Selected Essays: Communications on Stochastic Analysis
Dedicated to Professor Kiyosi Ito
Vol. 22 No. 2 (July-December, 2018) 103-122

A CENTRAL LIMIT TYPE THEOREM FOR A CLASS OF PARTICLE FILTERS

DAN CRISAN AND JIE XIONG*

Abstract. The optimal filter $\pi = \{\pi_t, t \geq 0\}$ for a general observation model is approximated by a probability measure valued process $\pi^n = \{\pi_t^n, t \geq 0\}$. The process π^n is the empirical measure of a system of weighted particles that at time 0 consists of n particles. The particles branch at equally spaced time instances $jn^{-2\alpha}$ where $j = 1, 2, ...$ and $0 < \alpha < 1$. We prove the convergence of the process π^n to π and derive sharp upper bounds for the mean square error. We also prove a central limit theorem to characterize the convergence rate of the approximate filter. A similar result is obtained for the unweighted, unnormalized version introduced in [8]. As a corollary, we show that $\alpha = \frac{1}{3}$ is the optimal exponent for that version.

Received: 23rd July 2017 Revised: 14th December 2017 Accepted: 15th January 2018

1. Introduction

The approximation of the optimal nonlinear filter by means of particle approximations has been studied extensively in last ten years (see, for example, [5], [7], [10], [12] and the references therein). The use of particle approximations stems from the fact that the unnormalized filter can be approximated by a weighted particle system. Since the weights have variances which grow exponentially fast, the particle system needs to be corrected after small time steps to control the error. At each time step, the particles will be replaced by a random number of "offsprings". The expected number of offsprings is the weight of the corresponding particle decided according to its path during the period prior to that time step.

In the following, we will work within a very general filtering framework. Namely, we will assume that the observation process takes values in a space of measures (rather than the usual k-dimensional Euclidean space). In particular, the observation process can be given by a random measure in space and time. Moreover, we will allow the observation and signal noises to be correlated. Let us now introduce the filtering model in more detail.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P_0)$ be a filtered probability space and $(H, \langle \cdot, \cdot \rangle_H)$, respectively $\left(H_1, \langle \cdot, \cdot \rangle_{H_1}\right)$ be two real separable Hilbert spaces such that $H \subseteq H_1$ and the injection from H to H_1 is a Hilbert-Schmidt operator. On $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P_0)$ we define an $\mathcal{F}_t$-adapted d-dimensional Brownian motion $B = \{B_t, \mathcal{F}_t, t \geq 0\}$ and an $\mathcal{F}_t$-adapted H-cylindrical Brownian motion (H-c.B.m.) $W = \{W_t, \mathcal{F}_t, t \geq 0\}$

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 is supported partially by NSA.

independent of B. We also introduce a $d-$dimensional stochastic process $X = \{X_t, \mathcal{F}_t,\ t \geq 0\}$ (the *signal* process) which is the unique solution of the stochastic differential equation

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds + \int_0^t c(X_s)dW_s \tag{1.1}$$

and an H_1-valued stochastic process $Y = \{Y_t, \mathcal{F}_t,\ t \geq 0\}$ (the *observation* process) defined by the formula

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

where the coefficients $\sigma :\ \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $b :\ \mathbb{R}^d \to \mathbb{R}^d$, $c :\ \mathbb{R}^d \to \mathbb{R}^d \otimes H$ and $h :\ \mathbb{R}^d \to H$ are Lipschitz continuous maps. We are interested in approximating the optimal filter $\pi_t = P_0(\cdot|\mathcal{G}_t)$, that is the conditional distribution of X_t given the observation σ-field $\mathcal{G}_t = \sigma(Y_s :\ s \leq t)$ (the information available at time t).

Particular cases of the above framework are the finite dimensional case ($H = H_1 = \mathbb{R}^m$) and the filtering model

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t b(X_s)ds + \int_0^t \int_U \sigma(X_s, u)\tilde{W}(duds)$$

with observation

$$Y(A,t) = \int_0^t \int_A \tilde{h}(X_s, u)\mu(du)ds + \tilde{W}(A,t), \qquad \forall\ A \in \mathcal{B}(U)$$

where $(U, \mathcal{B}(U))$ is a measurable space, μ is a σ-finite measure on U and $\tilde{W}$ is a Gaussian random measure on $U \times \mathbb{R}_+$ with intensity measure μ. We can convert this model to (1.1-1.2) by defining $H = L^2(U, \mu)$ and

$$W_t^f = \int_0^t \int_U f(u)\tilde{W}(duds), \qquad \forall f \in H.$$

Then W_t is an H-c.B.M and the filtering problem is given by (1.1-1.2). In this example, H_1 is the completion of H with respect to the norm $\|\cdot\|_1$ given by

$$\|\cdot\|_1^2 = \sum_j \langle \cdot, h_j \rangle_H j^{-2},$$

where $\{h_j\}$ is a complete orthonormal basis of H.

Let P be a probability measure given by

$$\left.\frac{dP_0}{dP}\right|_{\mathcal{F}_t} = \exp\left(\int_0^T h(X_s)dY_s - \frac{1}{2}\int_0^t |h(X_s)|_H^2 ds\right),\ t \geq 0.$$

Using Girsanov theorem we see that Y becomes an H-c.B.m. under P which is independent of B. The signal can be rewritten as

$$X_t = X_0 + \int_0^t \sigma(X_s)dB_s + \int_0^t \tilde{b}(X_s)ds + \int_0^t c(X_s)dY_s$$

where

$$\tilde{b}(x) = b(x) - \langle c, h \rangle_H (x) \text{ and } \langle c, h \rangle_H (x) = \langle c(x), h(x) \rangle_H .$$

By Kallianpur-Striebel formula, the optimal filter can be written as

$$\langle \pi_t, f\rangle = \mathbb{E}^{P_0}\left(f(X_t)|\mathcal{G}_t\right) = \frac{\langle V_t, f\rangle}{\langle V_t, 1\rangle}, \qquad \forall\, f \in C_b(\mathbb{R}^d)$$

where

$$\langle V_t, f\rangle = \mathbb{E}\left(M(t)f(X_t)|\mathcal{G}_t\right)$$

and

$$dM(t) = M(t)h(X_t)dY_t.$$

As in the classical framework (cf, for example, [1]) one can show that V is the unique solution of the following linear equation, called the Zakai equation

$$\langle V_t, f\rangle = \langle V_0, f\rangle + \int_0^t \langle V_s, Lf\rangle\, ds + \int_0^t \langle V_s, \nabla^* fc + hf\rangle\, dY_s. \tag{1.3}$$

Next, we introduce the branching interacting particle system to be used to approximate the optimal filter. We start with n particles of weight $\frac{1}{n}$ each at $x_i^n,\ i = 1, 2, \cdots, n$ (the initial position of the particles may be random). We define $V_0^n = \frac{1}{n}\sum_{i=1}^n \delta_{x_i^n}$ and asume that, P_0-almost surely, $\lim_{n\to\infty} V_0^n = \pi_0$ in $M_F(\mathbb{R}^d)$, where $\mathcal{M}_F(\mathbb{R}^d)$ is the set of finite measures over the Borel σ-field on $\mathbb{R}^d$ and the above convergence is taken in the weak topology.

Let $\delta = \delta_n = n^{-2\alpha}$, $0 < \alpha < 1$. At time $t = j\delta$, there are m_j^n particles alive. During the time interval $(j\delta, (j+1)\delta)$, the particles move according to the following diffusions: For $i = 1, 2, \cdots, m_j^n$,

$$X_t^i = X_{j\delta}^i + \int_{j\delta}^t \sigma(X_s^i)dB_s^i + \int_{j\delta}^t \tilde{b}(X_s^i)ds + \int_{j\delta}^t c(X_s^i)dY_s,$$

where $B^i = \left\{B_t^i, \mathcal{F}_t,\ t \ge 0\right\}$ are $\mathcal{F}_t$-adapted d-dimensional Brownian motions independent of Y.

At the end of the interval, the i-th particle $(i = 1, 2, \cdots, m_j^n)$ branches (independent of others) into a random number ξ_{j+1}^i of offsprings such that

$$\begin{aligned}
\mathbb{E}\left(\xi_{j+1}^i|\mathcal{F}_{(j+1)\delta-}\right) &= \tilde{M}_{j+1}^n(X^i) \\
Var\left(\xi_{j+1}^i|\mathcal{F}_{(j+1)\delta-}\right) &= \gamma_{j+1}^n(X^i),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{M}_j^n(X^i) &= \frac{M_j^n(X^i)}{\frac{1}{m_{j-1}^n}\sum_{\ell=1}^{m_{j-1}^n} M_j^n(X^\ell)} \\
M_j^n(X^i) &= \exp\left(\int_{(j-1)\delta}^{j\delta} h(X_t^i)dY_t - \frac{1}{2}\int_{(j-1)\delta}^{j\delta} |h(X_t^i)|_H^2 dt\right).
\end{aligned}$$

To minimize γ_j^n, we take

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

where $\{x\} = x - [x]$ is the fraction of x. In this case

$$\gamma_j^n(X^i) = \{\tilde{M}_j^n(X^i)\}(1 - \{\tilde{M}_j^n(X^i)\}).$$

The approximation to the optimal filter is then given by the process $\pi^n = \{\pi^n_t, t \geq 0\}$ defined by

$$\pi^n_t = \frac{1}{n}\sum_{i=1}^{m^n_j} M^n_j(X^i, t)\delta_{X^i_t}, \qquad j\delta \leq t < (j+1)\delta,$$

where

$$M^n_j(X^i, s) = \exp\left(\int_{j\delta}^s h(X^i_t)dY_t - \frac{1}{2}\int_{j\delta}^s |h(X^i_t)|^2_H dt\right).$$

In the analysis that follows we will make use of an additional process $V^n = \{V^n_t, t \geq 0\}$ defined by $V^n_t = \pi^n_t \eta^n_t$, $t \geq 0$, where

$$\eta^n_t = \Pi_{j=0}^{k-1}\frac{1}{m^n_j}\sum_{\ell=1}^{m^n_j} M^n_{j+1}(X^\ell), \qquad \text{if } k\delta \leq t < (k+1)\delta.$$

We will show that V^n converges to V.

In [6], the optimal filter is approximated by a particle filter $\tilde{\pi}^n$ consisting of particles without weights but with the same motion law and branching mechanism as those used to construct π^n, namely

$$\tilde{\pi}^n_t = \frac{1}{n}\sum_{i=1}^{m^n_j}\delta_{X^i(t)}, \qquad j\delta \leq t < (j+1)\delta, \tag{1.4}$$

In [8], another unweighed particle approximation $\hat{V}^n = \{\hat{V}^n_t, t \geq 0\}$ was introduced, an approximation not to the optimal filter but to the unnormalised filter V. To obtain it, the conditional expectation of ξ^i_j given $\mathcal{F}_{j\delta-}$ was chosen to be $M^n_j(X^i)$ instead of $\tilde{M}^n_j(X^i)$ and the approximation was given by

$$\hat{V}^n_t = \frac{1}{n}\sum_{i=1}^{m^n_j}\delta_{X^i(t)}, \qquad j\delta \leq t < (j+1)\delta. \tag{1.5}$$

We would like to differentiate between unweighed particle filters such as $\tilde{\pi}^n$ and $\hat{V}^n$ and the above weighted approximation π^n. Since the particles that form π^n have both weights as in [18] and [19] and branching mechnisms as in [8] and [6], we will call π^n a *hybrid* filter. The approximation introduced in [5] is also a hybrid filter. It differs from π^n through the choice of the branching mechanism (the number of offsprings of the particles are no longer mutually independent so the total number of particles stays constant) and the fact that the weights are normalised so that the approximation is a probability measure. We have yet to understand the asymptotic behaviour of the branching mechanism used in [5]. That is why we use here the independent branching mechanism instead.

In the following we will prove the convergence of π^n (and $\tilde{\pi}^n$) to π as $n \to \infty$ and study the corresponding convergence rate. It turns out that the best rate cannot be achieved for π^n_t, while it is achieved for $\tilde{\pi}^n_t$ when $\alpha = \frac{1}{3}$. Nevertheless, the convergence rate for π^n_t when $\alpha < \frac{1}{3}$ is better than the optimal rate for $\tilde{\pi}^n_t$. We will prove this fact via a central limit type theorem in a modification of the

Schwartz distribution space. Hence attaching weights to particles is certainly advantageous.

As mentioned above, the limiting behavior of π^n is shown via the convergence of V^n to V. In particular, we prove that

$$\lim_{n\to 0} \mathbb{E}[\sup_{0\le t\le T} d(V_t^n, V_t)^2] = 0, \tag{1.6}$$

where $d(\cdot,\cdot)$ is a suitable distance defined on $M_F(\mathbb{R}^d)$. This result is stronger than, for example, the corresponding result in [8] where only the convergence of $\mathbb{E}[d(\hat{V}_t^n, V_t)]$ to 0 is proved (supplemented with the tightness of the sequence $\{\hat{V}^n\}$ in $D([0,T], M_F(\mathbb{R}^d)))$. Also the model presented here is more general than in any of the existing papers. The central limit theorem presented below is the first result of this type for any of the particle filters enumerated above. In [10], [11] and [12], similar results are proved for a class of unweighted particle filter which uses a *multinomial* branching mechanism. See also [2] and [17] for central limit theorems in the discrete time framework.

This paper is organized as follows: In Section 2 we prove the convergence of the approximating filter V_t^n to the optimal filter V_t, for arbitrary, but fixed, $t \ge 0$. This preliminary convergence result is used in Section 3 to prove the stronger version (1.6). Finally, in Section 4, we establish a central limit type theorem to characterize the convergence rate of the approximating filter. The corresponding results for $\tilde{\pi}_t^n$ and $\hat{V}_t^n$ are also briefly discussed.

Throughout this paper, we shall use K with a subscript to denote a constant whose value might be different in different proofs.

2. Preliminary results

Hereafter we will denote by $C_b^m\left(\mathbb{R}^d, \mathcal{X}\right)$ to be the set of all bounded continuous maps from $\mathbb{R}^d$ to $\mathcal{X}$ with bounded partial derivatives up to order m, where $\mathcal{X}$ is a Hilbert space. We endow $C_b^m\left(\mathbb{R}^d, \mathcal{X}\right)$ with the following norm

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

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

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

When $\mathcal{X}$ is clear from the context or $\mathcal{X} = \mathbb{R}$, we will drop it from the notation for simplicity. The main tool for showing the convergence of V_t^n to V_t for fixed t is the dual $\psi = \{\psi_s, s \in [0,t]\}$ of the process V. The dual of V is the solution of the backward SPDE.

$$\begin{cases} d\psi_s = -L\psi_s ds - \left(\nabla^*\psi_s c + h\psi_s\right) \hat{d}Y_s, & 0 \le s \le t \\ \psi_t = \phi \end{cases} \tag{2.1}$$

where $\hat{d}$ denotes the backward Itô's integral. Namely, we take the right point in the Riemann sum when defining the stochastic integral. The same idea has been used in previous papers ([5], [8], [6], etc.). The dificulty here is that the backward SPDE (2.1) is driven by an H-c.B.m hence all classical estimates (such as those that appear in Rozovskii [21]) are no longer available. We need to prove them ourselves and we do so shortly. Further, because the correlation of the noises (observation and signal), some of the estimates have to be carefully refined. Let us define

$$\tilde{Y}_s = Y_t - Y_{t-s} \text{ and } \tilde{\psi}_s = \psi_{t-s}.$$

Then $\{\tilde{\psi}_s, s \in [0,t]\}$ satisfies the following forward SPDE, written here in weak form

$$d\left\langle \tilde{\psi}_s, \varphi \right\rangle = \left\langle \tilde{\psi}_s, L^\star \varphi \right\rangle ds + \left\langle \tilde{\psi}_s, \nabla^* (\varphi c) + h\varphi \right\rangle d\tilde{Y}_s \tag{2.2}$$

with $\tilde{\psi}_0$ having density ϕ with respect to the Lebesgue measure and $L^\star$ being the adjoint of L. Using Theorem 3.4 from [18], provided $a_{ij} \in C_b^2(\mathbb{R}^d)$, $b_i, c \in C_b^1(\mathbb{R}^d)$, $h \in C_b^0(\mathbb{R}^d)$ and $\phi \in W_2^0\left(\mathbb{R}^d\right)$, the SPDE (2.2) has a solution which is a measure valued process with square integrable density for all $t \geq 0$. In particular ψ_s belongs to $W_2^0\left(\mathbb{R}^d\right)$ for all $s \geq 0$. However we need here the solution of (2.1) to be a process with values in $C_b^2\left(\mathbb{R}^d\right)$. To achieve this, we show that $\psi_s \in W_2^m\left(\mathbb{R}^d\right)$ where m is chosen so that $2(m-2) > d$ and then use a standard Sobolev imbedding argument. To fix the ideas, in the following we will choose $m = \left[\frac{d}{2}\right] + 3$. We have the following

Lemma 2.1. *Suppose that the following condition on boundedness of the derivatives holds:*

(BD): a, b, c, h, $\phi \in C_b^{m+2}(\mathbb{R}^d)$ and $\phi \in W_2^m(\mathbb{R}^d)$.

Then there exists a constant K_1 independent of ϕ and $s \in [0,t]$ such that

$$\mathbb{E}[\|\psi_s\|_{m,2}^2] \leq K_1 \|\phi\|_{m,2}^2 \tag{2.3}$$

As a consequence $\psi_s \in C_b^2\left(\mathbb{R}^d\right)$ and there exists a constant K independent of ϕ and $s \in [0,t]$ such that

$$\mathbb{E}[\|\psi_s\|_{2,\infty}^2] \leq K_1 \|\phi\|_{m,2}^2.$$

Proof. The bound on $\mathbb{E}[\|\psi_s\|_{0,2}^2]$ follows from the same arguments as in [18]. Next, we differentiate (smoothing out by a Brownian semigroup as in [18] if necessary) both sides of (2.2). For simplicity of notations, we assume $d = 1$. Then $\tilde{\psi}_s^1 \equiv \nabla \tilde{\psi}_s$ satisfies the following SPDE

$$d\tilde{\psi}_s^1 = L_1 \tilde{\psi}_s^1 ds + \left(\nabla^* \tilde{\psi}_s^1 c + c_1 \tilde{\psi}_s^1 + c_2 \tilde{\psi}\right) d\tilde{Y}_s$$

with initial $\nabla\phi$, where L_1 is a second order differential operator with bounded coefficients, c_i are bounded functions. Similar to the arguments as in [18] we can prove that

$$\mathbb{E}[\|\psi_s^1\|_{0,2}^2] \leq K_1 \|\phi\|_{1,2}^2$$

The higher derivative estimates follow by induction. The last inequality follows from the Sobolev's imbedding theorem. $\square$

Let $k\delta \le t < (k+1)\delta$. Note that

$$\begin{aligned}
\langle V_t^n, \phi\rangle - \langle V_0^n, \psi_0\rangle &= \langle V_t^n, \psi_t\rangle - \langle V_{k\delta}^n, \psi_{k\delta}\rangle \\
&\quad + \sum_{j=1}^{k} \left(\left\langle V_{j\delta}^n, \psi_{j\delta}\right\rangle - \mathbb{E}\left(\left\langle V_{j\delta}^n, \psi_{j\delta}\right\rangle |\mathcal{F}_{j\delta-} \vee \mathcal{G}_{j\delta,t}\right)\right) \\
&\quad + \sum_{j=1}^{k} \left(\mathbb{E}\left(\left\langle V_{j\delta}^n, \psi_{j\delta}\right\rangle |\mathcal{F}_{j\delta-} \vee \mathcal{G}_{j\delta,t}\right) - \left\langle V_{(j-1)\delta}^n, \psi_{(j-1)\delta}\right\rangle\right) \\
&\equiv I_1^n + I_2^n + I_3^n, \qquad (2.4)
\end{aligned}$$

where $\mathcal{G}_{j\delta,t} = \sigma\left(Y_s - Y_t :\ j\delta \le s \le t\right)$. Then

$$\begin{aligned}
I_1^n &= \eta_{k\delta}^n \frac{1}{n}\sum_{i=1}^{m_k^n}\left(M_k^n(X^i,t)\psi_t(X_t^i) - \psi_{k\delta}(X_{k\delta}^i)\right), \\
I_2^n &= \sum_{j=1}^{k}\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)(\xi_j^i - \tilde{M}_j^n(X^i)) \\
I_3^n &= \sum_{j=1}^{k}\left(\eta_{j\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)\tilde{M}_j^n(X^i) - \eta_{(j-1)\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\psi_{(j-1)\delta}(X_{(j-1)\delta}^i)\right) \\
&= \sum_{j=1}^{k}\eta_{(j-1)\delta}^n\frac{1}{n}\sum_{i=1}^{m_{j-1}^n}\left(\psi_{j\delta}(X_{j\delta}^i)M_j^n(X^i) - \psi_{(j-1)\delta}(X_{(j-1)\delta}^i)\right).
\end{aligned}$$

The following lemma can be proved by adapting the argument in [4]. We leave the details to the reader.

Lemma 2.2.

$$\psi_{(j+1)\delta}(X_{(j+1)\delta}^i)M_{j+1}^n(X^i) - \psi_{j\delta}(X_{j\delta}^i) = \int_{j\delta}^{(j+1)\delta} M_j^n(X^i,s)\nabla^*\psi_s\sigma(X_s^i)dB_s^i. \qquad (2.5)$$

By replacing $(j+1)\delta$ and $j\delta$ by t and 0 respectively, in (2.5) i.e., take $j=0$ and $\delta = t$, we also get that

$$\begin{aligned}
\langle V_t, \phi\rangle &= \mathbb{E}\left(\phi(X_t)\exp\left(\int_0^t h(X_s)dY_s - \frac{1}{2}\int_0^t |h(X_s)|_H^2 ds\right)|\mathcal{G}_t\right) \\
&= \mathbb{E}(\psi_0(X_0)|\mathcal{G}_t) = \langle \pi_0, \psi_0\rangle.
\end{aligned}$$

The following theorem establishes the rates of convergence of the approximating filter to the optimal one. For this we need to assume the following initial condition of V^n (valid, for example, if V_0^n consists of n independent samples from π_0).

(I): $\mathbb{E}|\left\langle V_0^n, \phi\right\rangle - \langle \pi_0, \phi\rangle|^2 \le K_2 n^{-1}||\phi||_{0,\infty}^2$ *and* $\phi \in C_b^0(\mathbb{R}^d)$.

Theorem 2.3. *Suppose that* $\phi \in C_b^{m+2}(\mathbb{R}^d) \cap W_2^m(\mathbb{R}^d)$ *with* $m = \left[\frac{d}{2}\right] + 3$, *and the conditions (BD) and (I) hold. Then there exists a constant* K_3, *independent of* ϕ, *such that*

$$\mathbb{E}|\langle V_t^n,\phi\rangle-\langle V_t,\phi\rangle|^2\le K_3 n^{-(1-\alpha)}||\phi||_{m,2}^2.$$

Proof. First let us note that

$$\langle V_t^n,\phi\rangle-\langle V_t,\phi\rangle=I_1^n+I_2^n+I_3^n+(\langle V_0^n,\psi_0\rangle-\langle\pi_0,\psi_0\rangle)$$

Since the control of the last term is immediate from (BD) and (I) it only remains to control I_1^n, I_2^n and I_3^n. Via a straightforward argument similar to the one in [5], one shows that there exists a constant K_2, independent of ϕ such that

$$\mathbb{E}((I_3^n)^2)\le K_2 n^{-2}\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right)||\phi||_{m,2}^2.$$

Note that

$$\begin{aligned}
\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right) &= \mathbb{E}\mathbb{E}\left(\left(m_j^n(\eta_{j\delta}^n)^2\right)\Big|\mathcal{F}_{j\delta-}\right)\\
&= \mathbb{E}\left(m_{j-1}^n(\eta_{j\delta}^n)^2\right)\\
&= \mathbb{E}\left(m_{j-1}^n(\eta_{(j-1)\delta}^n)^2\mathbb{E}\left(\left(\eta_{j\delta}^n/\eta_{(j-1)\delta}^n\right)^2|\mathcal{F}_{(j-1)\delta}\right)\right)\\
&\le e^{K^2\delta}\mathbb{E}\left(m_{j-1}^n(\eta_{(j-1)\delta}^n)^2\right)
\end{aligned}$$

where the last inequality follows from

$$\begin{aligned}
&\mathbb{E}\left(\left(\frac{1}{m_{j-1}^n}\sum_{k=1}^{m_{j-1}^n}M_j^n(X^k)\right)^2|\mathcal{F}_{(j-1)\delta}\right)\\
\le\ &\frac{1}{m_{j-1}^n}\sum_{k=1}^{m_{j-1}^n}\mathbb{E}\left(M_j^n(X^k)^2|\mathcal{F}_{(j-1)\delta}\right)\le e^{K^2\delta}.
\end{aligned}$$

By induction, we have $\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right)\le e^{K^2T}n$. Hence, there exists a constant K, independent of ϕ such that

$$\mathbb{E}((I_3^n)^2)\le Kn^{-1}||\phi||_{m,2}. \tag{2.6}$$

One also shows that

$$\mathbb{E}((I_2^n)^2)\le\sum_{j=1}^{k}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\psi_{j\delta}(X_{j\delta}^i)^2\gamma_j^n(X^i)(\eta_{j\delta}^n)^2.$$

which implies, similarly as in [5], that there exists a constant K, independent of ϕ such that

$$\mathbb{E}((I_2^n)^2)\le Kn^{-(1-\alpha)}||\phi||_{m,2}.$$

The result follows after estimating I_1^n in a similar manner as I_3^n. $\square$

In a similar manner one can treat the approximation $\tilde{\pi}^n$ as defined in (1.4). We define $\tilde{V}_t^n = \tilde{\pi}_t^n \eta_t^n$ and can write that

$$\begin{aligned}\left\langle \tilde{V}_t^n, \psi_t \right\rangle - \left\langle \tilde{V}_{k\delta}^n, \psi_{k\delta} \right\rangle &= \frac{1}{n} \sum_{i=1}^{m_k^n} \left(\psi_t(X_t^i) M_k^n(X^i, t) - \psi_{k\delta}(X_{k\delta}^i) \right) \eta_{k\delta}^n \\ &\quad + \frac{1}{n} \sum_{i=1}^{m_k^n} \phi_t(X_t^i) \left(1 - M_k^n(X^i, t) \right) \eta_{k\delta}^n .\end{aligned}$$

It can be proved that the second moment of the second term is bounded by $Kn^{-2\alpha}||\phi||_{m,2}^2$. Therefore, we have

$$\mathbb{E}| < \tilde{V}_t^n, \phi > - < V_t, \phi > |^2 \leq K \left(n^{-(1-\alpha)} \vee n^{-2\alpha} \right) ||\phi||_{m,2}^2 .$$

The same inequality holds for the approximation $\hat{V}^n$ as defined in (1.5).

3. Convergence of V^n

In this section, we study the convergence of V^n, regarding as a sequence of stochastic processes. More specifically, we derive the convergence rate uniformly for t in an interval. First we observe that

$$\begin{aligned}\langle V_t^n, f \rangle &= \langle V_0^n, f \rangle + \int_0^t \langle V_s^n, Lf \rangle \, ds + \int_0^t \langle V_s^n, \nabla^* f c + hf \rangle \, dY_s \\ &\quad + N_t^{n,f} + \hat{N}_t^{n,f} ,\end{aligned} \tag{3.1}$$

where

$$\begin{aligned}N_t^{n,f} &= \sum_{j=0}^{[t/\delta]} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta)\wedge t} \nabla^* f \sigma(X_s^i) dB_s^i \eta_{j\delta}^n \\ \hat{N}_t^{n,f} &= \sum_{j=1}^{[t/\delta]} \eta_{j\delta}^n \frac{1}{n} \sum_{i=1}^{m_{j-1}^n} (\xi_j^i - \tilde{M}_j^n(X^i)) f(X_{j\delta}^i).\end{aligned}$$

$N^{n,f}$ and $\hat{N}^{n,f}$ are uncorrelated martingales with quadratic variation processes

$$\begin{aligned}\left\langle N^{n,f} \right\rangle_t &= \sum_{j=0}^{[t/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta)\wedge t} |\nabla^* f \sigma(X_s^i)|^2 ds (\eta_{j\delta}^n)^2 \\ \left\langle \hat{N}^{n,f} \right\rangle_t &= \sum_{j=1}^{[t/\delta]} \frac{1}{n^2} \mathbb{E} \left(\left(\sum_{i=1}^{m_{j-1}^n} (\xi_j^i - M_j^n(X^i)) f(X_{j\delta}^i) \right)^2 | \mathcal{F}_{j\delta-} \right) (\eta_{j\delta}^n)^2 \\ &= \sum_{j=1}^{[t/\delta]} \frac{1}{n^2} \sum_{i=1}^{m_{j-1}^n} \gamma_j^n(X^i) f^2(X_{j\delta}^i) (\eta_{j\delta}^n)^2 .\end{aligned} \tag{3.2}$$

We need the following technical estimate.

Lemma 3.1. *For any $f \in C_b^2(\mathbb{R}^d)$ with $||f||_{2,\infty} \le 1$, there exists a constant K_4, independent of f such that*

$$\left|\mathbb{E}\left(\gamma^n_{j+1}(X^i)f^2(X^i_{(j+1)\delta})(\eta^n_{(j+1)\delta}/\eta^n_{j\delta})^2|\mathcal{F}_{j\delta}\right)\right.$$
$$\left.-\sqrt{2\pi^{-1}}|h(X^i_{j\delta})-\bar{h}_{j\delta}|_H f^2(X^i_{j\delta})\sqrt{\delta}\right| \quad \le \quad K_4\delta,$$

where $\bar{h}_r = \frac{1}{m^n_j}\sum_{k=1}^{m^n_j}\tilde{M}^n_j(X^k,r)h(X^k_r)$.

Proof. Let $\hat{M}^n_j(t) = \frac{1}{m^n_j}\sum_{k=1}^{m^n_j} M^n_j(X^k,t)$. Then $\eta^n_{(j+1)\delta}/\eta^n_{j\delta} = \hat{M}^n_j((j+1)\delta)$ and by (20) in [9],

$$\tilde{M}^n_j(X^i,t) = 1 + \int_{j\delta}^t \tilde{M}^n_j(X^i,r)\left\langle h(X^i_r)-\bar{h}_r, dY_r - \bar{h}_r dr\right\rangle_H.$$

Similarly, we can prove that

$$d\hat{M}^n_j(t)^2 = 2\hat{M}^n_j(t)^2\bar{h}_t dY_t + \hat{M}^n_j(t)^2|\bar{h}_t|^2dt.$$

By Itô's formula, we get

$$\begin{aligned} d\left(\hat{M}^n_j(t)^2 f^2(X^i_t)\right) &= \hat{M}^n_j(t)^2\left(|\bar{h}_t|^2 f^2(X^i_t) + Lf^2(X^i_t) + 2\nabla^* f^2 c\bar{h}_t\right)dt \\ &\quad +\hat{M}^n_j(t)^2\left(2f^2(X^i_t)\bar{h}_t + \nabla^* f^2 c(X^i_t)\right)dY_t \\ &\quad +\hat{M}^n_j(t)^2\nabla^* f^2\sigma(X^i_t)dB_t. \end{aligned}$$

Now we adapt the argument in the proof of Proposition 6 in [3]. Let $F(x) = \{x\}(1-\{x\})$. By Itô's formula, we have

$$\begin{aligned} &\gamma^n_{j+1}(X^i)(\eta^n_{(j+1)\delta}/\eta^n_{j\delta})^2 f^2(X^i_{(j+1)\delta}) \\ =\ & \int_{j\delta}^{(j+1)\delta} \hat{M}^n_j(t)^2 f^2(X^i_t)D^-F(\tilde{M}^n_j(X^i,t))\tilde{M}^n_j(X^i,t) \\ &\qquad \times\left\langle h(X^i_t)-\bar{h}_t, dY_t-\bar{h}_t dt\right\rangle_H \\ &-\int_{j\delta}^{(j+1)\delta}\hat{M}^n_j(t)^2 f^2(X^i_t)|\tilde{M}^n_j(X^i,t)|^2|h(X^i_t)-\bar{h}_t|^2_H dt \\ &+2\sum_{k\ge1}\int_{j\delta}^{(j+1)\delta}\hat{M}^n_j(t)^2 f^2(X^i_t)dL_t(k) \\ &+\int_{j\delta}^{(j+1)\delta}F(\tilde{M}^n_j(X^i,t))\hat{M}^n_j(t)^2\left(|\bar{h}_t|^2 f^2(X^i_t)+Lf^2(X^i_t)+2\nabla^* f^2 c\bar{h}_t\right)dt \\ &+\int_{j\delta}^{(j+1)\delta}F(\tilde{M}^n_j(X^i,t))\hat{M}^n_j(t)^2\left(2f^2(X^i_t)\bar{h}_t+\nabla^* f^2 c(X^i_t)\right)dY_t \\ &+\int_{j\delta}^{(j+1)\delta}F(\tilde{M}^n_j(X^i,t))\hat{M}^n_j(t)^2\nabla^* f^2\sigma(X^i_t)dB_t \\ &+\int_{j\delta}^{(j+1)\delta}\hat{M}^n_j(t)^2 f^2(X^i_t)D^-F(\tilde{M}^n_j(X^i,t))\tilde{M}^n_j(X^i,t) \end{aligned}$$

$$\begin{aligned}
&\times \left\langle h(X_t^i)-\bar h_t, 2f^2(X_t^i)\bar h_t+\nabla^* f^2 c(X_t^i)\right\rangle_H dt \\
\equiv\ & \sum_{i=1}^7 I_i
\end{aligned}$$

where D^-F is the left derivative which is bounded by 1, and $L_t(k)$ is the local time at k for the semimartingale $\tilde M_j^n(X^i,t),\ \ j\delta\le t\le (j+1)\delta$. Note that

$$\begin{aligned}
\mathbb{E}\left(I_1|\mathcal{F}_{j\delta}\right) &= \mathbb{E}\Big(\int_{j\delta}^{(j+1)\delta} \hat M_j^n(t)^2 f^2(X_t^i) D^-F(\tilde M_j^n(X^i,t)) \\
&\qquad \times \tilde M_j^n(X^i,t)\left\langle h(X_t^i)-\bar h_t, -\bar h_t dt\right\rangle_H \Big| \mathcal{F}_{j\delta}\Big) \\
&\le \mathbb{E}\left(\int_{j\delta}^{(j+1)\delta} \hat M_j^n(t)^2 \|f^2\|_{0,\infty} \tilde M_j^n(X^i,t) 2\|h\|_\infty^2 dt \Big| \mathcal{F}_{j\delta}\right) \le K\delta.
\end{aligned}$$

Similarly, we can prove that $\mathbb{E}\left(I_4+I_7|\mathcal{F}_{j\delta}\right)\le K\delta$. Further, we have that $I_2\le 0$ and that $\mathbb{E}\left(I_5+I_6|\mathcal{F}_{j\delta}\right)=0$. Thus, we only need to deal with I_3. Similar to the proof of Proposition 6 in [3], we can show that

$$\mathbb{E}\left(\sum_{k\ge 2}\int_{j\delta}^{(j+1)\delta} \hat M_j^n(t)^2 f^2(X_t^i) dL_t(k)\Big|\mathcal{F}_{j\delta}\right)\le K\delta.$$

Thus we only need to deal with the first term in the sum for I_3. Note that

$$\begin{aligned}
&\int_{j\delta}^{(j+1)\delta} \hat M_j^n(t)^2 f^2(X_t^i) dL_t(1) - L_{(j+1)\delta}(1) f^2(X_{j\delta}^i) \\
=\ & \int_{j\delta}^{(j+1)\delta}\left(\hat M_j^n(t)^2 f^2(X_t^i) - f^2(X_{j\delta}^i)\right) dL_t(1).
\end{aligned}$$

It is easy to prove that

$$\mathbb{E}\left(\sup_{j\delta\le t\le (j+1)\delta}\left|\hat M_j^n(t)^2 f^2(X_t^i)-f^2(X_{j\delta}^i)\right|^2 \Big|\mathcal{F}_{j\delta}\right)\le K\delta$$

and that $\mathbb{E}L_{(j+1)\delta}(1)^2\le K\delta.$. Hence,

$$\mathbb{E}\left(\left|\int_{j\delta}^{(j+1)\delta}\hat M_j^n(t)^2 f^2(X_t^i) dL_t(1) - L_{(j+1)\delta}(1) f^2(X_{j\delta}^i)\right| \Big|\mathcal{F}_{j\delta}\right)\le K\delta.$$

By Itô's formula, we have

$$\begin{aligned}
&\left|\int_{j\delta}^{(j+1)\delta}\tilde M_j^n(X^i,r)\left\langle h(X_r^i)-\bar h_r, dY_r-\bar h_r dr\right\rangle_H\right| \\
=\ & \left|\tilde M_j^n(X^i)-1\right| \\
=\ & \int_{j\delta}^{(j+1)\delta} sgn\left(\tilde M_j^n(X^i,t)-1\right)\tilde M_j^n(X^i,t)\left\langle h(X_r^i)-\bar h_r, dY_r-\bar h_r dr\right\rangle_H \\
& +2L_{(j+1)\delta}(1).
\end{aligned}$$

Thus,

$$\begin{aligned}
& \mathbb{E}\left(2L_{(j+1)\delta}(1)|\mathcal{F}_{j\delta}\right) \\
= & \mathbb{E}\left(\left|\int_{j\delta}^{(j+1)\delta}\tilde{M}_j^n(X^i,r)\left\langle h(X_r^i)-\bar{h}_r, dY_r-h(X_t)dt\right\rangle_H\right|\Big|\mathcal{F}_{j\delta}\right)+O(\delta) \\
= & \mathbb{E}\left(\sqrt{2\pi^{-1}}\sqrt{\int_{j\delta}^{(j+1)\delta}|h(X_r^i)-\bar{h}_r|_H^2 dr}\Big|\mathcal{F}_{j\delta}\right)+O(\delta) \\
= & \sqrt{2\pi^{-1}}|h(X_{j\delta}^i)-\bar{h}_{j\delta}|_H\sqrt{\delta}+O(\delta).
\end{aligned}$$

Here $O(\delta)$ represents a term which is bounded by $K\delta$ for K being a deterministic constant. □

We define now the following distance on the space of finite measures

$$d(\nu_1,\nu_2)=\sum_{k=0}^{\infty}2^{-k}\left(|\langle\nu_1-\nu_2,f_k\rangle|\wedge 1\right)$$

where $f_0=1$ and for $k\geq 1$, $f_k\in C_b^{m+4}(\mathbb{R}^d)\cap W_2^{m+2}(\mathbb{R}^d)$ with $||f_k||_{m+4,\infty}\leq 1$ and also $||f_k||_{2,m+2}\leq 1$.

Theorem 3.2. *Suppose that the conditions (BD) and (I) hold true and, additionaly, that $h\in C_b^{m+2}(\mathbb{R}^d)\cap W_2^m(\mathbb{R}^d)$. Then, there exists a constant K_5 such that*

$$\mathbb{E}\sup_{t\leq T}d(V_t^n,V_t)^2\leq K_5 n^{-(1-\alpha)}.$$

Proof. Note that

$$\mathbb{E}\sup_{t\leq T}d(V_t^n,V_t)^2\leq\sum_{k=1}^{\infty}2^{-k}\left(\mathbb{E}\sup_{t\leq T}\langle V_t^n-V_t,f_k\rangle^2\wedge 1\right)+\mathbb{E}\sup_{t\leq T}\langle V_t^n-V_t,1\rangle^2 \quad (3.3)$$

and

$$\begin{aligned}
\mathbb{E}\sup_{t\leq T}\langle V_t^n-V_t,f\rangle^2 \leq\ & K\mathbb{E}\langle V_0^n-V_0,f\rangle^2+K\int_0^T\mathbb{E}\langle V_t^n-V_t,Lf\rangle^2 dt \\
& +K\int_0^T\mathbb{E}\left|\langle V_t^n-V_t,\nabla^* fc+hf\rangle\right|_H^2 dt \\
& +K\mathbb{E}\sum_{j=0}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_j^n}\int_{j\delta}^{((j+1)\delta)\wedge t}|\nabla^* f\sigma(X_s^i)|^2 ds(\eta_{j\delta}^n)^2 \\
& +K\mathbb{E}\sum_{j=1}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m_{j-1}^n}\gamma_j^n(X^i)f^2(X_{j\delta}^i)(\eta_{j\delta}^n)^2. \quad (3.4)
\end{aligned}$$

Following condition (I), the first term is bounded by Kn^{-1}. Next, by Theorem 2.3, we see that the following two terms are bounded by $Kn^{-(1-\alpha)}$. Note that

$$4\text{th term}\leq K\sum_{j=1}^{[T/\delta]}\frac{\delta}{n^2}\mathbb{E}\left(m_j^n(\eta_{j\delta}^n)^2\right)\leq Kn^{-1}.$$

By Lemma 3.1, we have

$$\begin{aligned}
\text{5th term} &= K\mathbb{E}\sum_{j=1}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}}\mathbb{E}\left(\gamma^n_j(X^i)f^2(X^i_{j\delta})(\eta^n_{j\delta})^2\middle|\mathcal{F}_{(j-1)\delta}\right)\\
&\le K\sum_{j=0}^{[T/\delta]}\frac{\sqrt{\delta}}{n^2}\mathbb{E}\left(m^n_j(\eta^n_{j\delta})^2\right)\le Kn^{-(1-\alpha)}.
\end{aligned}$$

To complete the proof we consider the last term in (3.3). Taking $f=1$ in (3.4), we get

$$\begin{aligned}
\mathbb{E}\sup_{t\le T}\langle V^n_t-V_t,1\rangle^2 &\le K\int_0^T\mathbb{E}\left|\langle V^n_t-V_t,h\rangle\right|^2_H dt\\
&\quad+K\mathbb{E}\sum_{j=1}^{[T/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}}\gamma^n_j(X^i)(\eta^n_{j\delta})^2. \qquad (3.5)
\end{aligned}$$

It is clear that Lemma 3.1 remains true with $f=1$, and hence, the second term of (3.5) is bounded by $Kn^{-(1-\alpha)}$. By Theorem 2.3, we get that the first term of (3.5) is bounded by $Kn^{-(1-\alpha)}$. The conclusion then follows by plugging all the above estimates back into (3.3). □

Remark 3.3. For the case of $\tilde{\pi}^n_t$, the jump at $(j+1)\delta$ is

$$\eta^n_{(j+1)\delta}\frac{1}{n}\sum_{i=1}^{m^n_j}\left(\xi^i_{j+1}-\frac{\eta^n_{j\delta}}{\eta^n_{(j+1)\delta}}\right)\delta_{X^i_{(j+1)\delta}}.$$

Write

$$\xi^i_{j+1}-\frac{\eta^n_{j\delta}}{\eta^n_{(j+1)\delta}}=\left(\xi^i_{j+1}-\tilde{M}^n_j(X^i)\right)+\frac{M^n_j(X^i)-1}{\frac{1}{m^n_{j-1}}\sum_{k=1}^{m^n_{j-1}}M^n_j(X^k)}.$$

Then the new $\hat{N}^{n,f}$ can be written as two terms. A careful estimate of the second term leads to the bound $Kn^{-2\alpha}$. Thus, we have

$$\mathbb{E}\sup_{t\le T}d(\tilde{V}^n_t,V_t)^2\le K\left(n^{-2\alpha}\vee n^{-(1-\alpha)}\right).$$

The same inequality holds for $\hat{V}^n_t$.

4. A central limit type theorem

In this section, we prove the exact rate of convergence by a central limit type theorem. For $\alpha\in(0,1)$, let

$$U^n_t=n^{\frac{1-\alpha}{2}}(V^n_t-V_t),\qquad t\ge 0.$$

By (3.1) and Zakai equation, we have

$$\begin{aligned}
\langle U^n_t,f\rangle &= \langle U^n_0,f\rangle+\int_0^t\langle U^n_s,Lf\rangle ds+\int_0^t\langle U^n_s,\nabla^*fc+hf\rangle dY_s\\
&\quad+n^{\frac{1-\alpha}{2}}N^{n,f}_t+n^{\frac{1-\alpha}{2}}\hat{N}^{n,f}_t, \qquad (4.1)
\end{aligned}$$

Let $\Phi' = \cup_{k=0}^{\infty}\Phi_{-k}$ be the dual of the nuclear space Φ defined on page 333 in [20].

Theorem 4.1. *There exists κ such that $\{U^n\}$ is tight in $D_{\Phi_{-\kappa}}[0,\infty)$.*

Proof. For $u \le \epsilon$, we have

$$\mathbb{E}\left(\left\langle U^n_{t+u} - U^n_t, f\right\rangle^2 \middle| \mathcal{F}_t\right) \le \mathbb{E}\left(\sum_{i=1}^4 \zeta_f^{i,n}(\epsilon)\right)$$

where $\zeta_f^{1,n}(\epsilon) = \int_t^{t+\epsilon} \langle U^n_s, Lf\rangle^2 ds$, $\zeta_f^{2,n}(\epsilon) = \int_t^{t+\epsilon} \langle U^n_s, \nabla^* fc + hf\rangle^2 ds$ and

$$\begin{aligned}
\zeta_f^{3,n}(\epsilon) &= n^{1-\alpha} \sum_{t\le j\delta<t+\epsilon} \frac{1}{n^2}\sum_{i=1}^{m_j^n} \int_{j\delta}^{((j+1)\delta)\wedge t} |\nabla^* f\sigma(X^i_s)|^2 ds (\eta^n_{j\delta})^2 \\
\zeta_f^{4,n}(\epsilon) &= n^{1-\alpha} \sum_{t\le j\delta<t+\epsilon} \frac{1}{n^2}\sum_{i=1}^{m_j^n} \gamma_j^2(X^i) f^2(X^i_{j\delta})(\eta^n_{j\delta})^2.
\end{aligned}$$

Similar to the previous section, we can show that

$$\lim_{\epsilon\to 0}\sup_n \mathbb{E}\left(\sum_{i=1}^4 \zeta_f^{i,n}(\epsilon)\right) = 0.$$

As in the proof of Theorem 3.2 there exists a constant K such that

$$\mathbb{E}\sup_{t\le T}\langle U^n_t - U^n_0, f\rangle^2 \le K,$$

which implies the compact containment for $\{\langle U^n_t, f\rangle :\ n \ge 1,\ t \ge 0\}$. By Remark 8.7 (p138) in Ethier and Kurtz [13], we get the tightness of $\langle U^n, f\rangle$ in $D_{\mathbb{R}}[0,\infty)$. As in the proof of Theorem 3.1 in Kurtz and Xiong [20], applying Mitoma's theorem, we get the tightness of U^n in $D_{\Phi_{-\kappa}}[0,\infty)$. $\square$

It is easy to show that $n^{\frac{1-\alpha}{2}} N_t^{n,f} \to 0$. On the other hand,

$$\begin{aligned}
\left\langle n^{(1-\alpha)/2}\hat{N}^{n,f}\right\rangle_t &= n^{1-\alpha}\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}} \gamma^n_j(X^i) f^2(X^i_{j\delta})(\eta^n_{j\delta})^2 \\
&= n^{1-\alpha}\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}} \mathbb{E}\left(\gamma^n_j(X^i) f^2(X^i_{j\delta})(\eta^n_{j\delta})^2 \middle| \mathcal{F}_{(j-1)\delta}\right) \\
&\quad + n^{1-\alpha}\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}} \Bigg(\gamma^n_j(X^i) f^2(X^i_{j\delta})(\eta^n_{j\delta})^2 \\
&\qquad - \mathbb{E}\left(\gamma^n_j(X^i) f^2(X^i_{j\delta})(\eta^n_{j\delta})^2 \middle| \mathcal{F}_{(j-1)\delta}\right)\Bigg). \qquad (4.2)
\end{aligned}$$

By Lemma 3.1, the first term satisfies

$$\lim_{n\to\infty} n^{1-\alpha}\sum_{j=1}^{[t/\delta]}\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}}\sqrt{\frac{2}{\pi}}|h(X^i_{(j-1)\delta})-\bar{h}_{(j-1)\delta}|_H\sqrt{\delta}f^2(X^i_{(j-1)\delta})(\eta^n_{(j-1)\delta})^2$$
$$=\lim_{n\to\infty}\sqrt{\frac{2}{\pi}}\sum_{j=0}^{[t/\delta]-1}\left\langle V^n_{j\delta},|h-\bar{h}_{j\delta}|_H f^2\right\rangle \delta\left\langle V^n_{j\delta},1\right\rangle$$
$$=\sqrt{\frac{2}{\pi}}\int_0^t\left\langle V_s,|h-\pi_s h|_H f^2\right\rangle\langle V_s,1\rangle\, ds.$$

Note that $\gamma^n_j(X^i)^2\le\gamma^n_j(X^i)$ and

$$d\hat{M}^n_j(t)^4=4\hat{M}^n_j(t)^4\bar{h}_t dY_t+6\hat{M}^n_j(t)^4|\bar{h}_t|^2dt.$$

Similar to Lemma 3.1, we have

$$\mathbb{E}\left(\gamma^n_j(X^i)^2\left(\eta^n_{j\delta}/\eta^n_{(j-1)\delta}\right)^4\middle|\mathcal{F}_{(j-1)\delta}\right)\le K\sqrt{\delta}.$$

Hence, the second moment of the second term on the right hand side of (4.2) is bounded by

$$\begin{aligned}
&n^{2(1-\alpha)}\sum_{j=1}^{[t/\delta]}\mathbb{E}\left(\left(\frac{1}{n^2}\sum_{i=1}^{m^n_{j-1}}\gamma^n_j(X^i)f^2(X^i_{j\delta})(\eta^n_{j\delta})^2\right)^2\right)\\
\le\ &\|f\|^4_\infty n^{2(1-\alpha)}\sum_{j=1}^{[t/\delta]}\mathbb{E}\left(\left(\frac{m^n_{j-1}}{n^2}\right)^2\frac{1}{m^n_{j-1}}\sum_{i=1}^{m^n_{j-1}}\gamma^n_j(X^i)^2(\eta^n_{j\delta})^4\right)\\
\le\ &K_1 n^{-2-\alpha}\mathbb{E}\left((m^n_{j-1})^2(\eta^n_{(j-1)\delta})^4\right).
\end{aligned}$$

Finally, we estimate $\mathbb{E}\left((m^n_j)^2(\eta^n_{j\delta})^4\right)$ recursively as follows

$$\begin{aligned}
\mathbb{E}\left((m^n_j)^2(\eta^n_{j\delta})^4\right) &= \mathbb{E}\left(\mathbb{E}\left((m^n_j)^2(\eta^n_{j\delta})^4\middle|\mathcal{F}_{j\delta-}\right)\right)\\
&= \mathbb{E}\left((\eta^n_{j\delta})^4\mathbb{E}\left(\sum_{i=1}^{m^n_{j-1}}\xi^j_i\right)^2\middle|\mathcal{F}_{j\delta-}\right)\\
&\le \mathbb{E}\left((\eta^n_{j\delta})^4 m^n_{j-1}\sum_{i=1}^{m^n_{j-1}}\mathbb{E}\left((\xi^j_i)^2\right)\right)\\
&\le \mathbb{E}\left((\eta^n_{j\delta})^4(m^n_{j-1})^2(1+K\delta)\right)\\
&\le (1+K\delta)e^{2K^2\delta}\mathbb{E}\left((m^n_{j-1})^2(\eta^n_{(j-1)\delta})^4\right).
\end{aligned}$$

Thus, by induction, we have

$$\mathbb{E}\left((m^n_j)^2(\eta^n_{j\delta})^4\right)\le(1+K\delta)^j e^{2K^2 j\delta}n^2\le Kn^2.$$

Combining the estimates above, we get

Lemma 4.2. *As $n \to \infty$, we have*

$$n^{\frac{1-\alpha}{2}} N_t^{n,f} \to 0 \text{ and } n^{\frac{1-\alpha}{2}} \hat{N}_t^{n,f} \Longrightarrow M_t^f$$

which is a martingale uncorrelated to W and Y such that

$$\left\langle M_t^f \right\rangle_t = \sqrt{\frac{2}{\pi}} \int_0^t \left\langle V_s, |h - \pi_s h|_H f^2 \right\rangle \langle V_s, 1 \rangle \, ds.$$

Further, there exists a space-time white noise $B(dtdx)$ (independent of W and Y such that

$$M_t^f = \sqrt[4]{\frac{2}{\pi}} \int_0^t \int_{\mathbb{R}^d} \sqrt{|h(x) - \pi_s h|_H V(s,x) \langle V_s, 1 \rangle} f(x) B(dsdx),$$

where $V(s,x)$ is the density of the measure V_s.

Summarizing, we get

Theorem 4.3. *$U^n \Longrightarrow U$ which is the unique solution to*

$$\begin{aligned} \langle U_t, f \rangle &= \langle U_0, f \rangle + \int_0^t \langle U_s, Lf \rangle \, ds + \int_0^t \langle U_s, \nabla^* f c + hf \rangle \, dY_s \\ &\quad + \sqrt[4]{\frac{2}{\pi}} \int_0^t \int_{\mathbb{R}^d} \sqrt{|h(x) - \pi_s h|_H V(s,x) \langle V_s, 1 \rangle} f(x) B(dsdx). \quad (4.3) \end{aligned}$$

Proof. By Theorem 4.1, we can take U being a limit point. Without loss of generality, we assume that $U^n \Longrightarrow U$. By Lemma 4.2, it is easy to show that U satisfies (4.3). To prove the uniqueness, we take another solution $\tilde{U}$ of (4.3) and define $\hat{U}_t = U_t - \tilde{U}_t$. Then $\hat{U}_t$ satisfies the following homogeneous linear equation

$$\left\langle \hat{U}_t, f \right\rangle = \int_0^t \left\langle \hat{U}_s, Lf \right\rangle ds + \int_0^t \left\langle \hat{U}_s, \nabla^* f c + hf \right\rangle dY_s.$$

Similar to Lemma 4.2 in [20] we get $\hat{U} = 0$. ☐

Now we give some details on how to deal with the branching particle filter $\hat{V}_t^n$. In this case, there is an extra term $S_t^{n,f} = \sum_{i=1}^6 J_i^n(t)$ where

$$J_1^n(t) = \sum_{j=1}^{[t/\delta]-1} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{(j+1)\delta} M_j^n(X^i, s) h(X_s^i)(f(X_s^i) - f(X_{j\delta}^i)) dY_s,$$

$$J_2^n(t) = \sum_{j=1}^{[t/\delta]-1} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{(j+1)\delta} (M_j^n(X^i, s) - 1) Lf(X_s^i) ds,$$

$$J_3^n(t) = \sum_{j=1}^{[t/\delta]-1} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{(j+1)\delta} (M_j^n(X^i, s) - 1) \nabla^* f \sigma dB_s^i,$$

$$J_4^n(t) = \sum_{j=1}^{[t/\delta]-1} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{(j+1)\delta} (M_j^n(X^i, s) - 1) \nabla^* f c(X_s^i) dY_s,$$

$$\begin{aligned} J_5^n(t) &= \sum_{j=1}^{[t/\delta]-1} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{(j+1)\delta} M_j^n(X^i,s) \nabla^* f \left\langle c,h \right\rangle_H (X_s^i) ds \\ &\quad - \int_0^t \int_U \left\langle V_s^n, \nabla^* f \left\langle c,h \right\rangle_H \right\rangle ds \end{aligned}$$

and

$$J_6^n(t) = \sum_{j=1}^{[t/\delta]-1} \frac{1}{n} \sum_{i=1}^{m_j^n} \int_{j\delta}^{(j+1)\delta} M_{j+1}^n(X^i,s) h(X_s^i) dY_s f(X_{j\delta}^i) - \int_0^t \left\langle V_s^n, hf \right\rangle dY_s.$$

To deal with this term, we need the following

Lemma 4.4. *Let*

$$W_{k\ell}^\delta(t) = \int_0^t \sqrt{2\delta^{-1}} \left(Y_s^k - Y_{\eta_\delta(s)}^k \right) dY^\ell(s), \qquad k,\ell = 1,2,\cdots$$

where $Y^k(t) = \left\langle Y(t), e_k \right\rangle_H$ *and* $\{e_k\}$ *is a CONS of* H. *Then, as* $\delta \to 0$, $W_{k\ell}^\delta \to W_{k\ell}$ *and* $W_{k\ell},\ k,\ell \geq 1$ *are independent Brownian motions which are independent of* Y.

Proof. We adapt the proof of Lemma 5.2 in Kurtz and Xiong [20]. It is clear that $W_{k\ell}^\delta,\ k,\ell \geq 1$ are uncorrelated martingales and

$$\begin{aligned} [W_{k\ell}^\delta]_t &= \int_0^t 2\delta^{-1} \left(Y_s^k - Y_{\eta_\delta(s)}^k \right)^2 ds \\ &= 2t\delta^{-2} \int_0^\delta \delta t^{-1} \sum_{j=0}^{[t/\delta]-1} \left(\frac{Y_{s+j\delta}^k - Y_{j\delta}^k}{\sqrt{s}} \right)^2 s ds \to t. \end{aligned}$$

For $g \in H$, we have

$$[W_k^\delta, Y(g)]_t = \sqrt{2\delta^{-1}} \int_0^t \left(Y_s^k - Y_{\eta_\delta(s)}^k \right) ds \left\langle g, e_\ell \right\rangle_H \to 0$$

and the lemma follows by the martingale central limit theorem. □

We can define an $H \otimes H$-c.B.m $\tilde{W}_t$ by

$$\tilde{W}^{e_k \otimes e_\ell}(t) = W_{k\ell}(t).$$

Note that

$$\begin{aligned} f(X_s^i) - f(X_{j\delta}^i) &= \int_{j\delta}^s \tilde{L} f(X_r^i) dr + \int_{j\delta}^s \nabla^* f \sigma(X_r^i) dB_r^i \\ &\quad + \int_{j\delta}^s \nabla^* f c(X_r^i) dY_r. \end{aligned}$$

Approximate $X_r^i,\ j\delta \le r < (j+1)\delta$, by $X_{j\delta}^i$, we have

$$\begin{aligned} & n^{1/3}\sum_{j=1}^{[t/\delta]-1}\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta} M_j^n(X^i,s)h(X_s^i)\int_{j\delta}^s \nabla^* fc(X_r^i)dY_r dY_s \\ \approx\quad & n^{1/3}\sum_{j=1}^{[t/\delta]-1}\frac{1}{n}\sum_{i=1}^{m_j^n}\sum_{k,\ell=1}^{\infty}\left\langle h(X_{j\delta}^i),e_k\right\rangle_H \left\langle \nabla^* fc(X_{j\delta}^i),e_\ell\right\rangle_H \\ & \times\sqrt{\frac{\delta}{2}}\left(W_{k\ell}^\delta((j+1)\delta)-W_{k\ell}^\delta(j\delta)\right) \\ \to\quad & \frac{1}{\sqrt{2}}\int_0^t \langle V_s, h\otimes\nabla^* fc\rangle\, d\tilde{W}_s. \end{aligned}$$

Similarly, we can prove

$$n^{1/3}\sum_{j=1}^{[t/\delta]-1}\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta} M_j^n(X^i,s)h(X_s^i)\int_{j\delta}^s \tilde{L}f(X_r^i)drdY_s \to 0$$

and

$$n^{1/3}\sum_{j=1}^{[t/\delta]-1}\frac{1}{n}\sum_{i=1}^{m_j^n}\int_{j\delta}^{(j+1)\delta} M_j^n(X^i,s)h(X_s^i)\int_{j\delta}^s \nabla^* f\sigma(X_r^i)dB_r^i dY_s \to 0.$$

Hence

$$n^{1/3}J_1^n(t) \to \frac{1}{\sqrt{2}}\int_0^t \langle V_s, h\otimes\nabla^* fc\rangle\, d\tilde{W}_s.$$

By the same arguments, we can prove that

$$n^{1/3}J_i^n(t) \to \begin{cases} \frac{1}{\sqrt{2}}\int_0^t \langle V_s, h\otimes\nabla^* fc\rangle\, d\tilde{W}_s, & \text{if } i=4,\ 6, \\ 0, & \text{if } i=2,\ 3,\ 5. \end{cases}$$

¿From the discussion above, we get

Proposition 4.5.

$$n^{\frac{1}{3}}S_t^{n,f} \to S_t^f = \frac{3}{\sqrt{2}}\int_0^t \langle V_s, h\otimes\nabla^* fc\rangle\, d\tilde{W}_s.$$

Now we state the central limit theorem for $\alpha=\frac{1}{3}$.

Theorem 4.6. *For $\alpha=\frac{1}{3}$, we have $n^{\frac{1}{3}}\left(\hat{V}^n - V\right) \to U$ which is the unique solution to*

$$\begin{aligned} \langle U_t, f\rangle \quad = \quad & \langle U_0, f\rangle + \int_0^t \langle U_s, Lf\rangle\, ds + \int_0^t \langle U_s, \nabla^* fc + hf\rangle\, dY_s \\ & + \frac{3}{\sqrt{2}}\int_0^t \langle V_s, h\otimes\nabla^* fc\rangle\, d\tilde{W}_s \\ & + \int_0^t\int_{\mathbb{R}^d}(2\pi^{-1})^{1/4}\sqrt{|h(x)|_H V(s,x)}f(x)B(dsdx). \end{aligned}$$

Remark 4.7. If $\alpha > \frac{1}{3}$, then $n^{\frac{1-\alpha}{2}}(\hat{V}_t^n - V_t)$ converges to a non-trivial limit characterized by an equation above without the term next to last; if $\alpha < \frac{1}{3}$, then $n^{\alpha}(\hat{V}_t^n - V_t)$ converges to a non-trivial limit characterized by an equation above without the last term. Same result holds for $\tilde{V}$.

Finally, we convert the convergence result to that for the optimal filter.

Theorem 4.8. *$n^{\frac{1-\alpha}{2}}(\pi_t^n - \pi_t)$ converges weakly to a process ζ_t which is the unique solution to the following evolution equation:*

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

where $d\nu_t = dY_t - \pi_t(h)dt$ and $\pi(t,x)$ is the density of the probability measure π_t.

Proof. It is easy to show that

$$n^{\frac{1-\alpha}{2}}(\pi_t^n - \pi_t) = (V_t 1)^{-1} U_t^n - (V_t^n 1 V_t 1)^{-1} U_t^n 1 V_t^n$$

which converges to $\zeta_t \equiv (V_t 1)^{-1}\left(U_t - (V_t 1)^{-1} U_t 1 V_t\right)$. Let $\eta_t = (V_t 1)^{-1} U_t$. By Itô's formula, we have

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

By applying Itô's formula again, we get the equation for ζ. □

Acknowledgment: Most of this work was done while the second author visited Imperial College. The hospitality of Imperial College and financial support from London Mathematical Society are gratefully acknowledged.

References

1. Bensoussan, A.: *Stochastic control of partially observable systems.* Cambridge University Press, Cambridge, 1992.
2. Chopin, N.: Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann. Statist.* **32**, no. 6, (2004) 2385–2411.
3. Crisan, D.: Exact rates of convergence for a branching particle approximation to the solution of the Zakai equation. *Ann. Probab.* **31**, no. 2, (2003) 693–718.
4. Crisan, D.: Superprocesses in a Brownian environment. *Proc. R. Soc. Lond. A* **460**, (2004) 243-270.
5. Crisan, D.: Particle approximations for a class of stochastic partial differential equations. *To appear in the Applied Mathematics and Optimization Journal.* (2006).
6. Crisan, D., Del Moral, P. and Lyons, T.: Interacting particle systems approximations of the Kushner-Stratonovitch equation. *Adv. in Appl. Probab.* **31**, no. 3, (1999) 819–838.
7. Crisan, D. and Doucet, A.: A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.* **50**, no. 3, (2002) 736–746.
8. Crisan, D., Gaines, J. and Lyons, T.: Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Math.* **58**, no. 5, (1998) 1568–1590 (electronic).

9. Crisan, D. and Lyons, T.: A particle approximation of the solution of the Kushner-Stratonovitch equation. *Probab. Theory Related Fields* **115**, no. 4, (1999) 549–578.
10. Del Moral, P.: *Feynman-Kac formulae. Genealogical and interacting particle systems with applications.* Probability and its Applications, Springer-Verlag, New York, 2004.
11. Del Moral, P. and Guionnet, A.: Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.* **9**, no. 2, (1999) 275–297.
12. Del Moral, P. and Miclo, L.: 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 (2000) 1–145.
13. Ethier, S. N. and Kurtz, T. G.: *Markov processes : Characterization and convergence.* Wiley, New York, 1986.
14. Friedman, A.: *Stochastic Differential Equations and Applications.* Volume 1. Academic Press, 1975.
15. Gel'fand, I. M. and Vilenkin, N. Ya.: *Generalized functions. Vol. 4: Applications of harmonic analysis.* Academic Press, New York - London, 1964.
16. Kallianpur, G. and Xiong, J.: *Stochastic differential equations in infinite dimensional spaces.* IMS Lecture Notes-Monograph Series **26**, 1995.
17. Kunsch, H. R.: Recursive Monte Carlo filters: algorithms and theoretical analysis. *Ann. Statist.* **33**, no. 5, (2005) 1983–2021.
18. Kurtz, T. G. and Xiong, J.: Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications* **83**, (1999) 103-126.
19. Kurtz, T. G. and Xiong, J.: 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.* Birkhauser, (2000) 233-258.
20. Kurtz, T. G. and Xiong, J.: A stochastic evolution equation arising from the fluctuation of a class of interacting particle systems. *Communication Mathematical Sciences* **Vol. 2**, (2004) 325-358.
21. Rozovskii, B. L.: *Stochastic evolution systems. Linear theory and applications to nonlinear filtering.* Kluwer Academic Publishers Group, Dordrecht, 1990.

Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2BZ, UK.

E-mail address: `d.crisan@imperial.ac.uk`

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`