*The Annals of Probability*
2003, Vol. 31, No. 2, 693–718


# EXACT RATES OF CONVERGENCE FOR A BRANCHING PARTICLE APPROXIMATION TO THE SOLUTION OF THE ZAKAI EQUATION

BY DAN CRISAN

*Imperial College*

In Crisan, Gaines and Lyons [*SIAM J. Appl. Probab.* **58** (1998) 313–342] we describe a branching particle algorithm that produces a particle approximation to the solution of the Zakai equation and find an upper bound for the rate of convergence of the mean square error. In this paper, the exact rate of convergence of the mean square error is deduced. Also, several variations of the branching algorithm with better rates of convergence are introduced.

**1. Introduction.** Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space on which we define a pair of stochastic processes satisfying the following $(d+m)$-dimensional system of stochastic differential equations

$$dX_t = f(X_t)\,dt + \sigma(X_t)\,dV_t, \tag{1}$$

$$dY_t = h(X_t)\,dt + dW_t, \tag{2}$$

where $V = \{V_t, \mathcal{F}_t; t \geq 0\}$ and $W = \{W_t, \mathcal{F}_t; t \geq 0\}$ are mutually independent $n$-dimensional, respectively, $m$-dimensional standard Brownian motions. Let $\pi_t$ be the conditional distribution of the $X$ at time $t$ given the $\sigma$-field $\mathcal{Y}_t \triangleq \sigma(Y_s, 0 \leq s \leq t)$. Then $\pi_t$ is defined so that

$$\pi_t(\varphi) \triangleq \int_{\mathbb{R}^d} \varphi(x)\pi_t(dx) = E[\varphi(X_t)|\mathcal{Y}_t], \qquad P\text{-a.s.}$$

for any bounded measurable function $\varphi$. One can prove that $\pi_t$ has an unnormalized version that satisfies the linear stochastic partial differential equation (cf. [20]; see also [12] and [17])

$$p_t(\varphi) = \pi_0(\varphi) + \int_0^t p_s(A\varphi)\,ds + \int_0^t p_s(h^*\varphi)\,dY_s. \tag{3}$$

In (3), $h^*$ is the row vector $(h_1, \dots, h_m)$, $A$ is the infinitesimal generator associated to $X$ and $\varphi$ is an arbitrary continuous bounded function belonging to $\mathcal{D}(A)$, the domain of $A$.

In [4], we describe a system of moving and branching particles whose empirical distribution at time $t$, denoted by $U_N(t)$, converges almost surely to $p_t$. Further,

Received May 2001; revised March 2002.
*AMS 2000 subject classifications.* 93E11, 60G57, 65U05.
*Key words and phrases.* Zakai equation, branching algorithm, particle filters, filtering, Monte Carlo approximation.

we find the following upper bound for the rate of convergence of $U_N(t)$:

$$\tilde{E}\big[\big((U_N(t),\varphi) - p_t(\varphi)\big)^2\big] \le \frac{c(t)}{\sqrt{N}},$$

where $c(t)$ is a constant independent of $N$ and $\tilde{E}$ is the expectation with respect to a probability measure $\tilde{P}$ absolutely continuous with respect to $P$ [as defined in (6)]. Intuitively, the component particles explore the state space following the law of the signal. Successive branching steps are added to gradually incorporate the new information (the observation $Y$) into the system. The branching procedure also reduces the variance of the system by removing particles with unlikely positions given the accumulated information and multiplying those that stay on the right track, thus speeding up the convergence to $p_t$. However, in [4], we leave a number of questions unanswered.

First, based on numerical experiments, the branching particle approximation is superior to the Monte Carlo approximation, that is, to the approximation given by

$$\Theta_N(t) \triangleq \frac{1}{N}\sum_{i=1}^{N} \mu_i(V_i(t))\delta_{V_i(t)},$$

where $V_1(t), \dots, V_N(t)$ are independent realizations of the signal process $X_t$ and the weights $\mu_i(V_i(t))$ are defined as

$$\mu_i(V_i(t)) = \exp\biggl(\int_0^t h^*(V_i(s))\,dY_s - \tfrac{1}{2}\int_0^t \|h(V_i(s))\|^2\,ds\biggr).$$

Here, the accumulated information (the observation $Y$) is incorporated into the system only at the end of the procedure by attaching weights to the particles, weights that depends on $Y$. The motion of the particle is not influenced by the observation $Y$. It is universally accepted among practitioners (cf. e.g., [7]) that the (raw) Monte Carlo method is vastly improved when linked up with some sort of variance reduction procedure. The general term used for such a reduced-variance Monte Carlo method is that of a sequential Monte Carlo method or a particle filter. The algorithm based on the above construction is a member of this class of methods. It is easy to check that (see Section 5)

$$\tilde{E}\big[\big((\Theta_N(t),\varphi) - p_t(\varphi)\big)^2\big] = \frac{c_\Theta(t)}{N}, \tag{4}$$

where $c_\Theta(t)$ is a constant independent of $N$. Furthermore, in [14] and [15], the authors prove that an upper bound of the same order holds true in a very general setup. So, *apparently*, it seems that the Monte Carlo approximation offers, at least theoretically, better rates of convergence, *a fact that contradicts the numerical results*.

Second, in contrast to the continuous case, if the branching algorithm is set up in a discrete-time framework, then both methods (Monte Carlo method and branching method) have the same order of convergence (see [3]). Is there something that changes in the continuous-time setup and makes the branching method converge slower?

Third, the convergence proof in [4] does not say anything about the way in which the corrective branching procedure actually improves on the Monte Carlo method. In other words, the intuition does not get translated into the mathematics of the proofs at all. Moreover, the whole philosophy behind the proof in [4] is fallacious, as it represents the branching algorithm as a perturbation of the Monte Carlo approximation.

In this paper, we give a (partial) answer to the above questions. In the following, we determine the exact rates of convergence for the branching algorithm. That is, we prove that, if the length of the interbranching times is (of order) $1/N^{2\alpha}$, where $\frac{1}{3} < \alpha < 1$, then

$$\lim_{N\to\infty} N^{1-\alpha}\tilde{E}\big[\big((U_N(t),\varphi) - p_t(\varphi)\big)^2\big] = c(t). \tag{5}$$

Moreover, if $t$ is chosen to be a branching time for all $N$, then (5) holds true for $0 < \alpha < 1$. *Hence, the branching method will have a slower rate of convergence as long as the interbranching times are taken to converge to* 0.

However, if the interbranching times are fixed (independent of the number of particles), then the exact rate of convergence of the branching algorithm is indeed of order $1/N$, just as the rate of the Monte Carlo approximation, and it is a safe conjecture that $c_\Theta(t)$ increases exponentially faster than $c(t)$. In all numerical experiments, the interbranching times are fixed, the reason being that the observation $Y$ does not arrive in a continuous manner, but at discrete intervals and branchings occur only at these arrival times. By contrast, the number of particles can be increased as much as we want or, more likely, as much as the computer hardware constraints permit us to do so.

Finally, the analysis in Section 4 is fundamentally more refined than its equivalent in [4]. It hinges on the unexpected representation formula (23) of the variance of the branching mechanism in terms of the local time(s) of an exponential martingale.

The paper is structured as follows. In Section 2, the proper theoretical framework for the filtering problem is set up. Then, in Section 3, we review the construction of the branching particle system presented in [4] and state some preliminary results that are proved in [4]. In Section 4, we prove the main result of the paper, that is, the asymptotic rate of convergence of $U_N$, and in Section 5, we present several variations of the branching algorithm with improved rates of convergence.

**2. Filtering framework.** In the following, we will assume that $f:\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma:\mathbb{R}^d \to \mathcal{L}(\mathbb{R}^n,\mathbb{R}^d) \simeq \mathbb{R}^{nd}$ are globally Lipschitz and that $X_0$ is a square-integrable, $\mathcal{F}_0$-measurable random variable, independent of $V$ and $W$. Under these conditions, (1) has a unique solution. We will also assume that $h = (h_i)_{i=1}^m : \mathbb{R}^d \to \mathbb{R}^m$ is continuous and *bounded* and $Y_0 = 0$. The process $X$ is usually called the *signal* process and the process $Y$ is called the *observation* process.

The filtering problem consists of computing $\pi_t$, the conditional distribution of the signal at time $t$ given the observation accumulated in the interval $[0, t]$. As already stated, $\pi_t$ is the conditional distribution of the $X_t$ given the $\sigma$-field $\mathcal{Y}_t = \sigma(Y_s, 0 \le s \le t)$. Let also $\mathcal{Y}$ be the total observation $\sigma$-field, $\mathcal{Y} \triangleq \sigma(Y_s, 0 \le s < \infty)$. We note that $\pi_0$ coincides with the initial distribution of $X_0$ (as $Y_0 \equiv 0$), so we will use the same notation for both. One can prove that the (random) probability measure $\pi_t$ satisfies the Kushner–Stratonovitch equation (cf. [9] and [16]; see also [1] and [18])

$$\pi_t(\varphi) = \pi_0(\varphi) + \int_0^t \pi_s(A\varphi)\,ds + \int_0^t \big(\pi_s(\varphi h^*) - \pi_s(\varphi)\pi_s(h^*)\big)\big(dY_s - \pi_s(h)\,ds\big),$$

where $h^*$ is the row vector $(h_1 \dots h_m)$, $A : C_{\mathrm{b}}(\mathbb{R}^d) \to C_{\mathrm{b}}(\mathbb{R}^d)$ [$C_{\mathrm{b}}(\mathbb{R}^d)$ is the set of real-valued continuous bounded functions defined on $\mathbb{R}^d$] is the infinitesimal generator associated to $X$,

$$A \triangleq \sum_{i=1}^d f_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^d \sum_{k=1}^d \sigma_{ik}\sigma_{kj} \frac{\partial^2}{\partial x_i\,\partial x_j},$$

and $\varphi$ is an arbitrary continuous bounded function belonging to $\mathcal{D}(A)$. As stated in the Introduction, $\pi_t$ has an unnormalized version $p_t$, which satisfies (3), called the Zakai equation. The standard way to arrive at $p_t$ is as follows. First, we define a new probability measure $\widetilde{P}$ absolutely continuous with respect to $P$:

$$\widetilde{P}(A) = E[\mathbb{1}_A Z_t] \qquad \text{for all } A \in \mathcal{F}_t,\ t \ge 0, \tag{6}$$

where $E[\cdot]$ is the expectation with respect to $P$ and $Z = \{Z_t, \mathcal{F}_t; t \ge 0\}$ is the exponential martingale

$$Z_t = \exp\left(-\int_0^t h^*(X_s)\,dW_s - \tfrac{1}{2}\int_0^t \|h(X_s)\|_2^2\,ds\right). \tag{7}$$

In (7) and later, $h^*(X_s)$ is the row vector $(h_1(X_s) \dots h_m(X_s))$ and, if $\xi = (\xi_i)_{i=1}^m \in \mathbb{R}^m$, then we denote by $\|\xi\|_2^2$ the sum $\sum_{i=1}^m \xi_i^2$. Hence, in (7), $\|h(X_s)\|_2^2 = \sum_{i=1}^m h_i(X_s)^2$. Since $h$ is a continuous bounded function, we have

$$\|h\| \triangleq \max_{i=1,m} \sup_{x \in \mathbb{R}^m} |h_i(x)| < \infty. \tag{8}$$

Hence, $\|h(X_s)\|_2^2 \le m\|h\|^2$. Under the new measure $\tilde{P}$, $Y$ becomes a Brownian motion independent of $X$ (Girsanov's theorem). One defines, for all bounded measurable functions $\varphi$,

$$p_t(\varphi) \triangleq \tilde{E}\left[\varphi(X_t)\exp\left(\int_0^t h^*(X_s)\,dY_s - \frac{1}{2}\int_0^t \|h(X_s)\|^2\,ds\right)\Big|\mathcal{Y}_t\right], \tag{9}$$

where $\tilde{E}$ is the expectation with respect to $\tilde{P}$. Then $p_t$ as defined in (9) can be viewed as a (random) measure that satisfies the Kallianpur–Striebel formula (cf. [10]):

$$\pi_t(\varphi) = \frac{p_t(\varphi)}{p_t(1)}, \qquad P\text{-a.s.}, \tag{10}$$

and hence it is an un-normalized version of $\pi_t$. Further, (3) uniquely identifies $p_t$ as a measure-valued process. More precisely, under the conditions set up above, if $U_t$ is a $\mathcal{Y}_t$-adapted, cadlag, measure-valued process satisfying, for all $t \le T$ and for a suitably large class of test functions $\varphi$, the integral equation

$$U_t(\varphi) = \pi_0(\varphi) + \int_0^t U_s(A\varphi)\,ds + \int_0^t U_s(h^*\varphi)\,dY_s,$$

then $U_t = p_t$ for $t \le T$ almost surely (cf. [13] and [19]).

**3. Branching particle system.** Let $\{U_N(t);\ t \ge 0\}$ be a sequence of measure-valued processes representing empirical distributions of systems of branching particles. For each $N$, $U_N(0)$ is the empirical measure of $N$ (random) particles of equal mass $1/N$. More precisely,

$$U_N(0) = \frac{1}{N}\sum_{i=1}^N \delta_{x_i^N},$$

where $x_i^N$ are independent and identically distributed random variables with common distribution $\pi_0$ for every $i, N \in \mathbb{N}$. In general, $U_N(t)$ is the occupation measure of $m_N(t)$ particles of mass $1/N$. Note that the number of particles can vary, but their mass stays constant. We partition the time interval $[0, \infty)$ into subintervals of equal length $\delta t$ and describe the evolution of the particles on the interval $[i\,\delta t, (i+1)\,\delta t]$, $i = 0, 1, \dots$.

(a) During the interval, the particles move independently with the same law as the signal $X$. More precisely, if $V(s)$, $s \in [i\,\delta t, (i+1)\,\delta t)$, is the trajectory of a generic particle in the interval, then $V$ satisfies

$$dV(t) = f(V(t))\,dt + \sigma(V(t))\,dB_t,$$

where $B$ is a Brownian motion independent of $\mathcal{Y}$ and independent of all other random variables in the system.

(b) At the end of the interval, each particle branches into a random number of particles with a branching mechanism depending on its trajectory in the interval and the observation $Y_s$, $s \in [i\,\delta t, (i+1)\,\delta t]$. The branching is chosen so that the mean number of offspring of the particle with trajectory $V(s)$, $s \in [i\,\delta t, (i+1)\,\delta t)$, given the $\sigma$-field $\mathcal{F}_{(i+1)\,\delta t-} = \sigma(\mathcal{F}_s, s < (i+1)\,\delta t)$ of events up to time $(i+1)\,\delta t$, is

$$\mu_N^i(V) \triangleq \exp\biggl(\int_{i\,\delta t}^{(i+1)\,\delta t} h^*(V(t))\,dY_t - \frac{1}{2}\int_{i\,\delta t}^{(i+1)\,\delta t} h^*h(V(t))\,dt\biggr). \tag{11}$$

The variance, denoted by $\nu_N^i(V)$, of the number of offspring is minimal, consistent with the number of offspring being an integer. The particles branch independently of each other, given $\mathcal{F}_{(i+1)\,\delta t-}$, and each offspring inherits the space position of its parent.

The branching variance $\nu_N^i(V)$ satisfies

$$\nu_N^i(V) = \{\mu_N^i(V)\}\bigl(1 - \{\mu_N^i(V)\}\bigr)$$

and so is always less than $\frac{1}{4}$ ($\{x\}$ is the fractional part of $x$, $\{x\} \triangleq x - [x]$, where $[x]$ is the largest integer smaller that $x$).

We now state a number of preliminary results. They were stated and proved in [4] for the particular interbranching time $\delta t \equiv 1/N$. As the proofs for arbitrary interbranching times are identical, we will omit them here.

We aim to keep the same notation as in [4], that is:

- $m_N(t)$ is the number of particles alive at time $t$. Just before the $(i+1)$st branching, we will have $m_N(i\,\delta t)$ particles as there is no change in the number of particles in the interval $[i\,\delta t, (i+1)\,\delta t)$.
- $U_N((i+1)\,\delta t-)$ is the state of the process just before the $(i+1)$st branching.
- $V_N^j(s)$, $s \in [i\,\delta t, (i+1)\,\delta t)$, is the trajectory of the $j$th particle alive during the interval.
- $q_N^j((i+1)\,\delta t)$ is the number of offspring of the $j$th particle with $1 \le j \le m_N(i\,\delta t)$ at time $(i+1)\,\delta t$.
- $\lambda_N^{i,j}(r)$ is the exponential martingale

$$\exp\biggl(\int_{i\,\delta t}^{r} h^*(V_N^j(t))\,dY_t - \frac{1}{2}\int_{i\,\delta t}^{r} \|h(V_N^j(t))\|_2^2\,dt\biggr), \qquad r \in [i\,\delta t, (i+1)\,\delta t]. \tag{12}$$

Of course, $\mu_N^i(V_N^j) = \lambda_N^{i,j}((i+1)\,\delta t)$.

PROPOSITION 1. *The mass process $m_N(t)$ is an $\mathcal{F}_{[t/\delta t]\,\delta t}$-adapted square-integrable martingale which satisfies the following (we will always work under the new probability measure $\tilde{P}$ and all the expectations will be considered with respect to $\tilde{P}$):*

(i) $\tilde{E}[m_N(t)] = N$;
(ii) $\tilde{E}[m_N^2(t)] \le N^2 e^{\|h\|^2 [t/\delta t]\,\delta t} + \frac{1}{4}\sum_{k \le [t/\delta t]} e^{k\|h\|^2\,\delta t}$, $\forall n \ge 0$.

COROLLARY 2. *For any bounded measurable function $\varphi$, the process $(U_N(t), \varphi)$ is square integrable.*

PROPOSITION 3. *If $\varphi \in \mathcal{D}(A)$, then the process $(U_N(t), \varphi)$ satisfies the following evolution equation*:

$$
\begin{aligned}
(U_N(t), \varphi) &= (U_N(0), \varphi) + \int_0^t (U_N(s), A\varphi)\, ds + S_N^{\varphi}(t) + M_N^{\varphi}\left(\left[\frac{t}{\delta t}\right]\right) \\
&\quad + \sum_{i=1}^{[t/\delta t]} \frac{1}{N} \sum_{j=1}^{m_N(i\,\delta t)} \varphi(V_N^j(i\,\delta t))(\mu_N^i(V_N^j) - 1),
\end{aligned} \tag{13}
$$

*where $\{(S_N^{\varphi}(t), \mathcal{F}_t), t \in [0, 1]\}$ is a local martingale with quadratic variation process*

$$
\langle S_N^{\varphi}\rangle(t) = \frac{1}{N} \int_0^t (U_N(s), \|\sigma^* D\varphi\|_2^2)\, ds, \tag{14}
$$

$$
\|\sigma^* D\varphi\|_2^2 \triangleq \sum_{i=1}^m \left(\sum_{j=1}^m \sigma_{ij} \frac{\partial \varphi}{\partial x_i}\right)^2 = \sum_{j,k=1}^m \left(\sum_{i=1}^m \sigma_{ij}\sigma_{ik}\right) \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_k}
$$

*and $\{(M_N^{\varphi}(l), \mathcal{F}_{(l+1)\delta t-}), l = 0, 1, \ldots, n\}$ is a discrete martingale with angle-brackets process*

$$
\langle M_N^{\varphi}\rangle(l) = \frac{1}{N} \sum_{i=1}^l (U_N((i+1)\delta t-), \nu_N^i \varphi^2). \tag{15}
$$

By applying Itô's rule, we get from (13) that

$$
\begin{aligned}
(U_N(t), \varphi) &= (U_N(0), \varphi) + \int_0^t (U_N(s), A\varphi)\, ds + S_N^{\varphi}(t) + M_N^{\varphi}\left(\left[\frac{t}{\delta t}\right]\right) \\
&\quad + \frac{1}{N} \int_0^{[t/\delta t]\,\delta t} \sum_{j=1}^{m_N([s/\delta t]\,\delta t)} \varphi\left(V_N^j\left(\left(\left[\frac{s}{\delta t}\right] + 1\right)\delta t\right)\right) \\
&\qquad \times \lambda_N^{[s/\delta t], j}(s) h^*(V_N^j(s))\, dY_s.
\end{aligned} \tag{16}
$$

LEMMA 4. *For all $i = 1, 2, \ldots,$ we have*

$$
\tilde{E}\left[\sum_{j=1}^{m_N(i\,\delta t)} \nu_N^i(V_N^j)\right] \le cN\sqrt{\delta t}, \tag{17}
$$

*where $c$ is a constant independent of $N$ and $\delta t$.*

Now let $\varphi \in C_b(\mathbb{R}^d)$ be a continuous bounded function with bounded first- and second-order partial derivatives and let $\{\psi_s\}_{0 \le s \le t}$ be the solution of the following backward stochastic partial differential equation:

$$
\begin{aligned}
d\psi_s &= -A\psi_s\, ds - h^* \psi_s\, \bar{d}Y_s, \qquad s \le t, \\
\psi_t &= \varphi,
\end{aligned} \tag{18}
$$

the integral form of (18) being

$$\psi_r = \psi_s - \int_s^r A\psi_p\,dp - \int_s^r h^*\psi_p\,\bar{d}Y_p, \qquad 0 \le s \le r \le t, \tag{19}$$

where $\int_s^r h^*\psi_p\,\bar{d}Y_p$ is a backward Itô integral. Obviously, $\{\psi_s\}_{0\le s\le t}$ has the representation $\psi_s = \bar{\psi}_{t-s}$, where $\{\bar{\psi}_s\}_{0\le s\le t}$ is the solution of the (forward) stochastic partial differential equation

$$\begin{aligned} d\bar{\psi}_s &= A\bar{\psi}_s\,ds + h^*\bar{\psi}_s\,d\bar{Y}_s, \qquad s \le t,\\ \bar{\psi}_0 &= \varphi, \end{aligned}$$

where $\bar{Y}_s \triangleq Y_t - Y_{t-s}$. Equation (18) has been extensively studied; see [1] and [18] for results of the existence and uniqueness of a solution of (18) in appropriate spaces of solutions. In the following, we will assume that $f$, $\sigma$ and $h$ satisfy sufficient conditions so that $\psi_s \in \mathcal{D}(A)$ for all $s \in [0,t]$ and

$$\tilde{E}\Big[\sup_{s\in[0,t]} \|\psi_s\|_\infty^2\Big] + \tilde{E}\Big[\sup_{s\in[0,t]} \|\sigma^* D\psi_s\|_2^2\Big] < \infty \tag{20}$$

for test functions $\varphi \in \mathcal{M} \cup \{1\}$, where $\mathcal{M} = \{\varphi_k, k \ge 1\}$ is a countable set uniformly dense in the set of all continuous bounded functions over $\mathbb{R}^d$ with compact support. In essence, we want $\mathcal{M} \cup \{1\}$ to be convergence determining for the set of finite measures over $\mathbb{R}^d$ (see [8]). One can find sufficient conditions on $f$, $\sigma$ and $h$ under which (20) holds in [2].

THEOREM 5. *If $\varphi$ is a test function for which* (20) *holds true*, *then*

$$\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \tilde{E}\big[\big((U_N(t),\varphi) - p_t(\varphi)\big)^2\big] = 0. \tag{21}$$

REMARK 1. In (21), it is important that, for each $N$, the length $\delta t = \delta t_N$ of the corresponding interbranching times is chosen so that $\lim_{N\to\infty} N\sqrt{\delta t} = \infty$. If, say, $\lim_{N\to\infty} N\sqrt{\delta t}$ exists and it is finite, then $U_N$ converges (in distribution) to a different measure-valued process.

**4. Exact rate of convergence of particle approximation.** The essential ingredient in the proof of Theorem 5 is the upper bound (17) on the sum of $\nu_N^i(V_N^j)$, $j = 1, 2, \dots$, the variance of the individual branching mechanisms. The size of $\nu_N^i(V_N^j)$ is important as it measures the amount of extra randomness introduced in the system at branching times. To obtain the exact rate of convergence of $U_N$, we need a better estimate on $\nu_N^i(V_N^j)$. For the following result, we need to introduce the $\sigma$-fields: $\mathcal{Y}^{(i+1)\,\delta t} \triangleq \sigma(Y_s - Y_{(i+1)\,\delta t} \mid s \ge (i+1)\,\delta t)$ and $\mathcal{B}^{j,N} \triangleq \sigma(B^{j,N}(s), s \in [i\,\delta t, (i+1)\,\delta t])$, where $B^{j,N}$ is the Brownian motion that generates $V_N^j$ of the $j$th particle; in other words, $V_N^j$ satisfies

$$dV_N^j(t) = f\bigl(V_N^j(t)\bigr)\,dt + \sigma\bigl(V_N^j(t)\bigr)\,dB_t^{j,N}, \qquad t \in [i\,\delta t, (i+1)\,\delta t].$$

Also, let $\mathcal{H}^{i,j,N}$ be the $\sigma$-field $\mathcal{H}^{i,j,N} \triangleq \mathcal{F}_{i\,\delta t} \vee \mathcal{Y}^{(i+1)\,\delta t} \vee \mathcal{B}^{j,N}$. Then the following proposition holds true.

PROPOSITION 6. *For all $i = 0, 1, \dots,$ we have*

$$\left|\tilde{E}\bigl[\nu_N^i(V_N^j)|\mathcal{H}^{i,j,N}\bigr] - \sqrt{\frac{2}{\pi}}\sqrt{\int_{i\,\delta t}^{(i+1)\,\delta t} \bigl\|h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt}\right| \le c\,\delta t, \tag{22}$$

*where $c$ is a constant independent of $N$ and $\delta t$.*

PROOF. First, let us note that $\nu_N^i(V_N^j) = \digamma(\mu_N^i(V_N^j))$, where

$$\digamma(x) \triangleq \{x\}(1-\{x\}) = x(1-x) + 2\sum_{k\ge 1}(x-k)_+.$$

Hence, the function $\digamma$ is a linear combination of convex functions. By applying the generalized Itô formula to the exponential martingale $\lambda_N^{i,j}$ [as defined in (12)] and the function $\digamma$, we obtain the following representation for $\nu_N^i(V_N^j)$:

$$\begin{aligned}\nu_N^i(V_N^j) &= \int_{i\,\delta t}^{(i+1)\,\delta t} D^-\digamma\bigl(\lambda_N^{i,j}(t)\bigr)\lambda_N^{i,j}(t)h^*\bigl(V_N^j(t)\bigr)\,dY_t \\ &\quad - \int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(t)^2\bigl\|h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt + 2\sum_{k\ge 1} L_{(i+1)\,\delta t}(k).\end{aligned} \tag{23}$$

In (23), $L_r(k)$ is the local time at $k$ associated to the martingale $r \to \lambda_N^{i,j}(r)$ and $D^-\digamma$ is the left derivative of $\digamma$:

$$D^-\digamma(x) = \left(1 - 2x + 2\sum_{k\ge 1}\mathbb{1}_{(k,\infty)}(x)\right) = \begin{cases} -1, & \text{if } x \in \mathbb{N}, \\ (1-2\{x\}), & \text{if } x \notin \mathbb{N}, \end{cases} \qquad x > 0.$$

Hence, $|D^-\digamma(x)| \le 1$ for all $x > 0$. We estimate now each of the three terms on the right-hand side of (23).

Using the independent increments property of the Brownian motion $Y$ and the independence of $B^{j,N}$, it is easy to check that the process

$$r \to \int_{i\,\delta t}^{r} D^-\digamma\bigl(\lambda_N^{i,j}(t)\bigr)\lambda_N^{i,j}(t)h^*\bigl(V_N^j(t)\bigr)\,dY_t$$

is a genuine martingale with respect to the enlarged filtration $r \to \mathcal{F}_r \vee \mathcal{Y}^{(i+1)\,\delta t} \vee \mathcal{B}^{j,N}$, $r \ge i\,\delta t$. Hence, its conditional expectation with respect to $\mathcal{H}^{i,j,N}$ is 0.

For the second term in (23), we first observe that

$$\lambda_N^{i,j}(r)^2 = \exp\biggl(\int_{i\,\delta t}^r 2h^*\bigl(V_N^j(t)\bigr)\,dY_t - \int_{i\,\delta t}^r \bigl\|h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt\biggr)$$

$$= \exp\biggl(\int_{i\,\delta t}^r \bigl\|h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt\biggr)\exp\biggl(\int_{i\,\delta t}^r 2h^*\bigl(V_N^j(t)\bigr)\,dY_t - \frac{1}{2}\int_{i\,\delta t}^r \bigl\|2h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt\biggr) \tag{24}$$

and, since

$$r \to \exp\biggl(\int_{i\,\delta t}^r 2h^*\bigl(V_N^j(t)\bigr)\,dY_t - \frac{1}{2}\int_{i\,\delta t}^r \bigl\|2h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt\biggr)$$

is also an exponential martingale with respect to the enlarged filtration $r \to \mathcal{F}_r \vee \mathcal{Y}^{(i+1)\delta t} \vee \mathcal{B}^{j,N}$, we have

$$\tilde{E}\bigl[\lambda_N^{i,j}(t)^2|\mathcal{H}^{i,j,N}\bigr] = \exp\biggl(\int_{i\,\delta t}^r \bigl\|h\bigl(V(t)\bigr)\bigr\|_2^2\,dt\biggr) \le \exp\bigl((r - i\,\delta t)m\|h\|^2\bigr).$$

Thus (using the fact that $e^\theta - 1 \le e^\theta\theta$ for $\theta \ge 0$),

$$\begin{aligned}\tilde{E}\biggl[\int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(t)^2 \bigl\|h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt\,\Big|\,\mathcal{H}^{i,j,N}\biggr] &\le \bigl(e^{\delta t m\|h\|^2} - 1\bigr)\\ &\le m\|h\|^2 e^{\delta t m\|h\|^2}\,\delta t\\ &\le m\|h\|^2 e^{m\|h\|^2}\,\delta t.\end{aligned} \tag{25}$$

Finally, let us analyze the third term. First, observe that for all $x > 0$ we have

$$2\sum_{k\ge 2}(x-k)_+ \le 2\sum_{k\ge 2}\int_{k-1}^k (x-u)_+\,du = 2\int_1^\infty (x-u)_+\,du = (x-1)^2.$$

Hence,

$$2\sum_{k\ge 2}\bigl(\lambda_N^{i,j}((i+1)\,\delta t) - k\bigr)_+ \le \bigl(\lambda_N^{i,j}((i+1)\,\delta t) - 1\bigr)^2$$

and since

$$\begin{aligned}2\sum_{k\ge 2}\bigl(\lambda_N^{i,j}((i+1)\,\delta t) - k\bigr)_+ &= 2\int_{i\,\delta t}^{(i+1)\,\delta t} \sum_{k\ge 2}\mathbb{1}_{(k,\infty)}\bigl(\lambda_N^{i,j}(t)\bigr)h^*\bigl(V_N^j(t)\bigr)\,dY_t\\ &\quad + 2\sum_{k\ge 2} L_{(i+1)\,\delta t}(k),\end{aligned}$$

$$\begin{aligned}\bigl(\lambda_N^{i,j}((i+1)\,\delta t) - 1\bigr)^2 &= 2\int_{i\,\delta t}^{(i+1)\,\delta t} \bigl(\lambda_N^{i,j}((i+1)\,\delta t) - 1\bigr)h^*\bigl(V_N^j(t)\bigr)\,dY_t\\ &\quad + \int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(t)^2\bigl\|h\bigl(V_N^j(t)\bigr)\bigr\|_2^2\,dt,\end{aligned}$$

it follows, by taking conditional expectation, that

$$\tilde{E}\left[2\sum_{k\geq 2} L_{(i+1)\,\delta t}(k)\middle|\mathcal{H}^{i,j,N}\right] \leq \tilde{E}\left[\int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(t)^2 \|h(V_N^j(t))\|_2^2\,dt\middle|\mathcal{H}^{i,j,N}\right]$$

and, consequently, we obtain the following upper bound [using (25)]:

$$\tilde{E}\left[2\sum_{k\geq 2} L_{(i+1)\,\delta t}(k)\middle|\mathcal{H}^{i,j,N}\right] \leq m\|h\|^2 e^{m\|h\|^2\,\delta t}\,\delta t. \tag{26}$$

The only term left to estimate is $2L_{(i+1)\,\delta t}(1)$. Using again the generalized Itô formula, we obtain

$$\begin{aligned}
&\left|\int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(t) h^*(V_N^j(t))\,dY_t\right| \\
&\qquad = \left|\lambda_N^{i,j}((i+1)\delta t) - 1\right| \\
&\qquad = 2L_{(i+1)\,\delta t}(1) + \int_{i\,\delta t}^{(i+1)\,\delta t} \operatorname{sgn}(\lambda_N^{i,j}(t) - 1)\lambda_N^{i,j}(t) h^*(V_N^j(t))\,dY_t.
\end{aligned}$$

Thus,

$$\tilde{E}[2L_{(i+1)\,\delta t}(1)|\mathcal{H}^{i,j,N}] = \tilde{E}\left[\left|\int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(t) h^*(V_N^j(t))\,dY_t\right|\middle|\mathcal{H}^{i,j,N}\right]. \tag{27}$$

It is straightforward to show that

$$\left|\tilde{E}\left[\left|\int_{i\,\delta t}^{(i+1)\,\delta t} (\lambda_N^{i,j}(t) - 1) h^*(V_N^j(t))\,dY_t\right|\middle|\mathcal{H}^{i,j,N}\right]\right| \leq m\|h\|^2 e^{m\|h\|^2\,\delta t}\,\delta t. \tag{28}$$

From (27) and (28), we obtain

$$\left|\tilde{E}\left[2L_{(i+1)\,\delta t}(1) - \left|\int_{i\,\delta t}^{(i+1)\,\delta t} h^*(V_N^j(t))\,dY_t\right|\middle|\mathcal{H}^{i,j,N}\right]\right| \leq m\|h\|^2 e^{m\|h\|^2\,\delta t}\,\delta t. \tag{29}$$

Finally, since the process $r \to (Y_r - Y_{i\delta t})$ is independent of $\mathcal{H}^{i,j,N}$, then, given $\mathcal{H}^{i,j,N}$, the random variable $\int_{i\,\delta t}^{(i+1)\,\delta t} h^*(V_N^j(t))\,dY_t$ has a Gaussian distribution with mean 0 and variance $\int_{i\,\delta t}^{(i+1)\,\delta t} \|h(V_N^j(t))\|_2^2\,dt$:

$$\int_{i\,\delta t}^{(i+1)\,\delta t} h^*(V_N^j(t))\,dY_t \sim N\left(0, \int_{i\,\delta t}^{(i+1)\,\delta t} \|h(V_N^j(t))\|_2^2\,dt\right).$$

Hence,

$$\tilde{E}\left[\left|\int_{i\,\delta t}^{(i+1)\,\delta t} h^*(V_N^j(t))\,dY_t\right|\middle|\mathcal{H}^{i,j,N}\right] = \sqrt{\frac{2}{\pi}}\sqrt{\int_{i\,\delta t}^{(i+1)\,\delta t} \|h(V_N^j(t))\|_2^2\,dt},$$

which completes the proof of the proposition. □

LEMMA 7. *Let* $\{V_r, r \in [s,t]\}$ *be a realization of the signal process, that is,*

$$dV_r = f(V_r)\,dr + \sigma(V_r)\,dB_r,$$

*where* $B$ *is a Brownian motion independent of* $\mathcal{Y}$, *and let* $\xi_r$ *be the exponential martingale*

$$\xi_r = \exp\biggl(\int_s^r h^*(V_p)\,dY_p - \tfrac{1}{2}\int_s^r \|h(V_p)\|_2^2\,dp\biggr).$$

*Then we have the following formulas*:

$$\psi_r(V_r) = \psi_s(V_s) - \int_s^r \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p + \int_s^r \bigl(D\psi_p(V_p)\bigr)^*\sigma(V_p)\,dB_p, \tag{30}$$

$$\psi_r(V_r)\xi_r = \psi_s(V_s) + \int_s^r \xi_p\bigl(D\psi_p(V_p)\bigr)^{\star}\sigma(V_p)\,dB_p, \tag{31}$$

$$\begin{aligned}\psi_r(V_r)^2 = \psi_s(V_s)^2 &- 2\int_s^r \psi_p(V_p)^2 h^*(V_p)\,\bar{d}Y_p\\ &+ 2\int_s^r \psi_p(V_p)\bigl(D\psi_p(V_p)\bigr)^*\sigma(V_p)\,dB_p\\ &- \int_s^r \psi_p(V_p)^2\|h(V_p)\|_2^2\,ds + \int_s^r \|\sigma^* D\psi_p(V_p)\|_2^2\,dp.\end{aligned} \tag{32}$$

PROOF. We will only prove (30), as (31) and (32) have similar proofs. For this, we follow an argument similar to that contained in the proofs of Theorems 4.1.2 and 4.2.1 in [1]. Let us state first a density result whose proof is identical to that of Lemma 4.1.4, in [1]. Let $b(r)$ and $c(r)$, $r \in [s,t]$, be bounded, Borel measurable, deterministic functions with values in $\mathbb{R}^d$, respectively, $\mathbb{R}^m$. Le $\theta_b$ and $\theta_c$ be the following processes:

$$\theta_b(r) \triangleq \exp\biggl(i\int_s^r b^*(p)\,dY_p + \tfrac{1}{2}\int_s^r \|b(p)\|^2\,dp\biggr), \tag{33}$$

$$\theta_c(r) \triangleq \exp\biggl(i\int_s^r c^*(p)\,dB_p + \tfrac{1}{2}\int_s^r \|c(p)\|^2\,dp\biggr). \tag{34}$$

Then we have the following result.

PROPOSITION 8. *Let* $\mathcal{W}$ *be an integrable random variable, measurable with respect to the* $\sigma$*-field* $\mathcal{F}_s \vee \mathcal{Y}_s^t \vee \mathcal{B}_s^t$, *where* $\mathcal{Y}_s^t \triangleq \sigma(Y_r - Y_s \mid r \in [s,t])$ *and* $\mathcal{B}_s^t \triangleq \sigma(B_r - B_s \mid r \in [s,t])$ *such that*

$$\tilde{E}[\mathcal{W}\zeta\theta_b(t)\theta_c(t)] = 0$$

*for any choice of* $b$ *and* $c$ *in* (33), *respectively*, (34) *and any bounded* $\mathcal{F}_s$*-measurable random variable* $\zeta$. *Then necessarily* $\mathcal{W} = 0$, $\tilde{P}$ *almost surely.*

As a corollary to the above result, if we show that

$$\tilde{E}\big[\big(\psi_r(V_r)-\psi_s(V_s)\big)\zeta\theta_b(t)\theta_c(t)\big]$$

$$= \tilde{E}\bigg[\bigg(-\int_s^r \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p + \int_s^r \big(D\psi_p(V_p)\big)^*\sigma(V_p)\,dB_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg], \tag{35}$$

then we have proved (30). First, observe that

$$\tilde{E}\big[\psi_r(V_r)\zeta\theta_b(t)\theta_c(t)|\mathcal{F}_s\vee\mathcal{Y}_s^r\vee\mathcal{B}_s^r\big] = \Xi_r(V_r)\zeta\theta_b(r)\theta_c(r),$$

where

$$\Xi_r \triangleq \tilde{E}\bigg[\psi_r\frac{\theta_b(t)}{\theta_b(r)}\bigg|\mathcal{F}_s\vee\mathcal{Y}_s^r\vee\mathcal{B}_s^r\bigg] = \tilde{E}\big[\psi_r\tilde{\theta}_b(r)|\mathcal{F}_s\vee\mathcal{Y}_s^r\vee\mathcal{B}_s^r\big]$$

and

$$\tilde{\theta}_b(r) \triangleq \exp\bigg(i\int_r^t b^*(p)\,dY_p + \tfrac{1}{2}\int_r^t \|b(p)\|^2\,dp\bigg).$$

Hence,

$$\tilde{\theta}_b(r) = 1 - \int_r^t ib^*(p)\tilde{\theta}_b(p)\,\bar{d}Y_p. \tag{36}$$

Now since both $\psi_r$ and $\tilde{\theta}_b(r)$ are measurable with respect to the $\sigma$-field $\mathcal{Y}_r^t$, which is independent of $\mathcal{F}_s\vee\mathcal{Y}_s^r\vee\mathcal{B}_s^r$, we get that $\Xi_r = E[\psi_r\tilde{\theta}_b(r)|\mathcal{F}_s\vee\mathcal{Y}_s^r\vee\mathcal{B}_s^r] = E[\psi_r\tilde{\theta}_b(r)]$. Also,

$$\begin{aligned}\tilde{E}[\psi_r(V_r)\zeta\theta_b(t)\theta_c(t)|\mathcal{F}_s\vee\mathcal{B}_s^r] &= E[\Xi_r(V_r)\zeta\theta_b(r)\theta_c(r)|\mathcal{F}_s\vee\mathcal{B}_s^r]\\ &= \Xi_r(V_r)\zeta\theta_c(r).\end{aligned}$$

By Itô's rule, from (18) and (36) and the fact that $Y$ is a Brownian motion, we get that

$$d\psi_r\tilde{\theta}_b(r) = \big(-A\psi_r\tilde{\theta}_b(r) - ih^*b(r)\psi_r\tilde{\theta}_b(r)\big)\,dr - \big(h^*+ib^*(r)\big)\psi_r\tilde{\theta}_b(r)\,\bar{d}Y_r.$$

Thus, we have consecutively

$$\begin{aligned}d\,\Xi_r &= \big(-A\,\Xi_r - ih^*b(r)\,\Xi_r\big)\,dr,\\ d\,\Xi_r(V_r) &= -ih^*(V_r)b(r)\,\Xi_r(V_r)\,dr + \big(D\,\Xi_r(V_r)\big)^*\sigma(V_r)\,dB_r,\\ d\,\Xi_r(V_r)\theta_c(r) &= -ih^*(V_r)b(r)\,\Xi_r(V_r)\theta_c(r)\,dr + \big(D\,\Xi_r(V_r)\big)^*\sigma(V_r)\theta_c(r)\,dB_r\\ &\quad + i\,\Xi_r(V_r)\theta_c(r)c^*(r)\,dB_r + ic(r)\big(D\,\Xi_r(V_r)\big)^*\sigma(V_r)\theta_b(r)\,dr,\end{aligned}$$

which implies

$$
\begin{aligned}
&\tilde{E}\big[\big(\psi_r(V_r)-\psi_s(V_s)\big)\zeta\theta_b(t)\theta_c(t)\big]\\
&\qquad=\tilde{E}\big[\zeta\big(\Xi_r(V_r)\theta_c(r)-\Xi_s(V_s)\theta_c(s)\big)\big]\\
&\qquad=-i\int_s^r \tilde{E}\big[\zeta\theta_c(p)h^*(V_p)b(r)\Xi_r(V_r)\big]\,dp\\
&\qquad\quad+i\int_s^r \tilde{E}\big[\zeta\theta_c(p)c(r)\big(D\Xi_r(V_r)\big)^*\sigma(V_r)\big]\,dp.
\end{aligned}\tag{37}
$$

Next, we have, as before, that

$$
\begin{aligned}
&\tilde{E}\bigg[\bigg(\int_r^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg|\mathcal{F}_s\vee\mathcal{Y}_r^t\vee\mathcal{B}_s^t\bigg]\\
&\qquad=\bigg(\int_r^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\tilde{\theta}_b(r)\zeta\theta_c(t).
\end{aligned}
$$

Further, since

$$
\begin{aligned}
\bigg(\int_r^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\tilde{\theta}_b(r)&=\int_r^t \psi_p(V_p)\tilde{\theta}_b(p)\big(h^*(V_p)-ib^*(p)\big)\,\bar{d}Y_p\\
&\quad+i\bigg(\int_s^t h^*(V_r)b(r)\psi_p(V_p)\tilde{\theta}_b(p)\,dp\bigg),
\end{aligned}
$$

it follows that

$$
\begin{aligned}
&\tilde{E}\bigg[\bigg(\int_r^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg|\mathcal{F}_s\vee\mathcal{B}_s^t\bigg]\\
&\qquad=i\bigg(\int_r^t h^*(V_p)b(p)\Xi_r(V_p)\,dp\bigg)\zeta\theta_c(t)
\end{aligned}\tag{38}
$$

and, similarly, that

$$
\begin{aligned}
&\tilde{E}\bigg[\bigg(\int_s^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg|\mathcal{F}_s\vee\mathcal{B}_s^t\bigg]\\
&\qquad=i\bigg(\int_s^t h^*(V_r)b(r)\Xi_r(V_r)\,dp\bigg)\zeta\theta_c(t).
\end{aligned}\tag{39}
$$

From (38) and (39), we deduce that

$$
\begin{aligned}
&\tilde{E}\bigg[\bigg(\int_s^r \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg]\\
&\qquad=\tilde{E}\bigg[\bigg(\int_s^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg]\\
&\qquad\quad-\tilde{E}\bigg[\bigg(\int_r^t \psi_p(V_p)h^*(V_p)\,\bar{d}Y_p\bigg)\zeta\theta_b(t)\theta_c(t)\bigg]
\end{aligned}\tag{40}
$$

$$= \tilde{E}\left[i\left(\int_s^r h^*(V_r)b(r)\Xi_r(V_r)\,dp\right)\zeta\theta_c(t)\right]$$
$$= i\int_s^r \tilde{E}[\zeta\theta_c(p)h^*(V_p)b(r)\Xi_r(V_r)]\,dp.$$

Finally, one proves in a similar fashion that

$$\begin{aligned}\tilde{E}\left[\left(\int_s^r \big(D\psi_p(V_p)\big)^*\sigma(V_p)\,dB_p\right)\zeta\theta_b(t)\theta_c(t)\right]\\ = i\int_s^r \tilde{E}\big[\zeta\theta_c(p)c(r)\big(D\Xi_r(V_r)\big)^*\sigma(V_r)\big]\,dp\end{aligned}\tag{41}$$

and, from (37), (40) and (41), we deduce (35). □

PROPOSITION 9. *If* (20) *is satisfied*, *then*:

(i) $p_r(\psi_r) = p_t(\varphi)$ *for all* $r \ge 0$, *in particular*, $p_t(\varphi) = \pi_0(\psi_0)$;

(ii) $\{U_N((i\,\delta t), \psi_{i\,\delta t}), \mathcal{F}_{i\,\delta t} \vee \mathcal{Y}\}$ *is a discrete martingale*, *in particular*, $\tilde{E}[U_N(i\,\delta t)|\mathcal{Y}] = p_{i\,\delta t}$.

PROOF. (i) From (31), it follows that

$$\psi_r(X_r)\xi_r = \psi_0(X_0) + \int_0^r \xi_p\big(D\psi_p(X_p)\big)^*\sigma(X_p)\,dV_p$$

and, since $r \to \int_0^r \xi_p(D\psi_p(X_p))^*\sigma(X_p)\,dV_p$, $r \in [0,t]$, is a martingale with respect to the filtration $\mathcal{F}_r \vee \mathcal{Y}$, we have

$$\begin{aligned}p_t(\varphi) &= \widetilde{E}[\varphi(X_t)\xi_t|\mathcal{Y}] = \widetilde{E}[\psi_t(X_t)\xi_t|\mathcal{Y}]\\ &= \widetilde{E}\big[\widetilde{E}[\psi_t(X_t)\xi_t|\mathcal{F}_r \vee \mathcal{Y}]|\mathcal{Y}\big] = \widetilde{E}[\psi_r(X_r)\xi_r|\mathcal{Y}] = p_r(\psi_r).\end{aligned}\tag{42}$$

In (42), we used the fact that $\psi_t \equiv \varphi$. In particular, as $p_0 = \pi_0$,

$$p_t(\varphi) = \pi_0(\psi_0).\tag{43}$$

(ii) For $i = 1, 2, \ldots,$ we have the following identity:

$$\begin{aligned}&\tilde{E}\big[\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big)\big|\mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big]\\ &\quad= \sum_{j=1}^{m_N(i\,\delta t)} \psi_{(i+1)\,\delta t}\big(V_N^j\big((i+1)\,\delta t\big)\big)\tilde{E}\big[q_N^j\big((i+1)\,\delta t\big)\big|\mathcal{F}_{(i+1)\,\delta t-}\big]\\ &\quad= \sum_{j=1}^{m_N(i\,\delta t)} \psi_{(i+1)\,\delta t}\big(V_N^j\big((i+1)\,\delta t\big)\big)\mu_N^i(V_N^j).\end{aligned}$$

But from (31), $\tilde{E}[\psi_{(i+1)\,\delta t}(V_N^j((i+1)\,\delta t))\mu_N^i(V_N^j)|\mathcal{F}_{i\,\delta t}\vee\mathcal{Y}]=\psi_{i\,\delta t}(V_N^j(i\,\delta t))$, so

$$\tilde{E}\big[\big(U_N\big((i+1)\,\delta t\big),\psi_{(i+1)\delta t}\big)\big|\mathcal{F}_{i\,\delta t}\vee\mathcal{Y}\big]$$

$$(44)\qquad =\sum_{j=1}^{m_N(i\,\delta t)}\tilde{E}\big[\psi_{(i+1)\,\delta t}\big(V_N^j\big((i+1)\,\delta t\big)\big)\mu_N^i(V_N^j)\big|\mathcal{F}_{i\,\delta t}\vee\mathcal{Y}\big]$$

$$=\big(U_N(i\,\delta t),\psi_{i\,\delta t}\big).$$

Hence, $\{U_N(i\,\delta t,\psi_{i\,\delta t}),\ \mathcal{F}_{i\,\delta t}\vee\mathcal{Y}\}$ is a discrete martingale. Finally, as $\tilde{E}[(U_N(0),\psi_0)|\mathcal{Y}]=\pi_0(\psi_0)=p_t(\varphi)$, we have that

$$\begin{aligned}\tilde{E}\big[(U_N(i\,\delta t),\varphi)|\mathcal{Y}\big]&=\tilde{E}\big[(U_N(i\,\delta t),\psi_{i\,\delta t})|\mathcal{Y}\big]\\&=\tilde{E}\big[\tilde{E}\big[\big(U_N(i\,\delta t),\psi_{i\,\delta t}\big)\big|\mathcal{F}_0\vee\mathcal{Y}\big]|\mathcal{Y}\big]\\&=\tilde{E}\big[\big(U_N(0),\psi_0\big)|\mathcal{Y}\big]\\&=p_t(\varphi)\end{aligned}$$

for all functions $\varphi$ for which (20) holds true. Since these functions form a convergence-determining set and hence, separating, it follows that $\tilde{E}[U_N(i\,\delta t)|\mathcal{Y}]=p_t$. $\square$

We now have all the results needed to prove our main theorem. In the theorem below, we will assume that the interbranching times $\delta t=\delta t(N)$ are chosen so that the fixed time $t$ is an integer multiple of $\delta t$, for all $N$, $t=i(N)\,\delta t(N)=i\,\delta t$. In the corollary following the theorem, we also look at times $t$ that may fall in between branching times.

THEOREM 10. *If $\varphi$ satisfies* (20), *$h$ is assumed to be Lipschitz and $t$ is a branching time for all $N$, then*

$$(45)\qquad \lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} N\sqrt{\delta t}\,\tilde{E}\big[\big((U_N(t),\varphi)-p_t(\varphi)\big)^2\big]=\sqrt{\frac{2}{\pi}}\int_0^t\tilde{E}\big[p_s(\psi_s^2\|h\|_2)\big]\,ds.$$

PROOF. Since $\{U_N((i\,\delta t),\psi_{i\,\delta t}),\ \mathcal{F}_{i\,\delta t}\vee\mathcal{Y}\}$ is a discrete martingale, we have that

$$\begin{aligned}&\tilde{E}\big[\big((U_N(t),\varphi)-p_t(\varphi)\big)^2\big]\\&\quad=\sum_{i=0}^{t/\delta t-1}\tilde{E}\big[\big(\big(U_N\big((i+1)\,\delta t\big),\psi_{(i+1)\,\delta t}\big)-\big(U_N(i\,\delta t),\psi_{i\,\delta t}\big)\big)^2\big]\\&\qquad+\tilde{E}\big[\big(\big(U_N(0),\psi_0\big)-\pi_0(\psi_0)\big)^2\big].\end{aligned}$$

Also, we have

$$
\begin{aligned}
&\tilde{E}\big[\big(\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) - \big(U_N(i\,\delta t), \psi_{i\,\delta t}\big)\big)^2\big] \\
&\quad = \tilde{E}\big[\big(\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) \\
&\qquad\qquad - \widetilde{E}\big[\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) | \mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big]\big)^2\big] \\
&\qquad + \tilde{E}\big[\big(\widetilde{E}\big[\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) | \mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big] \\
&\qquad\qquad\qquad - \big(U_N(i\,\delta t), \psi_{i\,\delta t}\big)\big)^2\big].
\end{aligned}
\tag{46}
$$

From (31), we get that

$$
\begin{aligned}
&\widetilde{E}\big[\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) | \mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big] - \big(U_N(i\,\delta t), \psi_{i\,\delta t}\big) \\
&\quad = \sum_{j=1}^{m_N(i\,\delta t)} \psi_{(i+1)\,\delta t}\big(V_N^j\big((i+1)\,\delta t\big)\big)\mu_N^i(V_N^j) - \psi_{i\,\delta t}\big(V_N^j(i\,\delta t)\big) \\
&\qquad + \sum_{j=1}^{m_N(i\,\delta t)} \int_{i\,\delta t}^{(i+1)\,\delta t} \lambda_N^{i,j}(p)\big(D\psi_p\big(V_N^j(p)\big)\big)^* \sigma\big(V_N^j(p)\big)\, dB_p^{j,N}.
\end{aligned}
$$

Thus, the last term in (46) is of order $\delta t/N$:

$$
\begin{aligned}
&\tilde{E}\Big[\Big(\widetilde{E}\big[\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) \big| \mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big] - \big(U_N(i\,\delta t), \psi_{i\,\delta t}\big)\Big)^2\Big] \\
&\quad \le c\frac{\delta t}{N}.
\end{aligned}
\tag{47}
$$

The first term in (46) satisfies

$$
\begin{aligned}
&\tilde{E}\Big[\big(\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) \\
&\qquad - \widetilde{E}\big[\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) \big| \mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big]\big)^2\Big] \\
&\quad = \frac{1}{N^2}\widetilde{E}\Bigg[\sum_{j=1}^{m_N(i\,\delta t)} \big(\psi_{(i+1)\,\delta t}\big(V_N^j\big((i+1)\,\delta t\big)\big)\big)^2 \nu_n^i(V_N^j)\Bigg].
\end{aligned}
\tag{48}
$$

Finally, since $U_N(0) = (1/N)\sum_{i=1}^N \delta_{x_i^N}$, where $x_i^N$ are i.i.d. random variables with common distribution $\pi_0$ and $\psi_0 \in \mathcal{D}(A) \subset C_b(E)$, $\tilde{P}$-a.s., we have that

$$
\widetilde{E}\big[\big((U_N(0), \psi_0) - \pi_0(\psi_0)\big)^2\big] = \frac{\widetilde{E}[\pi_0(\psi_0^2)]}{N}.
\tag{49}
$$

In (49), $0 \le \pi_0(\psi_0^2) \le \|\psi_0\|^2$. Hence, based on (20), the integral on the right-hand side of (49) is finite. From (46)–(49), it follows that

$$\begin{aligned}&\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} N\sqrt{\delta t}\,\tilde{E}\big[\big((U_N(t),\varphi)-p_t(\varphi)\big)^2\big]\\ &\quad=\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \frac{\sqrt{\delta t}}{N}\sum_{i=0}^{t/\delta t-1}\widetilde{E}\Bigg[\sum_{j=1}^{m_N(i\,\delta t)}\big(\psi_{(i+1)\,\delta t}\big(V_N^j((i+1)\,\delta t)\big)\big)^2\nu_n^i(V_N^j)\Bigg].\end{aligned}\tag{50}$$

Now (50) and Proposition 6 give

$$\begin{aligned}&\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \frac{\sqrt{\delta t}}{N}\tilde{E}\big[\big((U_N(t),\varphi)-p_t(\varphi)\big)^2\big]\\ &\quad=\sqrt{\frac{2}{\pi}}\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \frac{\sqrt{\delta t}}{N}\sum_{i=0}^{t/\delta t-1}\widetilde{E}\Bigg[\sum_{j=1}^{m_N(i\,\delta t)}\big(\psi_{(i+1)\,\delta t}\big(V_N^j((i+1)\,\delta t)\big)\big)^2\\ &\qquad\qquad\times\sqrt{\int_{i\,\delta t}^{(i+1)\,\delta t}\big\|h\big(V_N^j(t)\big)\big\|_2^2\,dt}\,\Bigg]\end{aligned}\tag{51}$$

and, using (32) and the Lipschitz condition on $h$, we get that

$$\begin{aligned}&\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \frac{\sqrt{\delta t}}{N}\sum_{i=0}^{t/\delta t-1}\widetilde{E}\Bigg[\sum_{j=1}^{m_N(i\,\delta t)}\big(\psi_{(i+1)\,\delta t}\big(V_N^j((i+1)\,\delta t)\big)\big)^2\\ &\qquad\qquad\times\sqrt{\int_{i\,\delta t}^{(i+1)\,\delta t}\big\|h\big(V_N^j(t)\big)\big\|_2^2\,dt}\,\Bigg]\\ &\quad=\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \frac{1}{N}\sum_{i=0}^{t/\delta t-1}\widetilde{E}\Bigg[\sum_{j=1}^{m_N(i\,\delta t)}\big(\psi_{i\,\delta t}\big(V_N^j(i\,\delta t)\big)\big)^2\big\|h\big(V_N^j(i\,\delta t)\big)\big\|_2\Bigg]\delta t\\ &\quad=\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \int_0^t\tilde{E}\bigg[\bigg(U_N\bigg(\bigg[\frac{s}{\delta t}\bigg]\delta t\bigg),\psi^2_{[s/\delta t]\,\delta t}\|h\|_2\bigg)\bigg]\,ds.\end{aligned}\tag{52}$$

But since $\tilde{E}[U_N(i\,\delta t)|\mathcal{Y}]=p_{i\,\delta t}$ for $i=0,1,\dots$ (Proposition 9), we finally get, using (51) and (52), that

$$\begin{aligned}&\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} N\sqrt{\delta t}\,\tilde{E}\big[\big((U_N(t),\varphi)-p_t(\varphi)\big)^2\big]\\ &\quad=\sqrt{\frac{2}{\pi}}\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}}\int_0^t\tilde{E}\big[p_{[s/\delta t]\,\delta t}\big(\psi^2_{[s/\delta t]\,\delta t}\|h\|_2\big)\big]\,ds\\ &\quad=\sqrt{\frac{2}{\pi}}\int_0^t\tilde{E}\big[p_s\big(\psi_s^2\|h\|_2\big)\big]\,ds\end{aligned}$$

by using the dominated convergence theorem with the upper bound

$$\tilde{E}\big[p_{[s/\delta t]\delta t}(\psi^2_{[s/\delta t]\delta t}\|h\|_2)\big] \le \|h\|\sqrt{m}\,\tilde{E}\Big[\sup_{s\in[0,t]}\|\psi_s\|^2_\infty\Big]. \qquad \square$$

COROLLARY 11. *If the interbranching times $\delta t = \delta t(N)$ are chosen so that $\lim_{N\to\infty} N(\delta t)^{3/2} = 0$, then the limit* (45) *holds true for arbitrary $t$* (*not just for $t$ being an integer multiple of $\delta t$*).

PROOF. From Proposition 9, we get that $p_t(\varphi) = p_{[t/\delta t]\delta t}(\psi_{[t/\delta t]\delta t})$. Also, from (30), it is easy to show that $\tilde{E}[((U_N(t),\varphi) - (U_N([t/\delta t]\delta t), \psi_{[t/\delta t]\delta t}))^2]$ is of order $O(\delta t)$. Hence, since we assumed that $\lim_{N\to\infty} N(\delta t)^{3/2} = 0$, it follows that

$$\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} N\sqrt{\delta t}\,\tilde{E}\Big[\Big((U_N(t),\varphi) - \Big(U_N\Big(\Big[\frac{t}{\delta t}\Big]\delta t\Big), \psi_{[t/\delta t]\delta t}\Big)\Big)^2\Big] = 0. \tag{53}$$

Also, from the proof of Theorem 10, it follows that

$$\begin{aligned}\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} \Bigg| N\sqrt{\delta t}\,\tilde{E}\Big[\Big(\Big(U_N\Big(\Big[\frac{t}{\delta t}\Big]\delta t\Big), \psi_{[t/\delta t]\delta t}\Big) - p_{[t/\delta t]\delta t}(\psi_{[t/\delta t]\delta t})\Big)^2\Big] \\ - \sqrt{\frac{2}{\pi}}\int_0^{[t/\delta t]\delta t} \tilde{E}[p_s(\psi_s^2\|h\|_2)]\,ds\Bigg| = 0.\end{aligned} \tag{54}$$

The claim then follows from (53) and (54). $\square$

COROLLARY 12. *If the length of the interbranching times is $1/N^\alpha$, where $\alpha \in (\frac{2}{3}, 2)$, then*

$$\begin{aligned}&\lim_{N\to\infty} N^{1-\alpha/2}\tilde{E}\big[\big((U_N(t),\varphi) - p_t(\varphi)\big)^2\big] \\ &\quad = \sqrt{\frac{2}{\pi}}\int_0^t \tilde{E}[p_s(\psi_s^2\|h\|_2)]\,ds.\end{aligned}$$

*In particular, if $\delta t = \delta t(N) = 1/N$, then*

$$\begin{aligned}&\lim_{N\to\infty} \sqrt{N}\,\tilde{E}\big[\big((U_N(t),\varphi) - p_t(\varphi)\big)^2\big] \\ &\quad = \sqrt{\frac{2}{\pi}}\int_0^t \tilde{E}[p_s(\psi_s^2\|h\|_2)]\,ds.\end{aligned} \tag{55}$$

PROOF. Direct consequence of the previous corollary. $\square$

It is clear now that the upper bound proved in [4] was sharp. However, the choice $\delta t = 1/N$ for the length of the interbranching times is suboptimal. We chose

$\delta t = 1/N$ in [4] because this particle approximation was inspired by the Dawson–Watanabe construction of a superprocess. We discussed this in greater detail in an earlier paper [6], where we followed closely the original construction and used branching mechanisms with fixed variance and not with minimal variance. In effect, in [6] we construct a superprocess in a random environment, the environment being the given trajectory of the observation process $Y$. By choosing a fixed value for the variance, we introduce much more randomness into the system and, as a result, the limiting process is not the solution of the Zakai equation, but rather a measure-valued process whose conditional expectation, given the environment $Y$, is $p_t$.

As we saw in the last corollary, the larger the length of the interbranching times is, the better the rate is. However, the order of the length of the interbranching times cannot be larger than $1/N^{2/3}$ as the last part of the evolution of the system is not corrected and hence a bias is introduced. We can resolve this problem by attaching weights to the particles, weights that only correspond to this last part of their path, that is, the path corresponding to the interval between the last branching time and the current time. The result is a partially weighted approximation that converges to $p_t$ no matter how large the interbranching times are (see the next section for details).

At the other end of the spectrum, if the interbranching times are of order $1/N^2$ (hence $N\sqrt{\delta t} \nrightarrow \infty$), then $U_N$ no longer converges to $p_t$. By branching so often, the randomness introduced in the system at branching times overpowers the corrective effect and, as a result, just as in the case when the branching variance is fixed, the limiting process is a measure-valued process whose conditional expectation, given the environment $Y$, is $p_t$.

**5. Improved branching algorithms.** We saw in the last section that the branching particle approximation has the asymptotic rate of convergence $c(t)/N^{1-\alpha/2}$, when the interbranching times are of order $1/N^{\alpha}$ and $c(t)$ is the constant

$$c(t) \triangleq \sqrt{\frac{2}{\pi}} \int_0^t \tilde{E}\big[p_s(\psi_s^2 \|h\|_2)\big]\,ds.$$

By comparison, the rate of convergence of $\Theta_N(t)$, the particle approximation obtained by using the Monte Carlo method, is of order $1/N$. As stated in the Introduction, $\Theta_N(t)$ is the approximation given by

$$\Theta_N(t) \triangleq \frac{1}{N}\sum_{i=1}^N \mu_i^t \delta_{\{V_i(t)\}},$$

where $V_1(t), V_2(t), \dots, V_M(t)$ are independent realizations of the signal (and independent of $Y$) and $\mu_i^t$ are their corresponding likelihoods/weights

$$\mu_i^t = \exp\biggl(\int_0^t h^*(V_i(t))\,dY_s - \frac{1}{2}\int_0^t \|hV_i(t)\|_2^2\,ds\biggr). \tag{56}$$

PROPOSITION 13. *If $\Theta_N(t)$ is the above approximation, then*

$$\tilde{E}\big[\big((\Theta_N(t),\varphi) - p_t(\varphi)\big)^2\big] = \frac{c_\Theta(t)}{N},$$

*where $c(t)$ is a constant independent of $N$, which admits the following representations*:

$$c_\Theta(t) = E\left[\exp\left(\int_0^t \|h(V_p)\|_2^2\,dp\right)(\varphi(V_t))^2\right] - E[p_t(\varphi)^2] \tag{57}$$

$$= \int_0^t \tilde{E}\left[\exp\left(\int_0^s \|h(V_p)\|_2^2\,dp\right)\|\sigma^* D\psi_s(V_s)\|_2^2\right]ds$$
$$+\,\tilde{E}[\pi_0(\psi_0^2) - \pi_0(\psi_0)^2], \tag{58}$$

*where $V_t$ is a realization of the signal process.*

PROOF. The representation (57) follows immediately, using the i.i.d. property of the random variables $V_i(t)$ and the identities

$$\tilde{E}[\varphi(V_i(t))\mu_i^t|\mathcal{Y}] = p_t(\varphi),$$
$$\tilde{E}\big[\big(\mu^t\varphi(V_t)\big)^2\big] = \tilde{E}\left[\exp\left(\int_0^t \|h(V_p)\|_2^2\,dp\right)(\varphi(V_t))^2\right].$$

Further, using Proposition 9 and identity (31), we deduce (58). □

From (45) and Proposition 13, it is clear now that the order of rate of convergence of the Monte Carlo approximation is better than the order of the approximation given by the branching algorithm as long as we choose asymptotically small interbranching times. As we stated at the end of Section 4, we can improve the branching algorithm by attaching weights to the particles, weights that only correspond to this last part of their path, that is, the path corresponding to the interval between the last branching time and the current time. Let $V_N^1(t),\dots,V_N^{m_N(t)}(t)$ be the positions of the $m_N(t)$ particles alive at time $t$ obtained using the branching algorithms and define

$$\bar{U}_N(t) = \frac{1}{N}\sum_{J=1}^{m_N(t)} \bar{\mu}_{j,N}^t \delta_{\{V_N^j(t)\}},$$

where $\bar{\mu}_{j,N}^t$ are defined as

$$\bar{\mu}_{j,N}^t = \exp\left(\int_{[t/\delta]\,\delta t}^t h^*\big(V_N^j(t)\big)\,dY_s - \tfrac{1}{2}\int_{[t/\delta]\,\delta t}^t \big\|h\big(V_N^j(t)\big)\big\|_2^2\,ds\right).$$

Let us observe that $\bar{U}_N(t) = U_N(t)$ if $t$ is a branching time, that is, $t = i\,\delta t$. Hence, if we are only interested in the value of the approximation at branching times, then we cannot distinguish between the two. This remark is important as in the following we will show that we can get better rates of convergence for $\bar{U}_N(t)$ than for $U_N(t)$, and hence these rates will also apply to $U_N(t)$ if $t$ is always a branching time regardless of $N$.

PROPOSITION 14. *If $\varphi$ is chosen so that* (20) *is satisfied, then* $\{\bar{U}_N((r), \psi_r), \mathcal{F}_r \vee \mathcal{Y}\}$, $r \in [0, t]$, *is a square-integrable martingale. In particular,* $\tilde{E}[\bar{U}_N(t)|\mathcal{Y}] = p_t$ *for all* $t \geq 0$.

PROOF. Similar to that of Proposition 9. □

The following theorem shows that $\bar{U}_N(t)$ has the same asymptotic rate of convergence as $U_N(t)$. More important, now we no longer need to impose the constraint that the interbranching times should be larger (in order) than $1/N^{2/3}$. Hence, we can obtain rates of convergence as close to $1/N$ as we want to, as the corollary following the theorem shows. *It can be also interpreted as an interpolation result between the case when the interbranching times converge to* 0 *as the number of particles increases and the case when the interbranching times are kept fixed regardless of the number of particles.*

THEOREM 15. *If $\varphi$ is chosen so that* (20) *is satisfied and $h$ is assumed to be Lipschitz, then*

$$\lim_{\substack{N\to\infty,\delta t\to 0\\ N\sqrt{\delta t}\to\infty}} N\sqrt{\delta t}\,\tilde{E}\big[\big((\bar{U}_N(t),\varphi) - p_t(\varphi)\big)^2\big] = \sqrt{\frac{2}{\pi}}\int_0^t \tilde{E}[p_s(\psi_s^2\|h\|_2)]\,ds. \tag{59}$$

PROOF. From Propositions 9 and 14, it follows that

$$\begin{aligned}
&\tilde{E}\big[\big((\bar{U}_N(t),\varphi) - p_t(\varphi)\big)^2\big]\\
&\quad = \tilde{E}\big[(\bar{U}_N(t),\varphi)^2\big] - \tilde{E}[p_t(\varphi)^2]\\
&\quad = \tilde{E}\left[\left((\bar{U}_N(t),\varphi) - \left(\bar{U}_N\left(\left[\frac{t}{\delta t}\right]\delta t\right), \psi_{[t/\delta t]\,\delta t}\right)\right)^2\right]\\
&\qquad + \tilde{E}\left[\left(\left(\bar{U}_N\left(\left[\frac{t}{\delta t}\right]\delta t\right), \psi_{[t/\delta t]\,\delta t}\right) - p_{[t/\delta t]\,\delta t}(\psi_{[t/\delta t]\,\delta t})\right)^2\right].
\end{aligned}$$

From (31), one proves that the term $\tilde{E}[((\bar{U}_N(t),\varphi) - (\bar{U}_N([t/\delta t]\,\delta t), \psi_{[t/\delta t]\,\delta t}))^2]$ is of order $\delta t/N$. Then, since $\bar{U}_N([t/\delta t]\,\delta t) = U_N([t/\delta t]\,\delta t)$, as the weights $\bar{\mu}^t_{j,N}$ are reinitialized to 1 at branching times, we have, following (54), the required asymptotic rate for $\bar{U}_N(t)$. □

COROLLARY 16. *If the length of the interbranching times is* $1/N^{\alpha}$, *where* $\alpha \in (0,2)$, *then*

$$\lim_{N\to\infty} N^{1-\alpha/2}\tilde{E}\big[\big((U_N(t),\varphi)-p_t(\varphi)\big)^2\big]=\sqrt{\frac{2}{\pi}}\int_0^t \tilde{E}\big[p_s(\psi_s^2\|h\|_2)\big]\,ds.$$

PROOF. Direct consequence of the previous corollary. □

For the following theorem, we need to introduce a couple of functions. First, let us observe that, based on the independent increments property of $Y$ and $B^{j,N}$ (the Brownian path that generates $V_N^j$), we have

$$\begin{aligned}&\tilde{E}\bigg[\exp\bigg(\int_{i\,\delta t}^{s}\|h(V_N^j(s))\|_2^2\,dp\bigg)\|\sigma^* D\psi_s(V_N^j(s))\|_2^2|\mathcal{F}_{i\,\delta t}\bigg]\\ &\qquad=\tilde{E}\bigg[\exp\bigg(\int_{i\,\delta t}^{s}\|h(V_N^j(s))\|_2^2\,dp\bigg)\|\sigma^* D\psi_s(V_N^j(s))\|_2^2|V_N^j(i\,\delta t)\bigg]\end{aligned}$$

for $s\in[i\,\delta t,(i+1)\,\delta t]$. Hence, there exists a measurable function $\Upsilon_s^{i\,\delta t}(x)$ such that

$$\tilde{E}\bigg[\exp\bigg(\int_{i\,\delta t}^{s}\|h(V_N^j(s))\|_2^2\,dp\bigg)\|\sigma^* D\psi_s(V_N^j(s))\|_2^2|\mathcal{F}_{i\,\delta t}\bigg]=\Upsilon_s^{i\,\delta t}\big(V_N^j(i\,\delta t)\big). \tag{60}$$

Similarly, there exists a measurable function $\hat{\Upsilon}^{i\,\delta t}(x)$ such that

$$\tilde{E}\big[\big(\psi_{(i+1)\,\delta t}(V_N^j((i+1)\,\delta t))\big)^2\nu_N^i(V_N^j)|\mathcal{F}_{i\,\delta t}\big]=\hat{\Upsilon}^{i\,\delta t}\big(V_N^j(i\,\delta t)\big). \tag{61}$$

THEOREM 17. *If* (20) *is satisfied and* $h$ *is assumed to be Lipschitz*, *then*

$$\tilde{E}\big[\big((\bar{U}_N(t),\varphi)-p_t(\varphi)\big)^2\big]=\frac{\bar{c}(t)}{N},$$

*where* $\bar{c}(t)$ *is a constant independent of* $N$ *with the following representation*:

$$\begin{aligned}\bar{c}(t)=&\int_0^t \tilde{E}\big[p_{[t/\delta t]\,\delta t}(\Upsilon_s^{[t/\delta t]\,\delta t})\big]\,ds+\sum_{i=1}^{[t/\delta t]}\tilde{E}\big[p_{i\,\delta t}(\hat{\Upsilon}^{i\,\delta t})\big]\\ &+E\big[\pi_0(\psi_0^2)-\pi_0(\psi_0)^2\big].\end{aligned} \tag{62}$$

PROOF. Since $t\to(\bar{U}_N(t),\psi_t)$ is a square-integrable martingale,

$$\begin{aligned}&\tilde{E}\big[\big((\bar{U}_N(t),\varphi)-(\bar{U}_N(0),\psi_0)\big)^2\big]\\ &\qquad=\tilde{E}\bigg[\bigg((\bar{U}_N(t),\varphi)-\bigg(\bar{U}_N\bigg(\bigg[\frac{t}{\delta t}\bigg]\delta t\bigg),\psi_{[t/\delta t]\,\delta t}\bigg)\bigg)^2\bigg]\\ &\qquad\quad+\sum_{i=0}^{[t/\delta t]-1}\tilde{E}\big[\big((U_N((i+1)\,\delta t),\psi_{(i+1)\,\delta t})-(U_N(i\,\delta t),\psi_{i\,\delta t})\big)^2\big]\end{aligned} \tag{63}$$

$$
\begin{aligned}
&+ \tilde{E}\big[\big((\bar{U}_N(0), \psi_0) - \pi_0(\psi_0)\big)^2\big] \\
&= \tilde{E}\bigg[\bigg((\bar{U}_N(t), \varphi) - \bigg(\bar{U}_N\bigg(\bigg[\frac{t}{\delta t}\bigg]\delta t\bigg), \psi_{[t/\delta t]\,\delta t}\bigg)\bigg)^2\bigg] \\
&\quad + \sum_{i=0}^{[t/\delta t]-1} \tilde{E}\big[\big(\big(U_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) - \big(U_N(i\,\delta t), \psi_{i\,\delta t}\big)\big)^2\big] \\
&\quad + \frac{E[\pi_0(\psi_0^2) - \pi_0(\psi_0)^2]}{N}.
\end{aligned}
$$

Then, using yet again (31), one shows that

$$
\begin{aligned}
&\tilde{E}\bigg[\bigg((\bar{U}_N(t), \varphi) - \bigg(\bar{U}_N\bigg(\bigg[\frac{t}{\delta t}\bigg]\delta t\bigg), \psi_{[t/\delta t]\,\delta t}\bigg)\bigg)^2\bigg] \\
&\qquad = \frac{1}{N}\int_{[t/\delta t]\,\delta t}^{t} \tilde{E}\bigg[\bigg(U_N\bigg(\bigg[\frac{t}{\delta t}\bigg]\delta t\bigg), \Upsilon_s^{[t/\delta t]\,\delta t}\bigg)\bigg]\,ds \qquad (64)\\
&\qquad = \frac{1}{N}\int_{[t/\delta t]\,\delta t}^{t} \tilde{E}\big[p_{[t/\delta t]\,\delta t}\big(\Upsilon_s^{[t/\delta t]\,\delta t}\big)\big]\,ds.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\tilde{E}\Big[\big(\tilde{E}\big[\big(\bar{U}_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big)|\mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big] - \big(\bar{U}_N(i\,\delta t), \psi_{i\,\delta t}\big)\big)^2\Big] \\
&\qquad = \frac{1}{N}\int_{i\,\delta t}^{(i+1)\,\delta t} \tilde{E}\big[\big(U_N(i\,\delta t), \Upsilon_s^{i\,\delta t}\big)\big]\,ds \qquad (65)\\
&\qquad = \frac{1}{N}\int_{i\,\delta t}^{(i+1)\,\delta t} \tilde{E}\big[p_{i\,\delta t}\big(\Upsilon_s^{i\,\delta t}\big)\big]\,ds
\end{aligned}
$$

and

$$
\begin{aligned}
&\tilde{E}\Big[\big(\big(\bar{U}_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big) \\
&\qquad - \tilde{E}\big[\big(\bar{U}_N\big((i+1)\,\delta t\big), \psi_{(i+1)\,\delta t}\big)|\mathcal{F}_{(i+1)\,\delta t-} \vee \mathcal{Y}\big]\big)^2\Big] \qquad (66)\\
&\qquad = \frac{1}{N^2}\tilde{E}\Bigg[\sum_{j=1}^{m_N(i\,\delta t)} \big(\psi_{(i+1)\,\delta t}\big(V_N^j\big((i+1)\,\delta t\big)\big)\big)^2 \nu_n^i\big(V_N^j\big)\Bigg] = \frac{1}{N} p_{i\,\delta t}\big(\hat{\Upsilon}^{i\,\delta t}\big).
\end{aligned}
$$

Finally, by plugging (64)–(66) into (63), we get the required expectation. □

Hence, when the interbranching times have fixed size, then both the Monte Carlo approximation $\Theta_N(t)$ and the (weighted) branching approximation $\bar{U}_N(t)$ have the same order $1/N$.

A variation of the previous algorithm is to use random interbranching times. For example, let $V_1(t), \dots, V_{m_N(t)}(t)$ be the positions of the $m_N(t)$ particles alive at time $t$ obtained using the branching algorithms when the branching times are

$$\tau_1 = \inf\left\{t \geq 0; \max_{i=1,\dots,n} \exp\left(\int_0^t \|h(V_N^j(t))\|_2^2\, ds\right) \geq 2\right\},$$

$$\tau_{j+1} = \inf\left\{t \geq \tau_j; \max_{i=1,\dots,m_N(\tau_j)} \exp\left(\int_{\tau i}^t \|h(V_N^j(t))\|_2^2\, ds\right) \geq 2\right\}, \qquad j = 1, 2, \dots.$$

Then define $\hat{U}_N(t) \triangleq (1/N)\sum_{i=1}^{m_N(t)} \hat{\mu}_i^t \delta_{\{V_i(t)\}}$, where, for $t \in [\tau_n, \tau_{n+1})$, $\hat{\mu}_i^t$, $i = 1, 2, \dots, m_N(t)$, are defined as $\hat{\mu}_i^t = \exp(\int_{\tau_N}^t h^*(V_i(t))\, dY_s - \frac{1}{2}\int_{\tau_N}^t \|V_i(t)\|_2^2\, ds)$. The reason we chose the particular bound 2 is because at $\tau_j$ there will be at least one particle so that $\tilde{E}[(\hat{\mu}_i^t)^2 | V_i] = \exp(\int_{\tau i}^t \|h(V_N^j(s))\|_2^2\, ds) = 2$. Hence, on average, $\hat{\mu}_i^t$ is "considerably" larger than 1, so it is time to branch (we can replace the lower bound 2 by any constant $k > 1$). Then it is easy to see that, for all $j = 0, 1, \dots,$ there exist $i = i_j$ such that

$$\exp(d\|h\|^2(\tau_{i+1} - \tau_i)) \geq \exp\left(\int_{\tau i}^{\tau_{i+1}} \|h(V_{i_j}(t))\|_2^2\, ds\right) = 2.$$

Thus, $\tau_{i+1} - \tau_i \geq \ln 2/d\|h\|^2$. Therefore, we have only a finite number of branching times and we would obtain the same upper bound as the upper bound corresponding to $\bar{U}_N(t)$ with *deterministic* interbranching times $\ln 2/d\|h\|^2$. But we should expect $\hat{U}_N(t)$ to perform better since we introduce less randomness into the system by branching less often.

Finally, we can modify the previous algorithm and only branch those particles whose corresponding weights reach 2. More precisely, let $V_1(0), \dots, V_{m_N(t)}(t)$ be the positions of the $m_N(t)$ particles alive at time $t$ and let $\hat{\mu}_i^t$, $i = 1, 2, \dots, m_N(t)$, be their corresponding weights

$$\hat{\mu}_i^t = \exp\left(\int_{\tau_{t,V_i}}^t h^*(V_i(t))\, dY_s - \frac{1}{2}\int_{\tau_{t,V_i}}^t \|V_i(t)\|_2^2\, ds\right),$$

where $\tau_{t,V_i}$ is the last time before time $t$, when the $i$th particle branched. The $i$th particle will branch again the first time that $\hat{\mu}_i^t$ is equal to 2. At branching times, each particle splits into two particles with the same positions as their mother's. As before, in between branchings, the particles follow the equation of the signal. In this case, there will be no errors introduced at branching time with the added bonus that all the weights will stay bounded by 2. The drawback is that particles with small weights are not eliminated and it is likely that the number of particles will increase exponentially (one can prove that the number of particles does not explode in finite time).

## REFERENCES

[1] BENSOUSSAN, A. (1992). *Stochastic Control of Partially Observable Systems*. Cambridge Univ. Press.

[2] CRISAN, D. (1996). The problem of nonlinear filtering. Ph.D. thesis, Edinburgh.

[3] CRISAN, D., DEL MORAL, P. and LYONS, T. (1999). Discrete filtering using branching and interacting particle systems. *Markov Process. Related Fields* **5** 293–318.

[4] CRISAN, D., GAINES, J. and LYONS, T. (1998). Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Probab.* **58** 1568–1591.

[5] CRISAN, D. and LYONS, T. (1997). Nonlinear filtering and measure valued processes. *Probab. Theory Related Fields* **109** 217–244.

[6] CRISAN, D. and LYONS, T. (1999). A particle approximation of the solution of the Kushner–Stratonovitch equation. *Probab. Theory Related Fields* **115** 549–578.

[7] DOUCET, A., DE FREITAS, N. and GORDON, N. (2001). *Sequential Monte Carlo Methods in Practice*. *Statistics for Engineering and Information Science*. Springer, New York.

[8] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.

[9] FUJISAKI, M., KALLIANPUR, G. and KUNITA, H. (1972). Stochastic differential equations for the non linear filtering problem. *Osaka J. Math.* **9** 19–40.

[10] KALLIANPUR, G. and STRIEBEL, C. (1968). Estimation of stochastic systems: Arbitrary system process with additive white noise observation errors. *Ann. Math. Statist.* **39** 785–801.

[11] KARATZAS, I. and SHREVE, S. E. (1988). *Brownian Motion and Stochastic Calculus*. Springer, New York.

[12] KRYLOV, N. V. and ROZOVSKII, B. L. (1978). Conditional distributions of diffusion processes. *Izv. Akad. Nauk SSSR Ser. Mat.* **42** 356–378.

[13] KURTZ, T. G. and OCONE, D. L. (1988). Unique characterisation of conditional distributions in nonlinear filtering. *Ann. Probab.* **16** 80–107.

[14] KURTZ, T. G. and XIONG, J. (1999). Particle representation for a class of nonlinear SPDE's. *Stochastic Process. Appl.* **83** 103–106.

[15] KURTZ, T. G. and XIONG, J. (2001). Numerical solutions for a class of SPDEs with application to filtering. In *Stochastics in Finite and Infinite Dimensions* (T. Hida, R. L. Karandikar, H. Kunita, B. S. Rajput, S. Watanabe and J. Xiong, eds.) 233–258. Birkhäuser, Boston.

[16] KUSHNER, H. J. (1967). Dynamical equations for optimal nonlinear filtering. *J. Differential Equations* **3** 179–190.

[17] PARDOUX, E. (1979). Stochastic partial differential equations and filtering of diffusion processes. *Stochastics* **3** 127–167.

[18] PARDOUX, E. (1991). Filtrage non linéaire et equations aux dérivées partielles stochastiques associées. *Ecole d'Eté de Probabilités de Saint-Flour XIX. Lecture Notes in Math.* **1464**. Springer, New York.

[19] YOR, M. (1977). Sur les théories du filtrage et de la prédiction. *Séminaire de Probabilités XI. Lecture Notes in Math.* **581** 257–297. Springer, New York.

[20] ZAKAI, M. (1969). On the optimal filtering of diffusion processes. *Z. Wahrsch. Verw. Gebiete* **11** 230–243.

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE
180 QUEEN'S GATE
LONDON SW7 2BZ
UNITED KINGDOM
E-MAIL: d.crisan@ic.ac.uk