal{F}_{k\varepsilon-}\vee\mathcal{Y}_t]\\
&=\frac{\psi_{k\varepsilon}^2}{n^2}\sum_{j',l'=1}^n\tilde{E}[(\lambda_{j'}^{n,k\varepsilon}-n\bar{a}_{j'}^{n,k\varepsilon})(\lambda_{l'}^{n,k\varepsilon}-n\bar{a}_{l'}^{n,k\varepsilon})|\mathcal{F}_{k\varepsilon-}\vee\mathcal{Y}_t]\\
&\times v(t,X_{j'}^n(k\varepsilon))v(t,X_{l'}^n(k\varepsilon)).
\end{aligned}\tag{5.67}$$

Recall one of the properties of random variables $\{\lambda_{j'}^{n,k\varepsilon},\quad j'=1,\cdots,n\}$ saying that they are non-positively correlated, it follows that

$$\begin{aligned}
&\tilde{E}[(\rho_{k\varepsilon}^n(v(t,X_{k\varepsilon}))-\rho_{k\varepsilon-}^n(v(t,X_{k\varepsilon-})))^2|\mathcal{F}_{k\varepsilon-}\vee\mathcal{Y}_t] \\
&\qquad\le\frac{\psi_{k\varepsilon}^2}{n^2}\sum_{j'=1}^n\tilde{E}[(\lambda_{j'}^{n,k\varepsilon}-n\bar{a}_{j'}^{n,k\varepsilon})^2|\mathcal{F}_{k\varepsilon-}\vee\mathcal{Y}_t](v(t,X_{j'}^n(k\varepsilon)))^2 \\
&\qquad\le\frac{\psi_{k\varepsilon}^2}{n^2}\sum_{j'=1}^n\{n\bar{a}_{j'}^{n,k\varepsilon}\}(1-\{n\bar{a}_{j'}^{n,k\varepsilon}\})(v(t,X_{j'}^n(k\varepsilon)))^2.
\end{aligned}\tag{5.68}$$

Finally using Young's inequality $q(1-q)\le\frac{1}{4}$ for $q=\{n\bar{a}_{j'}^{n,k\varepsilon}\}$, (5.64) and Lemma 5.9 with $p=4$, it follows that

$$\tilde{E}[(\rho_{k\varepsilon}^n(v(t,X_{k\varepsilon}))-\rho_{k\varepsilon-}^n(v(t,X_{k\varepsilon-})))^2]\le\frac{1}{4n}\sqrt{c_1^{t,4}c_2^{t,4}}\|\phi\|_\infty^2.\tag{5.69}$$

For the last term of (5.59), note that $v(t,X_0)$ is $\mathcal{Y}_t$-measurable, therefore using the mutual independence of the initial points $X_j^n(0)$, and the fact that

$$\tilde{E}[v(t,X_j^n(0))|\mathcal{Y}_t]=\pi_0(v(t,X_0)),\tag{5.70}$$

we obtain

$$\begin{aligned}
\tilde{E}[(\pi_0^n(v(t,X_0))-\pi_0(v(t,X_0)))^2|\mathcal{Y}_t]&=\frac{1}{n^2}\sum_{j=1}^n\tilde{E}[(v(t,X_j^n(0)))^2|\mathcal{Y}_t]-(\pi_0(v(t,X_0)))^2 \\
&\le\frac{1}{n^2}\sum_{j=1}^n\tilde{E}[(v(t,X_j^n(0)))^2|\mathcal{Y}_t].
\end{aligned}\tag{5.71}$$

Hence using (5.64) and Lemma 5.9 with $p=4$,

$$\begin{aligned}
\tilde{E}[(\pi_0^n(v(t,X_0))-\pi_0(v(t,X_0)))^2]&\le\frac{1}{n^2}\sum_{j=1}^n\tilde{E}[v(t,X_j^n(0))^2] \\
&\le\frac{1}{n}\sqrt{c_1^{t,4}c_2^{t,4}}\|\phi\|_\infty^2.
\end{aligned}\tag{5.72}$$

We get the conclusion by substituting all above individual estimates back to (5.59). □

References

1. Amirdjanova, A. and Kallianpur, G.: Stochastic vorticity and associated filtering theory, *Appl. Math. Optim.* **46** (2002) 89–96.
2. Anderson, C. and Greengard, C.: On vortex methods, *SIAM J. Numer. Anal.* **22** (1985) 413–440.
3. Aref, H.: Point vortex dynamics: a classical mathematics playground, *J. Math. Phys.* **48** (2007) 1–23.
4. Bain, A. and Crisan, D.: *Fundamentals of Stochastic Filtering*, Springer, New York, 2008.
5. Baras, J. S., Blankership, G. L. and Hopkins, W. E.: Existence, uniqueness, and asymptotic behavior of solutions to a class of Zakai equations with unbounded coefficients, *IEEE Trans. Automat. Control*, **AC-28** (1983) 203–214.

6. Besala, P.: Fundamental solution and Cauchy problem for a parabolic system with unbounded coefficients, *J. Differential Equations* **33** (1979) 26–38.
7. Budhiraja, A., Chen, L. and Lee, C.: A survey of numerical methods for nonlinear filtering problems, *Phys. D* **230, 1-2** (2007) 27–36.
8. Chorin, A. J.: Numerical study of slightly viscous flow, *J. Fluid Mech.* **57** (1973) 785–796.
9. Chorin, A. J.: *Vorticity and Turbulence*, Springer-Verlag, Berlin and Heidelberg, 1994.
10. Cottet, G. H. and Koumoutsakos, P. D.: *Vortex Methods: Theory and Practice*, Second ed., Cambridge University Press, New York, 2000.
11. Crisan, D.: Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation, *Ann. Probab.* **32(2)** (2003) 693–718.
12. Crisan, D., Del Moral, P. and Lyons, T.: Discrete filtering using branching and interacting particle systems, *Markov Process. Related Fields* **5** (1999) 299–318.
13. Crisan, D., Del Moral, P. and Lyons, T.: Interacting particle systems approximations of the Kushner-Stratonovitch equation, *Adv. in Appl. Probab.*, **31(3)** (1999) 819–838.
14. Crisan, D. and Lyons, T.: Nonlinear filtering and measure valued processes, *Probab. Theory Related Fields* **109** (1997) 217–244.
15. Crisan, D. and Lyons, T.: A particle approximation of the solution of the Kushner-Stratonovitch equation, *Probab. Theory Related Fields* **115(4)** (1999) 549–578.
16. Doucet, A., De Freitas, N. and Gordon, N.: *Sequential Monte Carlo Methods in Practice*, Springer, New York, 2001.
17. Eidelman, S. D.: *Parabolic Systems*, Wolters-Noordhoff, Groningen, 1969.
18. Ferrario, B.: Absolute continuity of laws for semilinear stochastic equations with additive noise, *Commun. Stoch. Anal.* **2(2)** (2008) 209–227.
19. Fujisaki, M., Kallianpur, G. and Kunita, H.: Stochastic differential equations for the nonlinear filtering problem, *Osaka J. Math.* **9** (1972) 19–40.
20. Gihman, I. I. and Skorokhod, A. V.: *Stochastic Differential equations*, Springer-Verlag, New York, 1972.
21. Gyöngy, I. and Krylov, N. V.: Stochastic partial differential equations with unbounded coefficients and applications I, *Stochastics* **32** (1990) 53–91.
22. Gyöngy, I. and Krylov, N. V.: Stochastic partial differential equations with unbounded coefficients and applications II, *Stochastics* **32** (1990) 165–180.
23. Gyöngy, I. and Krylov, N. V.: Stochastic partial differential equations with unbounded coefficients and applications III, *Stochastics* **40** (1992) 77–115.
24. Hobbs, S. and Sritharan, S. S.: Nonlinear filtering of stochastic reacting and diffusing systems, *Probability and Modern Analysis, N. Gretsky, J. Goldstein and J.J. Uhl, eds.*, Marcel Dekker, 1996.
25. Ikeda, N. and Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, Elsevier Science Ltd., Amsterdam, 1981.
26. Ito, K.: Approximation of the Zakai equation for nonlinear filtering, *SIAM J. Control and Optim.* **34(2)** (1996) 620–634.
27. Kallianpur, G.: *Stochastic Filtering Theory*, Springer-Verlag, New York, 1980.
28. Kallianpur, G. and Karandikar, R. L.: The nonlinear filtering problem for the unbounded case, *Stochastics Process. Appl.* **18** (1984) 57–66.
29. Kotelenez, P.: A stochastic Navier-Stokes equation for the vorticity of a two-dimensional fluid, *Ann. Appl. Probab.*, **5(4)** (1995) 1126–1160.
30. Krasny, R.: Desingularization of periodic vortex sheet roll-up, *J. Comput. Phys.* **65** (1986) 292–313.
31. Krasny, R.: A study of singularity formation in a vortex sheet by the point vortex approximation, *J. Fluid Mech.* **167** (1986) 65–93.
32. Krylov, N. V. and Rozovskii, B. L.: Characteristics of degenerating second-order parabolic Ito equations, *J. Math. Sci.* **32(4)** (1986) 336–348.
33. Kushner, H. J.: *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic Press, New York, 1977.

34. Le Grand, F., Musso, C. and Oudjane, N.: An analysis of regularized interacting particle methods for nonlinear filtering, in: *Proceedings of the IEEE European Workshop on Computer-intersive Methods in Control and Data Processing* (1998) 167–174, Prague.
35. Marchioro, C. and Pulvirenti, M.: Hydrodynamics in two dimensions and vortex theory, *Comm. Math. Phys.* **84** (1982) 483–503.
36. McKean, H. P.: *Stochastic Integrals*, American Mathematical Society, New York, 2005.
37. Milstein, G. N. and Tretyakov, M. V.: Monte Carlo methods for backward equations in nonlinear filtering, *Adv. Appl. Probab.* **41(1)** (2009) 63–100.
38. Oksendal, B. K.: *Stochastic Differential Equations: An Introduction with Applications*, Fifth ed., Springer, New York, 2002.
39. Pardoux, E.: Stochastic partial differential equations and filtering of diffusion processes, *Stochastics* **3** (1979) 127–167.
40. Pardoux, E. Nonlinear filtering, prediction and smoothing, *Stochastic Systems: The Mathematics of Filtering and Identification and Applications, M. Hazewinkel and J. C. Willems, eds.* (1980) 529–557, Reidel, Dordrecht.
41. Pardoux, E.: Équations du filtrage non linéaire, de la prédiction et du lissage, *Stochastics* **6** (1982) 193–231.
42. Popa, S. and Sritharan, S. S.: Nonlinear filtering of Ito-Lévy stochastic differential equations with continuous observations, *Accepted for publication in Commun. Stoch. Anal.* (2009) to appear.
43. Protter, P. E.: *Stochastic Integration and Differential Equations*, Second ed., Springer-Verlag, Berlin Heidelberg New York, 2005.
44. Raviart, P. A.: An analysis of particle methods, *Lecture Notes in Math.* **1127** (1985) 243–324.
45. Rosenhead, L.: The formation of vortices from a surface of discontinuity, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **134** (1931) 170–192.
46. Rozovskii, B. L.: *Stochastic Evolution Systems*, Kluwer, Dordrecht Boston London, 1990.
47. Rozovskii, B. L.: *Lecture notes on linear stochastic partial differential equations*, Technical Report 25, Department of Mathematics, The University of North Carolina at Charlotte, 1990.
48. Rozovskii, B. L.: A simple proof of uniqueness for Kushner and Zakai equation, *Stochastic Analysis, E. Mayer-Wolf, E. Merzbach and A. Schwartz, eds.* (1991) 449–458, Academic Press, New York.
49. Sritharan, S. S.: Nonlinear filtering of stochastic Navier-Stokes equations, *Nonlinear Stochastic PDEs: Burgers Turbulence and Hydrodynamic Limit, T. Funaki and W. A. Woyczynski, eds.* (1995) 247–260, Springer-Verlag, New York.
50. Sritharan, S. S.: *Optimal Control of Viscous Flow*, SIAM, Philadelphia, PA, 1998.
51. Sritharan, S. S.: Stochastic Navier-Stokes equations: solvability, control and filtering, *Stochastic Partial Differential Equations and Applications-VII, G. Da Prato and L. Tubaro, eds.* (2006) 273-280, Chapman and Hall/CRC, Boca Raton, FL.
52. Zakai, M.: On the optimal filtering of diffusion processes, *Probab. Theory and Related Fields* **11** (1969) 230–243.

Sivagruru S. Sritharan: The Graduate School of Engineering and Applied Sciences, Naval Postgraduate School, Monterey, California, 93943, United States
E-mail address: `sssritha@nps.edu`

Meng Xu: Department of Mathematics, University of Wyoming, Laramie, Wyoming, 82071, United States
E-mail address: `mxu@uwyo.edu`