PARTICLE APPROXIMATIONS TO THE FILTERING PROBLEM IN CONTINUOUS TIME

JIE XIONG

Abstract. In this chapter, we survey some recent results on the particle system approximations to stochastic filtering problems in continuous time. First, a weighted particle system representation of the optimal filter is given and a numerical scheme based on this representation is proposed. Its convergence to the optimal filter, together with the rate of convergence is proved. Secondly, to reduce the estimate error due to the exponential growth of the variance for individual weights, a branching weighted particle system is defined and an approximate filter based on this particle system is proposed. Its approximate optimality is proved and the rate of convergence is characterized by a central limit type theorem. Thirdly, as an alternative approach in reducing the estimate error, an interacting particle system (with neither branching nor weights) to direct the particles toward more likely regions is proposed. A convergence result for this system is established. Finally, we use weighted branching particle systems to approximate the optimal filter for the model with point process observations.

1. Introduction

There are two related stochastic processes in each filtering problem: The signal process which we want to estimate and the observation process which provides the information we can use. In this chapter, we assume that the signal process is a d-dimensional diffusion governed by the following stochastic differential equation (SDE):

$$dX_t = b(X_t)dt + c(X_t)dW_t + \sigma(X_t)dB_t, \qquad (1.1)$$

where B and W are independent Brownian motions of dimensions d and m, respectively, and $b:\ \mathbb{R}^d \to \mathbb{R}^d$, $c:\ \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $\sigma:\ \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are continuous mappings. To ensure the existence and uniqueness for the solution to (1.1) and for the convenience of the estimates, the following condition (BC1) will be assumed throughout this chapter.

Condition (BC1): The mappings b, c, σ are bounded and Lipschitz continuous.

For the observation process, we consider two models: The classical one and the one with point processes as its observations. In the classical model, the observation

---

2000 *Mathematics Subject Classification.* Primary: 60H15; Secondary: 60K35, 35R60, 93E11.

Key words and phrases. Stochastic partial differential equation, particle system approximation, Zakai equation, nonlinear filtering.

Research of Xiong is supported partially by NSA.

process is an m-dimensional stochastic process given by

$$Y_t = \int_0^t h(X_s)ds + W_t, \tag{1.2}$$

where $h:\ \mathbb{R}^d \to \mathbb{R}^m$ is a continuous mapping. For the point process model, the observation process is an m-dimensional process given by

$$Y_t^i = N_i\left(\int_0^t \lambda_i(X_s, s)ds\right), \qquad i = 1, 2, \cdots, m, \tag{1.3}$$

where $N_1,\ N_2,\ \cdots,\ N_m$ are independent unit Poisson processes which are independent of X, and $\lambda_i:\ \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$, $i = 1, 2, \cdots, m$, are measurable mappings.

The optimal filter π_t is a $\mathcal{P}(\mathbb{R}^d)$-valued process given by

$$\langle \pi_t, f\rangle \equiv \mathbb{E}\left(f(X_t)\Big|\mathcal{G}_t\right), \qquad \forall f \in C_b(\mathbb{R}^d),$$

where $\mathcal{G}_t = \mathcal{F}_t^Y$ is the σ-field generated by $Y_s,\ s \le t$, $\langle \mu, f\rangle$ stands for the integral of f with respect to the measure μ, and $\mathcal{P}(\mathbb{R}^d)$ is the collection of all Borel probability measures on $\mathbb{R}^d$. Stochastic differential equations on $\mathcal{P}(\mathbb{R}^d)$, called the filtering equations, satisfied by π_t for these two models will be derived in Sections 2 and 5, respectively.

Explicit solutions to the filtering equations are rarely available. Thus, to solve the filtering problems, we have to resort to numerical schemes. Particle approximations is a class of the effective numerical schemes. The main idea is to represent the solution to a stochastic partial differential equation (SPDE) through a system of weighted particles whose locations and weights are governed by SDEs which can be solved numerically. This numerical scheme based on the weighted particle system, regarded as a direct Monte-Carlo method, will be introduced in Section 2.

As the error in the Monte-Carlo approximation increases exponentially fast when the time parameter tends to infinity due to the exponential growth of the variances of the weights of the particles in the system, we need to modify the weight of each particle. However, the total mass has to be kept constant for the approximate filter to take values in the space of probability measures. To this end, the number of particles in the system will be changed from time to time. We use a branching particle system to match the change of the number of particles in the system. This numerical scheme based on branching weighted particle systems, called the hybrid filter, will be studied in Section 3.

Another method in reducing the error is to use interacting particle systems, namely, there is no branching in the system while the motions of the particles are directed to the region where the optimal filter has a higher density. This interacting particle system will be studied in Section 4.

Finally, we consider the filtering model with point processes as its observations. This model arises from the study of ultra-high frequency date in mathematical finance. A branching particle system will be utilized to approximate the optimal filter in this setup. The difference between this filtering model and the classical one will be presented in Section 5.

2. Filtering using weighted particles

In this section, we establish a weighted particle system representation for the optimal filter of the classical model. Based on this representation, a numerical scheme will be proposed and its convergence to the optimal filter will be proved, together with the rate of convergence derived. The material of this section is taken from the papers of Kurtz and Xiong [26], [27], [28].

To derive the filtering equation for π_t, it is convenient to apply Girsanov's formula to transform the probability measure to a new one such that Y_t becomes an m-dimensional Brownian motion, independent of the process B_t, under the new probability measure. To this end and for the convenience of the estimates later, we assume the following condition (BC2) on h.

Condition (BC2): The mapping h is bounded and Lipschitz continuous.

Let $\hat{P} \sim P$ be the probability measure given by

$$\left.\frac{dP}{d\hat{P}}\right|_{\mathcal{F}_t} = M_t \equiv \exp\left(\int_0^t h(X_s)^* dY_s - \frac{1}{2}\int_0^t |h(X_s)|^2 ds\right),$$

where v^* stands for the transpose of a vector (or matrix) v. Then, under $\hat{P}$, (B_t, Y_t) is a $(d+m)$-dimensional Brownian motion, and the signal process X_t is governed by

$$dX_t = (b - ch)(X_t)dt + c(X_t)dY_t + \sigma(X_t)dB_t. \tag{2.1}$$

The conditional expectation with respect to P, as appeared in the definition of the optimal filter, can be represented according to the conditional expectation with respect to $\hat{P}$ by Bayes' formula which is called the Kallianpur-Striebel formula (cf. Kallianpur-Striebel [23], [24]) in the filtering setup. The advantage of using $\hat{P}$ is that the signal is a functional of B and Y, while B is independent of $\mathcal{G}_t$ and Y is measurable with respect to $\mathcal{G}_t$, and the conditional expectations in both cases (individually) are easy to find.

Let $C_b(\mathbb{R}^d)$ be the collection of all bounded continuous real-valued functions.

Theorem 2.1 (Kallianpur-Striebel formula)**.** *The optimal filter π_t can be represented as*

$$\langle \pi_t, f\rangle = \frac{\langle V_t, f\rangle}{\langle V_t, 1\rangle}, \qquad \forall\ f \in C_b(\mathbb{R}^d), \tag{2.2}$$

where

$$\langle V_t, f\rangle = \hat{\mathbb{E}}(M_t f(X_t)|\mathcal{G}_t) \tag{2.3}$$

and $\hat{\mathbb{E}}$ refers to the expectation with respect to the measure $\hat{P}$.

On the probability space $(\Omega, \mathcal{F}, \hat{P})$, let B^i, $i = 1, 2, \cdots$, be independent copies of B, and let them be independent of Y. Now we consider an interacting particle system: For $i = 1, 2, \cdots$,

$$dX_t^i = (b - ch)(X_t^i)dt + c(X_t^i)dY_t + \sigma(X_t^i)dB_t^i \tag{2.4}$$

and

$$dM_t^i = M_t^i h^*(X_t^i)dY_t, \qquad M_0^i = 1 \tag{2.5}$$

By (2.3) and the conditional (given $\mathcal{G}_t$) law of large numbers, we get

Theorem 2.2. *Suppose that* $\{X_0^i,\ i=1,2,\cdots\}$ *are i.i.d. random vectors with common distribution* π_0 *on* $\mathbb{R}^d$*. Then*

$$\langle V_t, f\rangle = \lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k} M_t^i f(X_t^i), \tag{2.6}$$

where $\{(M^i, X^i):\ i=1,2,\cdots\}$ *is the unique strong solution to the particle system (2.4)-(2.5).*

Applying Itô's formula to (2.4)-(2.6), we get the Zakai equation for the unnormalized filter V_t.

Theorem 2.3 (Zakai's equation)**.** *The unnormalized filter* V_t *satisfies the following stochastic differential equation:*

$$\langle V_t, f\rangle = \langle V_0, f\rangle + \int_0^t \langle V_s, Lf\rangle\, ds + \int_0^t \langle V_s, \nabla^* f c + f h^*\rangle\, dY_s, \tag{2.7}$$

where

$$Lf = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_{ij}^2 f + \sum_{i=1}^{d} b_i\partial_i f$$

is the generator of the signal process, and the $d\times d$ *matrix* $a=(a_{ij})$ *is given by* $a = cc^* + \sigma\sigma^*$*.*

Then, applying Itô's formula to (2.2) and (2.7), we get the following filtering equation.

Theorem 2.4 (Kushner-FKK equation)**.** *The optimal filter* π_t *satisfies the following stochastic differential equation: For all* $f\in C_b^2(\mathbb{R}^d)$,

$$\begin{aligned}\langle \pi_t, f\rangle &= \langle \pi_0, f\rangle + \int_0^t \langle \pi_s, Lf\rangle\, ds \qquad (2.8)\\ &\quad + \int_0^t \left(\langle \pi_s, \nabla^* f c + f h^*\rangle - \langle \pi_s, f\rangle\langle \pi_s, h^*\rangle\right) d\nu_s,\end{aligned}$$

where the innovation process ν_t*, given by*

$$d\nu_t = dY_t - \langle \pi_t, h\rangle\, dt, \tag{2.9}$$

is an m*-dimensional* $\mathcal{G}_t$*-Brownian motion under the original probability measure.*

To propose a numerical approximation to the unnormalized filter, we apply Euler scheme to approximate the solution of a finite system of n particles. For $\delta>0$, let

$$\eta_\delta(t) = j\delta \qquad \text{for } j\delta\le t<(j+1)\delta.$$

Note that M_t^i is positive. To keep this positivity property in the approximation, we consider the Euler scheme of the process $Z_t^i\equiv\log M_t^i$ which satisfies the following equation

$$dZ_t^i = h(X_t^i)^* dY_t - \frac{1}{2}|h(X_t^i)|^2 dt.$$

Now we define the finite system $\{(X^{\delta,i}, Z^{\delta,i}, V^{n,\delta}):\ i=1,2,\cdots,n\}$ as follows:

$$\begin{cases} dX_t^{\delta,i} = (b-ch)(X^{\delta,i}_{\eta_\delta(t)})dt + c(X^{\delta,i}_{\eta_\delta(t)})dY_t + \sigma(X^{\delta,i}_{\eta_\delta(t)})dB_t^i \\ dZ_t^{\delta,i} = h\left(X^{\delta,i}_{\eta_\delta(t)}\right)^* dY_t - \frac{1}{2}\left|h\left(X^{\delta,i}_{\eta_\delta(t)}\right)\right|^2 dt \\ V_t^{n,\delta} = \frac{1}{n}\sum_{i=1}^n \exp\left(Z_t^{\delta,i}\right)\delta_{X_t^{\delta,i}}. \end{cases} \tag{2.10}$$

To prove the convergence of $V_t^{n,\delta}$ to V_t, we need a metric on the space $\mathcal{M}_F(\mathbb{R}^d)$ of finite Borel measures on $\mathbb{R}^d$. We shall use Wasserstein's metric.

For $\nu_1, \nu_2 \in \mathcal{M}_F(\mathbb{R}^d)$, the Wasserstein metric is given by

$$\rho(\nu_1,\nu_2) = \sup\{|\langle \nu_1,\phi\rangle - \langle \nu_2,\phi\rangle| : \phi \in \mathbb{B}_1\},$$

where

$$\mathbb{B}_1 = \left\{\phi:\ \mathbb{R}^d \to \mathbb{R};\ |\phi(x)-\phi(y)| \le |x-y|, |\phi(x)| \le 1, \forall x,y \in \mathbb{R}^d\right\}.$$

Under this metric, $\mathcal{M}_F(\mathbb{R}^d)$ becomes a Polish space.

We need the following

Condition (I): The initial positions $\{x_i^n:\ i=1,2,\cdots,n\}$ of the particles are i.i.d. random vectors in $\mathbb{R}^d$ with the common distribution $\pi_0 \in \mathcal{P}(\mathbb{R}^d)$ which satisfies $\int_{\mathbb{R}^d}|x|^2\pi_0(dx) < \infty$.

Theorem 2.5. *Let $\bar{V}_t^n = V_t^{n,1/n}$. Assume Conditions (I), (BC1) and (BC2) hold. Then there exists a constant $K_1(t)$ such that*

$$\hat{\mathbb{E}}\rho(\bar{V}_t^n, V_t) \le \frac{K_1(t)}{\sqrt{n}}.$$

To show that $\frac{1}{\sqrt{n}}$ is indeed the order for the rate of convergence, we define

$$S_t^n = \sqrt{n}\left(\bar{V}_t^n - V_t\right),$$

and prove the tightness for $\{S^n\}$ in an appropriate space. For simplicity of notation, we restrict our calculations to space dimensions $d=m=1$ in the rest of this section.

As in Hitsuda and Mitoma [22], we use a modified Schwartz space Φ. Let $\rho(x) = C\exp\left(-1/(1-|x|^2)\right)1_{|x|<1}$, where C is a constant such that $\int \rho(x)dx = 1$. Let

$$\psi(x) = \int e^{-|y|}\rho(x-y)dy.$$

Let

$$\Phi = \{\phi : \phi\psi \in \mathcal{S}\},$$

where $\mathcal{S}$ is the Schwartz space. For $\kappa = 0,1,2,\ldots$, define

$$\|\phi\|_\kappa^2 = \sum_{0\le k\le\kappa}\int_{\mathbb{R}}(1+|x|^2)^{2\kappa}\left|\frac{d^k}{dx^k}(\phi(x)\psi(x))\right|^2 dx.$$

Let Φ_κ be the completion of Φ with respect to $\|\cdot\|_\kappa$. Then Φ_κ is a Hilbert space with inner product

$$\langle\phi_1,\phi_2\rangle_\kappa = \sum_{0\le k\le\kappa}\int_{\mathbb{R}}(1+|x|^2)^{2\kappa}\left(\frac{d^k}{dx^k}(\phi_1(x)\psi(x))\right)\left(\frac{d^k}{dx^k}(\phi_2(x)\psi(x))\right)dx.$$

Note that $\Phi_\kappa \supset \Phi_{\kappa+1}$ and that Φ_0 is $L^2(\mu_\psi)$, where $\mu_\psi(dx) = \psi^2(x)dx$. For $\hat{\phi} \in \Phi_0$ and $\phi \in \Phi_\kappa$,

$$\langle \hat{\phi}, \phi \rangle \equiv \langle \hat{\phi}, \phi \rangle_0 = \int_{\mathbb{R}} \hat{\phi}(x)\phi(x)\psi^2(x)dx$$

defines a continuous linear functional on Φ_κ with norm

$$\|\hat{\phi}\|_{-\kappa} = \sup_{\phi \in \Phi_\kappa} \frac{|\langle \hat{\phi}, \phi \rangle|}{\|\phi\|_\kappa},$$

and we let $\Phi_{-\kappa}$ denote the completion of Φ_0 with respect to this norm. Then $\Phi_{-\kappa}$ is a representation of the dual of Φ_κ. If $\{\phi_j^\kappa\}$ is a complete, orthonormal system for Φ_κ, then the inner product for $\Phi_{-\kappa}$ can be written as

$$\langle \hat{\phi}_1, \hat{\phi}_2 \rangle_{-\kappa} = \sum_{j=1}^{\infty} \langle \hat{\phi}_1, \phi_j^\kappa \rangle \langle \hat{\phi}_2, \phi_j^\kappa \rangle. \tag{2.11}$$

By a slight modification of Theorem 7, page 82, of Gel'fand and Vilenkin [19], these norms determine a nuclear space, so in particular, for each κ there exists a $\kappa' > \kappa$ such that the embedding $T_\kappa^{\kappa'} : \Phi_{\kappa'} \to \Phi_\kappa$ is a Hilbert-Schmidt operator. The adjoint $T_\kappa^{\kappa' *} : \Phi_{-\kappa} \to \Phi_{-\kappa'}$ is also Hilbert-Schmidt. $\Phi' = \cup_{k=0}^{\infty} \Phi_{-k}$ gives a representation of the dual of Φ. (See [19], page 59.) We prove the tightness for $\{S^n\}$ in $C_{\Phi_{-\kappa}}[0,\infty)$ for an appropriate κ, where $C_{\Phi_{-\kappa}}[0,\infty)$ denotes the collection of all continuous mappings from $[0,\infty)$ to $\Phi_{-\kappa}$.

Theorem 2.6. *Suppose that Conditions (I), (BC1) and (BC2) hold. Then there exists κ such that $\{S^n\}$ is tight in $C_{\Phi_{-\kappa}}[0,\infty)$.*

Finally, we characterize the limit. Let M be a $\Phi_{-\kappa}$-valued local martingale with $\langle M^\phi, Y \rangle_t = 0$ for every $\phi \in \Phi$, and

$$\langle M^\phi \rangle_t = \int_0^t \langle V_s^2, |\sigma\phi'|^2 \rangle \, ds,$$

where $M_t^\phi = \langle M_t, \phi \rangle$,$V_t^2$ is a Φ'-valued process given by

$$V_t^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (M_t^i)^2 \delta_{X_t^i}.$$

The following condition is needed in the proof of the uniqueness part of the next theorem.
Condition (E): There exists a constant $\delta > 0$ such that $\sigma^2 - \delta c^2 \geq 0$.

Theorem 2.7. *In addition to the conditions of Theorem 2.6, we assume that Condition (E) holds. Then $\{S^n\}$ converges weakly in $C_{\Phi_{-\kappa}}[0,\infty)$ to a process S which is the unique solution of the following stochastic evolution equation: $\forall \phi \in \Phi$,*

$$\langle S_t, \phi \rangle = \langle S_0, \phi \rangle + \langle M_t, \phi \rangle + \int_0^t \langle S_u, L\phi \rangle \, dy + \int_0^t \langle S_u, h\phi + c\phi' \rangle \, dY_u. \tag{2.12}$$

3. Filtering using branching particle systems

In this section, we introduce the branching particle system approximation of the optimal filter. The main purpose is to reduce the variances of the weights of the particles in the system. The idea is to divide the time interval into small subintervals and the weight for each particle at any partition time is modified as an exponential martingale which depends on the signal and the noise in the small interval prior to that time instead of on the whole interval starting from 0. This section is based on a joint paper with Crisan [12].

Now we proceed to giving the definition of the branching particle system. Initially, there are n particles of weight 1 each at locations x_i^n, $i = 1, 2, \cdots, n$, satisfying the initial condition (I).

Let $\delta = \delta_n = n^{-2\alpha}$, $0 < \alpha < 1$. For $j = 0, 1, 2, \cdots$, there are m_j^n number of particles alive at time $t = j\delta$. Note that $m_0^n = n$.

During the time interval $(j\delta, (j+1)\delta)$, the particles move according to the following diffusions: For $i = 1, 2, \cdots, m_j^n$,

$$X_t^i = X_{j\delta}^i + \int_{j\delta}^t \sigma(X_s^i) dB_s^i + \int_{j\delta}^t (b - ch)(X_s^i) ds + \int_{j\delta}^t c(X_s^i) dY_s. \tag{3.1}$$

At the end of the interval, the ith particle ($i = 1, 2, \cdots, m_j^n$) branches (independent of others) into a random number ξ_{j+1}^i of offsprings which is chosen such that the conditional variance $Var^{\hat{P}}\left(\xi_{j+1}^i | \mathcal{F}_{(j+1)\delta-}\right)$ is minimized subject to the constraint

$$\hat{\mathbb{E}}\left(\xi_{j+1}^i | \mathcal{F}_{(j+1)\delta-}\right) = \tilde{M}_j^n(X^i, (j+1)\delta), \tag{3.2}$$

where

$$M_j^n(X^i, t) = \exp\left(\int_{j\delta}^t h^*(X_s^i) dY_s - \frac{1}{2}\int_{j\delta}^t |h(X_s^i)|^2 ds\right) \tag{3.3}$$

and

$$\tilde{M}_j^n(X^i, t) = \frac{M_j^n(X^i, t)}{\frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n} M_j^n(X^\ell, t)}.$$

It is clear that

$$\xi_{j+1}^i = \begin{cases} [\tilde{M}_j^n(X^i, (j+1)\delta)] & \text{with probability } 1 - \{\tilde{M}_j^n(X^i, (j+1)\delta)\}, \\ [\tilde{M}_j^n(X^i, (j+1)\delta)] + 1 & \text{with probability } \{\tilde{M}_j^n(X^i, (j+1)\delta)\}, \end{cases}$$

where $\{x\} = x - [x]$ is the fraction of x, and $[x]$ is the largest integer which is not greater than x.

Denote the conditional variance of ξ_{j+1}^i by $\gamma_{j+1}^n(X^i)$. Then

$$\gamma_{j+1}^n(X^i) = \{\tilde{M}_j^n(X^i, (j+1)\delta)\}\left(1 - \{\tilde{M}_j^n(X^i, (j+1)\delta)\}\right).$$

Now we define the approximate filter as follows:

$$\pi_t^n = \frac{1}{m_j^n}\sum_{i=1}^{m_j^n} \tilde{M}_j^n(X^i, t)\delta_{X_t^i}, \qquad j\delta \le t < (j+1)\delta.$$

Namely, the ith particle has a time-dependent weight $\tilde{M}_j^n(X^i,t)$. At the end of the interval, i.e., $t=(j+1)\delta$, this particle dies and gives birth to a random number of offspring, whose conditional expectation is equal to the pre-death weight of the particle. The new particles start from their mother's position with weight 1 each.

The process π_t^n is called the *hybrid filter* since it involves a branching particle system and the empirical measure of these weighted particles. In the earlier stage of the study of particle approximation of the optimal filter, the particle approximation is defined as π_t^n without the weight, i.e., the *particle filter* is

$$\tilde{\pi}_t^n = \frac{1}{m_j^n}\sum_{i=1}^{m_j^n} \delta_{X_t^i}, \qquad j\delta \le t < (j+1)\delta \tag{3.4}$$

Thus, the current approximate filter π_t^n is a combination of the weighted filter introduced in Section 2 and the particle filter (3.4). That is the reason we call it the hybrid filter.

Since Zakai's equation for the unnormalized filter V_t is much simpler than the Kushner-FKK equation for the optimal filter π_t, to study the convergence of π_t^n to π_t, it is more convenient to consider an auxiliary process first. Let

$$\eta_k^n = \Pi_{j=0}^{k-1} \frac{1}{m_j^n}\sum_{\ell=1}^{m_j^n} M_j^n(X^\ell,(j+1)\delta).$$

For $k\delta \le t < (k+1)\delta$, we define

$$V_t^n = \frac{1}{n}\eta_k^n \pi_t^n \sum_{i=1}^{m_k^n} M_k^n(X^i,t) = \frac{1}{n}\eta_k^n \sum_{i=1}^{m_k^n} M_k^n(X^i,t)\delta_{X_t^i}.$$

We will prove that V_t^n converges to the unnormalized filter V_t. To this end, we need the following lemmas.

Lemma 3.1. *For each* $1\le j\le [T/\delta]$, *we have*

$$\hat{\mathbb{E}}\left(m_j^n(\eta_j^n)^2\right) \le ne^{K^2T}.$$

Lemma 3.2. *There exists a constant* K_1 *such that for any* $j\ge 0$ *and* $i=1,2,\cdots,m_j^n$, *we have*

$$\hat{\mathbb{E}}\left(\gamma_{j+1}^n(X^i)\left(\eta_{j+1}^n/\eta_j^n\right)^2|\mathcal{F}_{j\delta}\right) \le K_1\sqrt{\delta}.$$

A key ingredient in the proof of the convergence of V^n to V is the following dual of the Zakai equation:

$$\begin{cases} d\psi_s = -L\psi_s ds - (\nabla^*\psi_s c + h^*\psi_s)\,\hat{d}Y_s, & 0\le s\le t,\\ \psi_t = \phi, \end{cases} \tag{3.5}$$

where $\hat{d}$ denotes the backward Itô's integral. Namely, we take the right endpoints in the approximating Riemann sum in defining the stochastic integral.

Hereafter we will denote by $C_b^k\left(\mathbb{R}^d,\mathcal{X}\right)$ the set of all bounded continuous mappings from $\mathbb{R}^d$ to $\mathcal{X}$ with bounded partial derivatives up to order k, where $\mathcal{X}$ is a

Hilbert space. We endow $C_b^k\left(\mathbb{R}^d,\mathcal{X}\right)$ with the following norm

$$||\varphi||_{k,\infty}=\sum_{|\alpha|\le k}\sup_{x\in\mathbb{R}^d}\|D_\alpha\varphi(x)\|_{\mathcal{X}}\,,\ \ \varphi\in C_b^k\left(\mathbb{R}^d,\mathcal{X}\right),$$

where $\alpha=\left(\alpha^1,\cdots,\alpha^d\right)$ is a multi-index, $|\alpha|=\alpha^1+\cdots+\alpha^d$ and $D_\alpha\varphi=\partial_1^{\alpha^1}\cdots\partial_d^{\alpha^d}\varphi$. Also let $W_p^k\left(\mathbb{R}^d,\mathcal{X}\right)$ be the set of all functions with generalized partial derivatives up to order k with both the function and all its partial derivatives being p-integrable. We endow $W_p^k\left(\mathbb{R}^d,\mathcal{X}\right)$ with the following Sobolev norm

$$||\varphi||_{k,p}=\left(\sum_{|\alpha|\le k}\int_{\mathbb{R}^d}\|D_\alpha\varphi(x)\|_{\mathcal{X}}^p\,dx\right)^{\frac{1}{p}}.$$

The following Condition (BD) is needed in establishing a representation of ψ_s which plays a key role in the proof of the convergence of V^n.

Condition (BD): The mappings a, b, c, h, ϕ are in $C_b^k(\mathbb{R}^d,\mathcal{X})$ with $k=\left[\frac{d}{2}\right]+2$ and $\mathcal{X}$ being $\mathbb{R}^{d\times d}$, $\mathbb{R}^d$, $\mathbb{R}^{d\times m}$, $\mathbb{R}^m$ and $\mathbb{R}$ respectively. Also, we assume $\phi\in W_2^k(\mathbb{R}^d)$.

Under this condition, we can get an estimate for the supremum norm of ψ_s.

Lemma 3.3. *Suppose that Assumption (BD) holds. Then $\psi_s\in C_b^2\left(\mathbb{R}^d\right)$ a.s. and there exists a constant K_1 independent of ϕ and $s\in[0,t]$ such that*

$$\mathbb{E}[\|\psi_s\|_{2,\infty}^2]\le K_1\|\phi\|_{k,2}^2. \tag{3.6}$$

Now, we give the representation of ψ_s.

Lemma 3.4. *Suppose that Condition (BD) holds. Then, for every $t\ge 0$, we have*

$$\psi_t(X_t)M_t-\psi_0(X_0)=\int_0^t M_s\nabla^*\psi_s\sigma(X_s)dB_s,\qquad a.s.. \tag{3.7}$$

As a consequence, if $\phi\in C_b(\mathbb{R}^d)$ and $\pi_0\in L^2(\mathbb{R}^d)$, then $\langle V_t,\phi\rangle=\langle\pi_0,\psi_0\rangle$.

Note that

$$\langle V_t^n,\phi\rangle-\langle V_0^n,\psi_0\rangle=I_1^n+I_2^n+I_3^n,$$

where

$$I_1^n=\eta_k^n\frac{1}{n}\sum_{i=1}^{m_k^n}\left(M_k^n(X^i,t)\psi_t(X_t^i)-\psi_{k\delta}(X_{k\delta}^i)\right),$$

$$I_2^n=\sum_{j=1}^{k}\eta_j^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i-\tilde{M}_j^n(X^i))$$

and

$$I_3^n=\sum_{j=1}^{k}\eta_{j-1}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\left(\psi_{j\delta}(X_{j\delta}^i)M_j^n(X^i)-\psi_{(j-1)\delta}(X_{(j-1)\delta}^i)\right).$$

Applying Lemma 3.4, we get

$$I_3^n=\sum_{j=0}^{k-1}\eta_j^n\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta}M_j^n(X^i,s)\nabla^*\psi_s\sigma(X_s^i)dB_s^i$$

and hence, $\hat{\mathbb{E}}\left((I_3^n)^2\right) \le Kn^{-1}$. The term I_1^n can be estimated similarly.

Using the conditional independency for the terms in I_2^n, we get

$$\hat{\mathbb{E}}\left((I_2^n)^2\right) = \hat{\mathbb{E}}\sum_{j=1}^{k}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)^2\gamma_j^n(X^i)(\eta_j^n)^2. \tag{3.8}$$

By Lemmas 3.2 and 3.3, we get that

$$\begin{aligned}\hat{\mathbb{E}}\left((I_2^n)^2\right) &\le \hat{\mathbb{E}}\sum_{j=1}^{k}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\hat{\mathbb{E}}\left(\|\psi_{j\delta}\|_\infty^2\right)\hat{\mathbb{E}}\left(\gamma_j^n(X^i)\left((\eta_j^n)^2/\eta_{j-1}^n\right)^2\Big|\mathcal{F}_{(j-1)\delta}\right)(\eta_{j-1}^n)^2\\ &\le K_1\sqrt{\delta}\frac{1}{n^2}\sum_{j=1}^{k}\hat{\mathbb{E}}\left(m_{j-1}^n(\eta_{j-1}^n)^2\right)\\ &\le K_2\sqrt{\delta}\frac{1}{n^2}\frac{T}{\delta}n \le K_3n^{-(1-\alpha)}. \end{aligned}\tag{3.9}$$

Thus, we have

Theorem 3.5. *Suppose that the conditions (BD) and (I) hold. Then there exists a constant K_1 such that*

$$\hat{\mathbb{E}}|\langle V_t^n,\phi\rangle - \langle V_t,\phi\rangle|^2 \le K_1n^{-(1-\alpha)}\|\phi\|_{k,2}^2$$

where $k=\left[\frac{d}{2}\right]+2$ is given in Condition (BD).

To get an uniform estimate, we need the following equation satisfied by V_t^n:

$$\begin{aligned}\langle V_t^n,f\rangle &= \langle V_0^n,f\rangle + \int_0^t\langle V_s^n,Lf\rangle\,ds + \int_0^t\langle V_s^n,\nabla^*fc+h^*f\rangle\,dY_s\\ &\quad +N_t^{n,f}+\hat{N}_t^{n,f},\end{aligned}\tag{3.10}$$

where

$$N_t^{n,f} = \sum_{j=0}^{[t/\delta]}\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{((j+1)\delta)\wedge t}\nabla^*f\sigma(X_s^i)dB_s^i\eta_j^n$$

and

$$\hat{N}_t^{n,f} = \sum_{j=1}^{[t/\delta]}\eta_j^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}(\xi_j^i-\tilde{M}_j^n(X^i))f(X_{j\delta}^i)$$

are two uncorrelated martingales.

Define the usual distance on $\mathcal{M}_F(\mathbb{R}^d)$ by

$$d(\nu_1,\nu_2)=\sum_{i=0}^{\infty}2^{-i}\left(|\langle\nu_1-\nu_2,f_i\rangle|\wedge 1\right), \qquad \forall\,\nu_1,\ \nu_2\in\mathcal{M}_F(\mathbb{R}^d), \tag{3.11}$$

where $f_0=1$ and for $i\ge1$, $f_i\in C_b^{k+2}(\mathbb{R}^d)\cap W_2^{k+2}(\mathbb{R}^d)$ with $||f_i||_{k+2,\infty}\le1$ and also $||f_i||_{k+2,2}\le1$, where $k=\left[\frac{d}{2}\right]+2$ is given in Condition (BD).

Theorem 3.6. *Suppose that the conditions (BD) and (I) hold and, additionally, that* $h \in C_b^k(\mathbb{R}^d) \cap W_2^k(\mathbb{R}^d)$. *Then, there exists a constant* K_1 *such that*

$$\hat{\mathbb{E}} \sup_{t \le T} d(V_t^n, V_t)^2 \le K_1 n^{-(1-\alpha)}. \tag{3.12}$$

By Kallianpur-Striebel formula, we then get

Theorem 3.7. *Suppose that the conditions in Theorem 3.6 are satisfied. Then, there exists a constant* K_1 *such that*

$$\mathbb{E} \sup_{0 \le t \le T} d(\pi_t^n, \pi_t) \le K_1 n^{-\frac{1-\alpha}{2}}. \tag{3.13}$$

For the particle filter, we have the following estimate.

Remark 3.8. For the particle filter $\tilde{\pi}_t^n$, we have

$$\mathbb{E} \sup_{0 \le t \le T} d(\tilde{\pi}_t^n, \pi_t) \le K_1 \left(n^{-(1-\alpha)/2} \vee n^{-\alpha} \right).$$

Note that the optimal rate in this case is $\frac{1}{3}$ achieved at $\alpha = \frac{1}{3}$.

Finally, we characterize the convergence rate of π^n to π by studying the convergence of the sequence $\zeta_t^n = n^{\frac{1-\alpha}{2}} (\pi_t^n - \pi_t)$.

Theorem 3.9. ζ^n *converges weakly to a process* ζ *which is the unique solution to the following evolution equation:*

$$\begin{aligned} d\zeta_t &= \langle \zeta_t, Lf - (\pi_t(\nabla^* fc + hf) - \pi_t f \pi_t h) h \rangle \, dt \\ &\quad + \langle \zeta_t, \nabla^* fc + hf - f\pi_t h - h\pi_t f \rangle \, d\nu_t \\ &\quad - \sqrt[4]{2\pi^{-1}} \int_{\mathbb{R}^d} \frac{f - \pi_t f}{V_t 1} \sqrt{|h(x) - \pi_t h| V(t,x)} B(dtdx) \end{aligned}$$

where B *is a space-time white noise which is independent of* Y *and* ν *is the innovation process defined by (2.9).*

To bring this section to an end, we briefly mention, to the best of our knowledge, some of the related papers available in the literature.

Remark 3.10. Particle system approximation of optimal filter was studied in heuristic schemes in the beginning of the 1990s by Gordon, Salmon and Ewing [21], Gordon, Salmon and Smith [20], Kitagawa [25], Carvalho, Del Moral, Monin and Salut [1], Del Moral, Noyer and Salut [18]. The rigorous proof of the convergence results for the particle filter were published by Del Moral [15] in 1996, and independently, by Crisan and Lyons [10] in 1997. After that, many improvements have been made by various authors. Here we would like to mention only a few: Crisan and Lyons [11], Crisan [6], [4], [3], [5], Crisan, Del Moral and Lyons [7], [9], Crisan and Doucet [8], Crisan, Gaines and Lyons [9], Del Moral and Guionnet [16], Del Moral and Miclo [17].

4. Filtering using interacting particle systems

In this section, we give an interacting particle system representation of the optimal filter based on the papers of Crisan and Xiong [13], [14]. The main motivation is to seek a numerical scheme to approximate the optimal filter using neither branching nor weight. The idea is to direct the particles to move toward more likely regions, and hence, the coefficients should depend on the whole system configuration. For simplicity of notations, we take $c = 0$ in the signal equation (1.1).

Note that the innovation process ν_t can be approximated by a smooth process $\tilde{\nu}_t$. By a robust representation of the optimal filter due to Clark and Crisan [2], the optimal filter π_t can be approximated by the solution to the following PDE:

$$\frac{d}{dt}\langle \mathcal{I}_t, f\rangle = \langle \mathcal{I}_t, Lf\rangle + \left\langle \mathcal{I}_t, \alpha_t^{\mathcal{I}_t} f\right\rangle \tag{4.1}$$

where $\alpha_t = h\frac{d}{dt}\tilde{\nu}_t$ is a bounded smooth function and

$$\alpha^{\mathcal{I}} = \alpha - \langle \mathcal{I}, \alpha\rangle .$$

For mathematical convenience, we assume that the signal is a reflecting diffusion in a bounded domain $D = \{x \in \mathbb{R}^d : \ |x| \le R\}$ of $\mathbb{R}^d$, namely,

$$L = b^*\nabla + \frac{1}{2}\sum_{j,k=1}^d a^{jk}\partial^2_{jk}, \qquad D(L) = \{f : \ x^*\nabla f|_{\partial D} = 0\}$$

where $a = \sigma^*\sigma$. In fact, the diffusion on $\mathbb{R}^d$ can be approximated by such reflecting diffusions. Thus, the optimal filter with bounded signal can be regarded as an approximation of the original optimal filter.

By Conditions (BC1), (BC2) and (E), we see that the following conditions, which are needed in the study of the numerical approximation to (4.1), are satisfied:

(B) There exists a constant K such that for any $t \in [0, T]$ and $x \in D$, we have

$$|b(x)| + |\sigma(x)| + |\alpha_t(x)| \le K.$$

(Lip) For any $t \in [0, T]$ and $x, \ y \in D$, we have

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| + |\alpha_t(x) - \alpha_t(y)| \le K|x - y|.$$

(UE) There exists a constant $K_0 > 0$ such that for any $x \in D$, the matrix $a(x) - K_0 I$ is non-negative definite.

The following identity is the key for the interacting particle system representation of the optimal filter.

Lemma 4.1.

$$\int_D \frac{\nabla^* f(x)(y-x)}{\|y-x\|^d}dx = \begin{cases} -\int_{\partial D} f(x)dS + \omega_d f(y) & \text{if } y \neq 0 \\ \omega_d f(0) & \text{if } y = 0 \text{ and } f \in D(L), \end{cases}$$

where ω_d is the surface area of the d-dimensional unit sphere S_{d-1}.

Based on this identity, we can show that $\mathcal{I}_t$ has the following representation by an interacting infinite particle system.

Proposition 4.2.

$$\mathcal{I}_t = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \delta_{X_t^i} \tag{4.2}$$

where

$$dX_t^i = \tilde{b}_t(\mathcal{I}_t, X_t^i)dt + \sigma(X_t^i)dB_t^i + \mathcal{N}(X_t^i)dK_t^i, \qquad i = 1, 2, \cdots, \tag{4.3}$$

K_t^i *is the local time of* X_t^i *at the boundary of* D*,* $\mathcal{N}(x)$ *is the unit normal vector of* ∂D *at* $x \in \partial D$*,*

$$\tilde{b}_t(\mathcal{I}_t, x) = b(x) + \frac{\Lambda_{\mathcal{I}_t}\alpha_t(x)}{\mathcal{I}_t(x)}$$

and

$$\Lambda_{\mathcal{I}}\alpha(x) = \frac{1}{\omega_d}\int_{\mathbb{R}^d} \frac{(y-x)\alpha^{\mathcal{I}}(y)}{\|y-x\|^d}\mathcal{I}(dy).$$

Note that X_t^i given by (4.3) is a diffusion on D with reflecting boundary.

Remark 4.3. From the definition of $\tilde{b}$ we see that the particles move fast in unlikely region (with $\mathcal{I}_t(x)$ small) and slow in more likely region (with $\mathcal{I}_t(x)$ large). Therefore, the particles spend more time in more likely region.

Based on the representation above, we now propose an approximation to $\{\mathcal{I}_t\}$ by finite interacting particle systems. For each finite system, the empirical measure has no density, so we smooth it out by the operator T_ϵ given below. For the convenience of the estimates, we introduce an extra parameter $\delta > 0$ to make the coefficient $\tilde{b}$ bounded. Namely, we fix $n \in \mathbb{N}$ and $\epsilon,\ \delta > 0$ and consider the following finite system: For $i = 1, 2, \cdots n$,

$$dX_t^{n,\epsilon,\delta,i} = \tilde{b}_t^{\epsilon,\delta}(\mathcal{I}_t^{n,\epsilon,\delta}, X_t^{n,\epsilon,\delta,i})dt + \sigma(X_t^{n,\epsilon,\delta,i})dB_t^i + \mathcal{N}(X_t^{n,\epsilon,\delta,i})dK_t^{n,\epsilon,\delta,i} \tag{4.4}$$

and

$$\mathcal{I}_t^{n,\epsilon,\delta} = \frac{1}{n}\sum_{i=1}^n \delta_{X_t^{n,\epsilon,\delta,i}}, \qquad \tilde{b}_t^{\epsilon,\delta}(\mu, x) = b(x) + \frac{\Lambda_\mu \alpha_t(x)}{T_\epsilon\mu(x) + \delta} \tag{4.5}$$

where

$$T_\epsilon\mu(x) = (2\pi\epsilon)^{-\frac{d}{2}}\int \exp\left(-\frac{|x-y|^2}{2\epsilon}\right)\mu(dy).$$

Here is the main convergence theorem in this section.

Theorem 4.4. *For any* $t > 0$ *fixed, we have*

$$\lim_{\delta\to0}\lim_{\epsilon\to0}\lim_{n\to\infty} \mathcal{I}_t^{n,\epsilon,\delta} = \mathcal{I}_t.$$

Finally, we use Euler's scheme to solve the SDE system (4.4). The main difference from the Euler scheme in Section 2 is the involvement of the local time process $K_t^{n,\epsilon,\delta,i}$ in each equation. We use a penalization method (cf. Pettersson [29] and Slominski [30]) to deal with this problem. For simplicity of notation, we drop the superscripts $n,\ \epsilon,\ \delta$ in this part. The system (4.4) becomes

$$\begin{cases} dX_t^i = \tilde{b}_t(\mathcal{I}_t, X_t^i)dt + \sigma(X_t^i)dB_t^i + \mathcal{N}(X_t^i)dK_t^i \\ \mathcal{I}_t = \frac{1}{n}\sum_{i=1}^n \delta_{X_t^i}, \qquad \tilde{b}_t(\mu, x) = b(x) + \frac{\Lambda_\mu\alpha_t(x)}{T_\epsilon\mu(x)+\delta}. \end{cases}$$

Let $0 = t_0 < t_1 < \cdots < t_{c_\gamma} = T$ be a partition of $[0, T]$ with mesh size $\gamma = \max_{1\le k\le c_\gamma} \Delta t_k$ where $\Delta t_k = t_k - t_{k-1}$. Let Π_D be the projection to D (i.e. $\Pi_D x$ is the point in D which is the closest to x). For $0 \le t < t_1$, define $X^{\gamma,i}_t = x^i_0$; and for $t_k \le t < t_{k+1}$, $k \ge 1$,

$$\begin{cases} X^{\gamma,i}_t = \Pi_D\{X^{\gamma,i}_{t_{k-1}} + \tilde{b}_{t_{k-1}}(\mathcal{I}^\gamma_{t_{k-1}}, X^{\gamma,i}_{t_{k-1}})\Delta t_k + \sigma(X^{\gamma,i}_{t_{k-1}})\Delta B^i_{t_k}\} \\ \mathcal{I}^\gamma_t = \frac{1}{n}\sum_{i=1}^n \delta_{X^{\gamma,i}_t} \end{cases} \tag{4.6}$$

where $\Delta B^i_{t_k} = B^i_{t_k} - B^i_{t_{k-1}}$.

Let $\mathcal{I}^{n,\epsilon,\delta,\gamma}_t$ be the process $\mathcal{I}^\gamma_t$ given by (4.6). Then

Theorem 4.5. *There exists a constant $C_{\epsilon,\delta}$ such that*

$$\sup_{0\le t\le T} \mathbb{E}\rho(\mathcal{I}^{n,\epsilon,\delta,\gamma}_t, \mathcal{I}^{n,\epsilon,\delta}_t)^2 \le C_{\epsilon,\delta}\left(\gamma \log \frac{1}{\gamma}\right)^{1/2}.$$

Remark 4.6. In Section 2, we saw that the time-varying weights of the particles cause the approximate error to increase exponentially fast in time. To overcome this drawback, branching particle systems were introduced in Section 3 so that the individual weight of any particle depends on its path in small time intervals only. In the present section, we have provided an alternative with constant weights hoping to reduce the approximate error, or equivalently, to increase the convergence rate. It remains a challenging *open* problem to derive an estimate of the convergence rate and to compare it with that of the hybrid filter of Section 3.

Remark 4.7. As we have seen from this section, the numerical scheme based on interacting particle system with neither branching nor weight involves quite many approximating procedures ($\tilde{\nu} \to \nu$, $D \uparrow \mathbb{R}^d$, ϵ, $\delta \to 0$, $n \to \infty$). How do propose more efficient approximation (i.e., with less approximating procedures) remains a challenging *open* problem.

5. A filtering model with point process observations

The filtering model [(1.1), (1.3)] with point processes observations, called the Filtering Micromovement (FM) model, is proposed by Zeng [32]. The signal process X_t represents the intrinsic value process of d assets, which corresponds to the macro-movement in the empirical econometric literature or the continuous-time price process in the option pricing literature. Prices are observed only at random trading times which are modelled by a conditional Poisson process. Moreover, prices are distorted observations of the intrinsic value process at the trading times. The observed prices take values in the discrete set of m levels. Thus, an alternative description for the observation process is Y^i_t which counts the number of times that level i price are observed before time t. Then, Y^i_t, $i = 1, 2, \cdots, m$, are Poisson point processes whose intensities depend on the intrinsic value process.

The Zakai equation and the filtering equation are derived in [32]. The particle approximation for the optimal filter is studied by Xiong and Zeng [31] on which this section is based.

For simplicity of notations, we assume $c = 0$ in the signal equation (1.1). In addition to the Conditions (I) and (BC1), we assume the following Condition (P) on the intensities of the point processes introduced by (1.3).

Condition (P): i) The total intensity $a(x,t) \equiv \sum_{k=1}^m \lambda_k(x,t)$ *is bounded.*

ii) Let $a'(x,t) \equiv \frac{\partial}{\partial x} a(x,t)$ *and* $p'_k(x,t) \equiv \frac{\partial}{\partial x} p_k(x,t)$, *where*

$$p_k(x,t) = a(x,t)^{-1}\lambda_k(x,t).$$

Then $a'(x,t)$ *and* $p'_k(x,t)$ *are bounded and continuous in* x.

Let M_t be the solution to the following SDE:

$$dM_t = \sum_{k=1}^m \left(\lambda_k(X_s, s-) - 1\right) M_{t-} d\left(Y_t^k - t\right), \qquad M_0 = 1. \tag{5.1}$$

Let $\hat{P}$ be the probability measure given by $\left.\frac{dP}{d\hat{P}}\right|_{\mathcal{F}_t} = M_t$. By Girsanov's theorem, we know that, under $\hat{P}$, $Y_t^k,\ k = 1, \cdots, m$ are independent standard Poisson point processes, and they are independent of the Brownian motion B. The Kallianpur-Striebel formula in the current setup gives

$$\langle \pi_t, f\rangle = \frac{\langle V_t, f\rangle}{\langle V_t, 1\rangle},$$

where

$$\langle V_t, f\rangle = \hat{\mathbb{E}}\left(f(X_t)M_t \middle| \mathcal{G}_t\right).$$

Note that V_t has the following weighted particle system representation:

$$V_t = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n M_t^i \delta_{X_t^i}$$

where

$$dX_t^i = b(X_t^i)dt + \sigma(X_t^i)dB_t^i$$

and

$$dM_t^i = \sum_{k=1}^m \left(\lambda_k(X_s^i, s-) - 1\right) M_{t-}^i d\left(Y_t^k - t\right), \qquad M_0 = 1.$$

Applying Itô's formula to the above system, we get the Zakai equation in this model

$$\langle V_t, f\rangle = \langle V_0, f\rangle + \int_0^t \langle V_s, Lf\rangle ds + \sum_{k=1}^m \int_0^t \langle V_{s-}, (ap_k - 1)f\rangle d(Y_s^k - s). \tag{5.2}$$

Making use of the Kallianpur-Striebel and Itô formulas, we have the filtering equation:

$$\begin{aligned}\langle \pi_t, f\rangle =& \langle \pi_0, f\rangle + \int_0^t \Big[\langle \pi_s, Lf\rangle - \langle \pi_s, fa\rangle + \langle \pi_s, f\rangle\langle \pi_s, a\rangle\Big] ds \\ &+ \sum_{k=1}^m \int_0^t \Big[\frac{\langle \pi_{s-}, f\lambda_k\rangle}{\langle \pi_{s-}, \lambda_k\rangle} - \langle \pi_{s-}, f\rangle\Big] dY_s^k.\end{aligned} \tag{5.3}$$

When $a(X(t),t)=a(t)$, the above equation is simplified as:

$$\langle \pi_t, f\rangle =\langle \pi_0, f\rangle + \int_0^t \langle \pi_s, Lf\rangle ds + \sum_{k=1}^m \int_0^t \Big[\frac{\langle \pi_{s-}, f\lambda_k\rangle}{\langle \pi_{s-}, \lambda_k\rangle} - \langle \pi_{s-}, f\rangle\Big] dY_s^k. \tag{5.4}$$

Note that a branching weighted particle system can be defined similar to that in Section 3. We will not repeat its definition here. Instead, we shall only study its properties. To this end, we use the same notations as in Section 3. The main difference is the following lemma which is the counterpart of Lemma 3.2 in the present setup.

Lemma 5.1. *Let $F(x)=\{x\}(1-\{x\})$. Then, for bounded $f(x)$ with bounded Lf^2, as $\delta\to 0$, we have*

$$\Big|\mathbb{E}\Big(\gamma_{j+1}^n(X^i) f^2(X^i_{(j+1)\delta})(\eta^n_{j+1}/\eta^n_j)^2|\mathcal{F}_{j\delta}\Big) - (f^2 H^n_{j\delta})(X^i_{j\delta})\delta\Big| = o(\delta),$$

where H^n_s is a nonnegative function given by

$$H_s^n(x) = \sum_{k=1}^m F\left(\frac{\lambda_k(x,s)}{\bar{h}_k^n(s)+1}\right)(\bar{h}^n_k(s)+1)^2 \tag{5.5}$$

and

$$\bar{h}^n_k(s) = \langle \pi^n_s, \lambda_k - 1\rangle . \tag{5.6}$$

As a consequence of the lemma, we have that

$$\mathbb{E}\Big(\gamma_{j+1}^n(X^i)(\eta^n_{j+1}/\eta^n_j)^2|\mathcal{F}_{j\delta}\Big) \le K_1\delta,$$

which is different from Lemma 3.2.

Based on this estimate, it follows from the same argument as in (3.9) that

$$\hat{\mathbb{E}}\left((I_2^n)^2\right) \le K_1\delta\frac{1}{n^2}\frac{T}{\delta} n \le K_2 n^{-1},$$

which is in contrast with the bound $K_3 n^{-(1-\alpha)}$ for the classical case.

As we did in Section 3, we consider the dual of the Zakai equation which is the following backward SPDE:

$$\begin{cases} d\psi_s = -L\psi_s ds - \sum_{k=1}^w (ap_k - 1)\psi_{s+}\hat{d}(Y_s^k - s), & 0\le s\le t\\ \psi_t = \phi \end{cases} \tag{5.7}$$

where $\hat{d}$ denotes the backward Itô's integral and ϕ is a bounded function. Actually, the current setup makes this equation easier to handle because it can be studied as a piecewise PDE. Thus, we get the following lemma easily.

Lemma 5.2. *Suppose that Conditions (I), (BC1) and (P) hold for the FM model. Let $\psi_u'(x) = \frac{d}{dx}\psi_u(x)$. Then, there exists a constant K such that*

$$\hat{\mathbb{E}}[\sup_{0\le s\le t}\|\psi_s\|_\infty + \sup_{0\le s\le t}\|\psi_s'\|_\infty] \le K.$$

Based on this lemma, we can get the following key identity.

Lemma 5.3. *Suppose that Conditions (I), (BC1) and (P) hold. Then, for every* $t \geq 0$, *we have*

$$\psi_t(X_t)M_t - \psi_0(X_0) = \int_0^t M_s \nabla^* \psi_s \sigma(X_s) dB_s, \qquad a.s.. \tag{5.8}$$

Note that Lemma 5.3 is in the same form as Lemma 3.4. However, the restrictive Condition (BD) is not required here.

By the same arguments as those leading to Theorem 3.5, we get

Theorem 5.4. *Suppose that Conditions (I), (BC1) and (P) hold. Then there exists a constant* K_1 *such that*

$$\hat{\mathbb{E}} |\langle V_t^n, \phi \rangle - \langle V_t, \phi \rangle|^2 \leq K_1 n^{-1}.$$

Next, by an equation similar to (3.10), we can prove that

Theorem 5.5. *Under the assumptions of Theorem 5.4, there exists a constant* K *such that*

$$\mathbb{E} \sup_{0 \leq t \leq T} d(\pi_t^n, \pi_t) \leq K n^{-\frac{1}{2}}. \tag{5.9}$$

Finally, we characterize the rate of convergence by studying $\zeta_t^n = n^{\frac{1}{2}}(\pi_t^n - \pi_t)$. Let Φ be the nuclear space defined in Section 2.

Theorem 5.6. *Under the assumptions of Theorem 5.4,* ζ^n *converges weakly in* $D_{\Phi_{-\kappa}}[0, \infty)$ *to a process* ζ *which is the unique solution to the following evolution equation:* $\forall\ f \in \Phi$,

$$\begin{aligned} d\langle \zeta_t, f \rangle =& \langle \zeta_t, Lf - (a-w)f - f\langle \pi_t, a-w \rangle + (a-w)\langle \pi_t, f \rangle \rangle dt \\ &+ \sum_{k=1}^{w} \Big[\frac{\langle \zeta_{t-}, fap_k \rangle}{\langle \pi_{t-}, ap_k \rangle} - \frac{\langle \zeta_{t-}, ap_k \rangle \langle \pi_{t-}, fap_k \rangle}{\langle \pi_{t-}, ap_k \rangle^2} - \langle \zeta_{t-}, f \rangle \Big] dY_t^k \\ &+ \int_{\mathbb{R}} \frac{f(x) - \langle \pi_t, f \rangle}{\langle V_t, 1 \rangle} \sqrt{H_t(x) V_t(x) \langle V_t, 1 \rangle} W(dxdt), \end{aligned} \tag{5.10}$$

where W *is a space-time white noise independent of* Y, $H_s^n(x)$ *is a nonnegative function given by*

$$H_s(x) = \sum_{k=1}^{m} F\left(\frac{\lambda_k(x,s)}{\bar{h}_k(s) + 1} \right) (\bar{h}_k(s) + 1)^2 \tag{5.11}$$

and

$$\bar{h}_k(s) = \langle \pi_s, \lambda_k - 1 \rangle. \tag{5.12}$$

References

1. H. Carvalho, P. Del Moral, A. Monin and G. Salut (1997). Optimal nonlinear filtering in GPS/INS integration. *IEEE Trans. Aerosp. Electron. Syst.,* **33**, no. 3, 835-850.
2. J. M. C. Clark and D. Crisan (2005). On a robust version of the integral representation formula of nonlinear filtering. *Probab. Theory Related Fields* **133**, no. 1, 43-56.
3. D. Crisan (2001). Particle filters—a theoretical perspective. Sequential Monte Carlo methods in practice, 17–41, Stat. Eng. Inf. Sci., Springer, New York.

4. D. Crisan (2002). *Numerical methods for solving the stochastic filtering problem*, Numerical methods and stochastics (Toronto, ON, 1999), 1–20, Fields Inst. Commun., 34, Amer. Math. Soc., Providence, RI.
5. D. Crisan (2003). Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation. *Ann. Probab.* **31**, no. 2, 693–718.
6. D. Crisan (2004). Superprocesses in a Brownian environment. Stochastic analysis with applications to mathematical finance. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460**, *no. 2041*, 243-270.
7. D. Crisan, P. Del Moral and T. Lyons (1999). Interacting particle systems approximations of the Kushner-Stratonovitch equation. *Adv. in Appl. Probab.* **31**, no. 3, 819–838.
8. D. Crisan and A. Doucet (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.* **50**, no. 3, pp 736–746.
9. D. Crisan, J. Gaines and T. Lyons (1998). Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Math.* **58**, no. 5, 1568–1590 (electronic).
10. D. Crisan and T. Lyons (1997). Nonlinear filtering and measure-valued processes. *Probab. Theory Related Fields,* **109**, 217-244.
11. D. Crisan and T. Lyons (1999). A particle approximation of the solution of the Kushner-Stratonovitch equation. *Probab. Theory Related Fields* **115**, no. 4, 549–578.
12. D. Crisan and J. Xiong (2007). A central limit type theorem for particle filter. *Comm. Stoch. Analysis* **1**, no. 1, 103-122.
13. D. Crisan and J. Xiong (2008). Approximate McKean-Vlasov representations for a class of SPDEs. To appear in *Stochastics*.
14. D. Crisan and J. Xiong (2008). Numerical solutions for a class of SPDEs over bounded domains. *In preparation*.
15. P. Del Moral (1996). Non-linear filtering: interacting particle resolution. *Markov Process. Related Fields* **2**, no. 4, 555-581.
16. P. Del Moral and A. Guionnet (1999) Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.* **9**, no. 2, 275–297.
17. P. Del Moral and L. Miclo (2000). Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. *Séminaire de Probabilités, XXXIV, Lecture Notes in Math.,* **1729**, Springer, Berlin, 1-145.
18. P. Del Moral, J. C. Noyer and G. Salut (1995). Résolution particulaire et traitement non-linéaire du signal: application radar/sonar. In *traitement du signal* **12**, no. 4, 287-301.
19. I. M. Gel'fand and N. Ya. Vilenkin (1964). *Generalized functions. Vol. 4: Applications of harmonic analysis.* Academic Press, New York - London.
20. N. J. Gordon, D. J. Salmon and A. F. M. Smith (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F,* **140**, 107-113.
21. N. J. Gordon, D. J. Salmon and C. Ewing (1995). Bayesian state estimation for tracking and guidance using the bootstrap filter. *J. Guidance Control Dyn.,* **18**, no. 6, 1434-1443.
22. M. Hitsuda and I. Mitoma (1986). Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions. *J. Multivariate Anal.* 19, 311–328.
23. G. Kallianpur and C. Striebel (1968). Estimation of stochastic systems: arbitrary system process with additive noise observation errors. *Ann. Math. Statist.,* **39**, 785-801.
24. G. Kallianpur and C. Striebel (1969). Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes. *Teor. Veroyatn. Primen.,* **14**, no. 4, 597-622.
25. G. Kitagawa (1996). Monte-Carlo filter and smoother for non-Gaussian non-linear state space models. *J. Comput. and Graphical Stat.,* **5**, no. 1, 1-25.
26. T. Kurtz and J. Xiong (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.* **83**, 103-126.
27. T. Kurtz and J. Xiong (2000). Numerical solutions for a class of SPDEs with application to filtering. *Stochastics in Finite and Infinite Dimension: In Honor of Gopinath Kallianpur.* Edited by T. Hida, R. Karandikar, H. Kunita, B. Rajput, S. Watanabe and J. Xiong. *Trends in Mathematics*. Birkhäuser.

28. T. Kurtz and J. Xiong (2004). A stochastic evolution equation arising from the fluctuation of a class of interacting particle systems. *Communication Mathematical Sciences* **2**, 325-358.
29. R. Pettersson (1997). Penalization schemes for reflecting stochastic differential equations. *Bernoulli* **3**, No. 4, 403-414.
30. L. Slominski (2001). Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Processes and their Applications* **94**, No. 2, 317-337.
31. J. Xiong and Y. Zeng. A Branching Particle Approximation to the Filtering Problem with Counting Process Observations. *Submitted*.
32. Y. Zeng (2003). A partialy observed model for micromovement of asset prices with Bayes estimation via filtering, *Mathematical Finance*, **13** 411-444.

Department of Mathematics, University of Tennessee, Knoxville, TN 37996-1300, USA and Department of Mathematics, Hebei Normal University, Shijiazhuang 050016, PRC

E-mail address: jxiong@math.utk.edu