SIAM J. APPL. MATH.
Vol. 58, No. 5, pp. 1568–1590, October 1998

009

CONVERGENCE OF A BRANCHING PARTICLE METHOD TO THE SOLUTION OF THE ZAKAI EQUATION*

DAN CRISAN†, JESSICA GAINES‡, AND TERRY LYONS†

Abstract. We construct a sequence of branching particle systems U_n convergent in distribution to the solution of the Zakai equation. The algorithm based on this result can be used to solve numerically the filtering problem. The result is an improvement of the one presented in a recent paper [Crisan and T. Lyons, *Prob. Theory Related Fields*, 109 (1997), pp. 217–244], because it eliminates the extra degree of randomness introduced there.

Key words. nonlinear filtering, Zakai equation, particle filters, numerical solutions, stochastic partial differential equations

AMS subject classifications. 93E11, 60G57, 65U05

PII. S0036139996307371

1. Introduction. In its standard formulation ([3], [28], etc.), filtering theory studies the problem of estimating a "signal" process $X = \{X_t; t \geq 0\}$, given "observations"

$$Y_t = \int_0^t h(X_s)ds + W_t, \qquad t \geq 0,$$

where $\{W_t; t \geq 0\}$ is a standard Brownian motion; in other words, the conditional distribution of X_t, given $\mathcal{Y}_t = \sigma\{Y_s; s \in [0,t]\}$. In practical applications, one is interested in computing estimates of the form $\pi_t(\varphi) := E[\varphi(X_t)|\mathcal{Y}_t]$. Using an appropriate change of measure one can show that, in the case of X being a Markov process with generator A and evolving independently of W,

$$\pi_t(\varphi) = \frac{p_t(\varphi)}{p_t(1)},$$

where $p_t(\varphi)$ is the solution of the Zakai equation

$$dp_t(\varphi) = p_t(A\varphi)dt + p_t(h^*\varphi)dY_t.$$

In the following we will construct a sequence of branching particle systems U_n, whose laws will approximate p_t, i.e.,

$$\lim_{n\to\infty} (U_n(t), \varphi) = p_t(\varphi).$$

The particles will move according to the law of the signal, independently of each other and after fixed-length intervals will branch. The mean number of offspring of a

*Received by the editors July 29, 1996; accepted for publication (in revised form) June 13, 1997; published electronically July 17, 1998.
http://www.siam.org/journals/siap/58-5/30737.html

†Department of Mathematics, Imperial College, 180 Queen's Gate, London, SW7 2BZ, United Kingdom (d.crisan@ic.ac.uk, t.lyons@ic.ac.uk). The research of the first author was supported by EPSRC grant GR/K69704 and by the British Council. The research of third author was supported by EPSRC senior fellowship B/93/SF/445, EPSRC grant GR/K69704, and EEC grant SC1∗-CT92-0784.

‡Department of Mathematics, University of Edinburgh, JCMB, KB, Mayfield Road, Edinburgh, EH9 3JZ, United Kingdom (jessica@maths.ed.ac.uk).

particle will depend on the last part of its trajectory and on the observation process; the variance of the branching mechanism will be the minimum possible one.

We can use these particle systems to solve numerically the filtering problem since, as the number of particles is increased, the empirical measure associated with the cloud of particles converges to the solution of the Zakai equation. By starting with a number of particles constituting a sample approximating the initial distribution of X, we allow the system to evolve up to time t and then use the empirical law to estimate the distribution of the required statistic φ.

In [10], we constructed a sequence $\mathcal{X}_n$ of similar branching particle systems also solving the Zakai equation. In that case, the variance of the branching mechanism was a priori given. In this way, we introduced an extra degree of randomness and so only the (conditional) expectation of that sequence converged to p_t. Therefore, we needed a whole set of copies of the particle system in order to obtain a good approximation to the solution of the Zakai equation. Our new approach converges directly to p_t, and we don't need to estimate an average.

In the next section we give a more detailed description of the nonlinear filtering problem and set up the notation we will use subsequently. In the third section, we construct the branching particle system, which will approximate the solution of the Zakai equation, and prove several useful remarks. In the fourth section we prove that the laws of the systems are tight. In the fifth section, we show that $U_n(t)$ is convergent in measure (and hence in distribution) to the solution of the Zakai equation (p_t). In the sixth section, we describe the numerical implementation of the algorithm based on the approximating particle system, and in the last section we use the implementation to solve particular examples.

In this paper we analyze the numerical solution of the Zakai equation. There is evidence that still more efficient algorithms could be built using measure valued processes of minimal variance which directly approximate the solution of the Kushner–Stratonovitch equation. The reader is directed to [11] for details of this approach.

We do not want to give a comprehensive survey of other approaches here but should mention certain works which provide contrasting approaches to the numerical analysis of the standard filtering problem. These can be roughly classified into the following.

- **Linearization methods (extended Kalman filter).** The nonlinear problem is linearized within small time frames and then the linear (Kalman) filter is applied to it (see [28]).
- **Approximation by finite-dimensional nonlinear filters.** The filtering problem is solved using approximation by certain exact nonlinear filters (see [2], [12], [30]).
- **Particle methods.** The conditional measure is approximated by a system of particles on which a resampling procedure is applied every time a new observation is available (see [17], [13]).
- **Classical partial differential equation (PDE) methods.** The associated stochastic PDE (the Zakai equation), whose solution is the (unnormalized) density of the conditional distribution of the signal, is solved (see [4], [5], [15], [20],[25],[31]).
- **Wiener chaos expansions.** The Zakai equation is solved using a decomposition of its solution into Wiener integrals (see [22],[23]).
- **Moment methods.** The conditional distribution is approximated using its moments (see [6]).

Using the language of Glowinski and Sun [31], one might summarize this paper as introducing an implementable algorithm of particle type and proving that it has good convergence properties.

One point where we differ is that we prove estimates for the L^2 convergence of statistics and not of densities. This seems a more appropriate measure of convergence, as densities have little meaning or practical use in high dimensions and, in any case, would give an infinite distance between any particle approximation and the posterior distribution despite the clear, practical observation that particle approximations to posterior distributions can be very informative.

2. The nonlinear filtering problem. Let $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$ be a complete filtered probability space (we denote by $\mathcal{N}$ the collection of P-null sets), and let E be a locally compact complete separable metric space (a locally compact Polish space). We will denote by $C_b(E)$ the space of continuous bounded functions on E denote and by $C_0(E)$ the space of continuous functions which vanish at infinity, both endowed with the supremum norm $\|\cdot\|$.

Let $M_F(E)$ be the space of finite measures over E endowed with the topology of weak convergence, i.e., the topology in which $\mu_n \to \mu$ iff $(\mu_n, f) \to (\mu, f)$ for all $f \in C_b(E)$, and let $M_F'(E)$ be the space of finite measures over E endowed with the topology of vague convergence, i.e., the topology in which $\mu_n \to \mu$ iff $(\mu_n, f) \to (\mu, f)$ for all $f \in C_0(E)$. Let also $D_{M_F(E)}[0,1]$, $D_{M_F'(E)}[0,1]$ be the space of cádlág paths defined on $[0,1]$ with values in $M_F(E)$, respectively, $M_F'(E)$, endowed with the Skorohod topology.

Let $X = \{X_t, \mathcal{F}_t; t \geq 0\}$ be a diffusion process (the "signal" process), solution of the martingale problem associated with the infinitesimal generator A and the initial distribution $\pi_0 \in P(E)$. We assume that A is an operator on $C_b(E)$ with the domain $\mathcal{D}(A)$, $1 \in \mathcal{D}(A)$, and $A1 = 0$. Let A_0 be the restriction of A to $C_0(E)$. We assume that the domain of A_0 is a *dense algebra* in $C_0(E)$ (in particular that if $f \in \mathcal{D}(A_0)$, then $f^2 \in \mathcal{D}(A_0)$) and that the martingale problem for A_0 is well posed.

Let also $W = \{W_t, \mathcal{F}_t; t \geq 0\}$ be an m-dimensional standard Brownian motion, independent of X, and let $Y = \{(Y_t, \mathcal{F}_t; t \geq 0\}$ be the $\mathbb{R}^m$-valued process defined by the following formula:

$$Y_t = \int_0^t h(X_s)ds + W_t, \qquad t \geq 0, \tag{1}$$

where $h : E \to \mathbb{R}^m$ is a continuous, bounded Borel measurable function, and denote

$$\|h\| = \sup_{x \in E, 1 \leq i \leq m} |h_i(x)| < \infty.$$

The process Y is usually called the "observation" process. We denote by $\mathcal{Y}_t \stackrel{\text{def}}{=} \sigma(Y_s, 0 \leq s \leq t) \vee \mathcal{N}$. The filtering problem consists, basically, of computing

$$\pi_t(\varphi) \stackrel{\text{def}}{=} E[\varphi(X_t)|\mathcal{Y}_t], \qquad \forall t \geq 0, \;\; \varphi \in B(E).$$

To do this, one changes the measure so that Y_t becomes a Brownian motion under the new probability measure $\tilde{P}$, independent of X, the law of X remains unchanged, and

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

where

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

and $\tilde{E}$ is the expectation with respect to the new probability measure.

One can prove that p_t satisfies the following evolution equation, called the *Zakai* equation:

$$p_t(\varphi) = \pi_0(\varphi) + \int_0^t p_s(A\varphi)ds + \int_0^t p_s(h^*\varphi)dY_s, \quad \tilde{P}-\text{a.s.}, \quad \forall\varphi \in \mathcal{D}(A). \tag{3}$$

We note that p_t can be viewed as an $M_F(E)$-valued process defined by formula (2) (see [24], [32]). In fact, (3) uniquely identifies p_t as a measure valued process since we have the following uniqueness theorem (cf. [24]).

THEOREM 2.1. *Under the condition set up above, if $U(t)$ is a $\mathcal{Y}_t$-adapted, cádlág, $M_F(E)$-valued process satisfying, for all $t \le T$,*

$$(U(t),\varphi) = (\pi_0,\varphi) + \int_0^t (U(s),A\varphi)ds + \int_0^t (U(s),h^*\varphi)dY_s,\ a.s.,\ \forall\varphi \in \mathcal{D}(A_0)\cup\{1\},$$

then $U(t) = p_t$ for $t \le T$, a.s., as $M_F(E)$-valued processes.

We did not define the filtering problem, under the broadest possible terms. Indeed, h need not be bounded nor X be a diffusion. Also, both A and h can be taken to be time dependent, and W does not need to be independent of X. There is vast literature concerning the filtering problem; here are just a few articles and monographs: [3], [16], [18], [21], [28]. We restricted ourselves to the above framework because this was the most convenient one for the analysis involved in this paper. In a sequel to this paper, we will give the most general framework under which our result still holds.

3. The branching particle systems U_n. From now on, we work under the new probability measure $\tilde{P}$, and all the expectations will be considered with respect to $\tilde{P}$. We will construct the particle systems and, implicitly, prove the convergence results, up to a fixed horizon, i.e., on the fixed interval $[0,1]$, the construction being identical for any interval $[0,T]$. Then, using an argument based on the Carathèodory extension theorem, the construction can be extended for the whole positive axis.

Let $\{U_n(t),\mathcal{F}_t;\ 0 \le t \le 1\}$ be a sequence of branching particle systems on $(\Omega,\mathcal{F},\tilde{P})$ with values in $M_F(E)$ defined as follows.

Initial condition.

1. $U_n(0)$ is the empirical measure of n particles of mass $\frac{1}{n}$, i.e., $U_n(0) = \frac{1}{n}\sum_{i=1}^n \delta_{x_i^n}$, where $x_i^n \in E$, for every $i,n \in \mathbb{N}$.

2. The empirical measure of the particles tends weakly to π_0.

Evolution in time.

We describe the evolution of the processes on the interval $\left[\frac{i}{n},\frac{i+1}{n}\right]$, $i = 0,1,\dots,n-1$.

1. At time $\frac{i}{n}$, the process consists of the occupation measure of $m_n(\frac{i}{n})$ particles of mass $\frac{1}{n}$ (we will denote the number of particles alive at time t by $m_n(t)$).

2. During the interval the particles move independently with the same law as the signal X. Let $V(s)$, $s \in [\frac{i}{n},\frac{i+1}{n})$ be the trajectory of a generic particle in this interval.

3. At the end of the interval, each particle branches into a random number of particles with a mechanism depending on its trajectory in the interval. The mechanism

is chosen so that it has finite second moment; the mean number of offspring for a particle given the σ-field $\mathcal{F}_{\frac{i+1}{n}-} = \sigma(\mathcal{F}_s,\, s < \frac{i+1}{n})$ of events up to time $\frac{i+1}{n}$ is

$$\mu_n^i(V) \triangleq \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*(V(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} h^*h(V(t))dt\right) \tag{4}$$

and so that the variance $\nu_n^i(V)$ is minimal consistent with the number of offspring being an integer. The particles branch independently of each other, given $\mathcal{F}_{\frac{i+1}{n}-}$.

We remark that $\nu_n^i(V) = (\mu_n^i(V) - [\mu_n^i(V)])([\mu_n^i(V)] + 1 - \mu_n^i(V))$ and so is always less than $\frac{1}{4}$ ($[x]$ is the largest integer smaller that x).

We now make some preliminary estimates before showing the convergence of U_n in the next section. Just before the $(i+1)$th branching, we will have $m_n(\frac{i}{n})$ particles. Let us denote by $U_n(\frac{i+1}{n}-)$ the state of the process just before the $(i+1)$th branching and denote by $V_n^j(s)$, $s \in [\frac{i}{n}, \frac{i+1}{n})$ the trajectory of the jth particle alive during the interval. Let also $q_n^j(\frac{i+1}{n})$ be the number of offspring of the jth particle with $1 \le j \le m_n(\frac{i}{n})$ at time $\frac{i+1}{n}$.

PROPOSITION 3.1. *We have the following a priori upper bounds.*

(i). $\tilde{E}[m_n(t)] = n$, $\forall n \ge 0$, $t \in [0,1]$.

(ii). $\tilde{E}[m_n^2(t)] \le n^2 e^{\|h\|^2 \frac{[nt]}{n}} + \frac{1}{4}\sum_{k \le [nt]} e^{\frac{k\|h\|^2}{n}}$, $\forall n \ge 0$, $t \in [0,1]$.

Proof. (i) The number of particles does not change during the intervals $(\frac{k}{n}, \frac{k+1}{n})$, $k = 1, \dots, n-1$, so $m_n(t) = m_n(\frac{[tn]}{n})$. Therefore it suffices to prove that $\tilde{E}[m_n(\frac{i}{n})] = \tilde{E}[m_n(\frac{(i+1)}{n})]$ for $0 \le i < n$. Using (4), we have

$$\begin{aligned}
\tilde{E}\left[m_n\left(\frac{i+1}{n}\right)\right] &= \tilde{E}\left[\sum_{j=1}^{m_n(\frac{i}{n})} q_n^j\left(\frac{i+1}{n}\right)\right] \\
&= \tilde{E}\left[\sum_{j=1}^{m_n(\frac{i}{n})} \tilde{E}\left[q_n^j\left(\frac{i+1}{n}\right)|\mathcal{F}_{\frac{i+1}{n}-}\right]\right] \\
&= \tilde{E}\left[\sum_{j=1}^{m_n(\frac{i}{n})} \mu_n^i(V_n^j)\right] \\
&= \tilde{E}\left[\sum_{j=1}^{m_n(\frac{i}{n})} \tilde{E}[\mu_n^i(V_n^j)|\mathcal{F}_{\frac{i}{n}}]\right] \\
&= \tilde{E}\left[m_n\left(\frac{i}{n}\right)\right]
\end{aligned}$$

since $\tilde{E}[\mu_n^i(V_n^j)|\mathcal{F}_{\frac{i}{n}}] = 1$ ($\mu_n^i(V_n^j)$ is the value at time $\frac{i+1}{n}$ of an exponential martingale identically equal to 1 at time $\frac{i}{n}$).

(ii) From the construction of the branching mechanism of the particles we have that

$$\begin{aligned}
\tilde{E}\left[\left(q_n^j\left(\frac{(i+1)}{n}\right)\right)^2 |\mathcal{F}_{\frac{i+1}{n}-}\right] &= \nu_n^i(V_n^j) + (\mu_n^i(V_n^j))^2 \\
&\le \frac{1}{4} + e^{\frac{\|h\|^2}{n}} \exp\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} 2h^*(V_n^j(t))dY_t - \frac{1}{2}\int_{\frac{i}{n}}^{\frac{i+1}{n}} 4h^*h(V_n^j(t))dt\right).
\end{aligned}$$

This inequality and the independence of the particles implies (as in (i))

$$\begin{aligned}\tilde{E}\quad\left[\left(m_n\left(\frac{i+1}{n}\right)\right)^2\right] &= \tilde{E}\left[\sum_{j=1}^{m_n(\frac{i}{n})}\tilde{E}\left[\tilde{E}\left[\left(q_n^j\left(\frac{i+1}{n}\right)\right)^2|\mathcal{F}_{\frac{i+1}{n}-}\right]\mathcal{F}_{\frac{i}{n}}\right]\right]\\ &+2\tilde{E}\left[\sum_{1\le j_1<j_2\le l}^{m_n(\frac{i}{n})}\tilde{E}[\mu_n^i(V_n^{j_1})\mu_n^i(V_n^{j_2})|\mathcal{F}_{\frac{i}{n}}]\right.\\ &\le \frac{1}{4}+e^{\frac{\|h\|^2}{n}}\tilde{E}\left[m_n\left(\frac{i}{n}\right)\right]+e^{\frac{\|h\|^2}{n}}\tilde{E}\left[m_n\left(\frac{i}{n}\right)\left(m_n\left(\frac{i}{n}\right)-1\right)\right].\end{aligned}$$

It follows that

$$\tilde{E}\left[\left(m_n\left(\frac{i+1}{n}\right)\right)^2\right]\le e^{\frac{\|h\|^2}{n}}\tilde{E}\left[\left(m_n\left(\frac{i}{n}\right)\right)^2\right]+\frac{1}{4};$$

hence

$$\tilde{E}[(m_n(t))^2] = \tilde{E}\left[\left(m_n\left(\frac{[tn]}{n}\right)\right)^2\right]\le e^{\|h\|^2\frac{[nt]}{n}}m_n^2(0)+\frac{1}{4}\sum_{k\le[nt]}e^{\frac{k\|h\|^2}{n}},$$

which completes the proof of the proposition. □

REMARK 3.2. *For any $t\in[0,1]$ and $\varphi\in B(E)$, the processes $(U_n(t),\varphi)$ are square integrable.*

Proof. Let V_1, V_2,..., $V_{m_n(t)}$ be the positions of the $m_n(t)$ particles alive at time t. Using Proposition 3.1, we have

$$\tilde{E}[(U_n(t),\varphi)^2] = \tilde{E}\left[\left(\frac{1}{n}\sum_{j=1}^{m_n(t)}\varphi(V_j)\right)^2\right]\le\frac{\|\varphi\|}{n^2}\tilde{E}[(m_n(t))^2]<\infty. \qquad \square$$

REMARK 3.3. *If $\varphi\in\mathcal{D}(A_0)\cup\{1\}$, then the process $(U_n(t),\varphi)$ satisfies the following evolution equation:*

$$(U_n(t),\varphi) = (U_n(0),\varphi)+\int_0^t(U_n(s),A\varphi)ds+S_n^\varphi(t)+M_n^\varphi([nt])$$

$$+\sum_{i=1}^{[nt]}\frac{1}{n}\sum_{j=1}^{m_n(\frac{i-1}{n})}\varphi\left(V_n^j\left(\frac{i}{n}\right)\right)(\mu_n^i(V_n^j)-1), \tag{5}$$

where $\{(S_n^\varphi(t),\,\mathcal{F}_t),\ t\in[0,1]\}$ is a square integrable martingale with quadratic variation

$$\langle S_n^\varphi\rangle(t) = \frac{1}{n}\int_0^t(U_n(s),A\varphi^2-2\varphi A\varphi)ds \tag{6}$$

and $\{(M_n^\varphi(l),\,\mathcal{F}_{\frac{l+1}{n}-}),\ l=0,1,\dots,n\}$ is a discrete martingale with (conditional) quadratic variation

$$\langle M_n^\varphi\rangle(l) = \frac{1}{n}\sum_{i=1}^{l}\left(U_n\left(\frac{i+1}{n}-\right),\nu_n^i\varphi^2\right). \tag{7}$$

Proof. From the construction of the particle systems we have that

$$\tilde{E}\left[\left(U_n\left(\frac{i+1}{n}\right),\varphi\right)|\mathcal{F}_{\frac{i+1}{n}-}\right]=\frac{1}{n}\sum_{i=1}^{m_n(\frac{i}{n})}\varphi\left(V_n^j\left(\frac{i+1}{n}\right)\right)\mu_n^i(V_n^j) \tag{8}$$

and also

$$\tilde{E}\left[\left(U_n\left(\frac{i+1}{n}\right),\varphi\right)^2|\mathcal{F}_{\frac{i+1}{n}-}\right]-\left(\tilde{E}\left[\left(U_n\left(\frac{i+1}{n}\right),\varphi\right)|\mathcal{F}_{\frac{i+1}{n}-}\right]\right)^2 = \frac{1}{n}\left(U_n\left(\frac{i+1}{n}-\right),\nu_n^i\varphi^2\right). \tag{9}$$

In between two branches the particles move according to the prescribed law; hence for t in the interval $[\frac{i}{n},\frac{i+1}{n})$,

$$(U_n(t),\varphi)=\left(U_n\left(\frac{i}{n}\right),\varphi\right)+\int_{\frac{i}{n}}^t(U_n(s),A\varphi)ds+S_n^{\varphi,i}(t), \tag{10}$$

where $\{(S_n^{\varphi,i}(t),\mathcal{F}_t),\, t\in[\frac{i}{n},\frac{i+1}{n}]\}$ is a square integrable martingale (we use Remark 3.2) with the quadratic variation

$$\langle S_n^{\varphi,i}\rangle(t)=\frac{1}{n}\int_{\frac{i}{n}}^t(U_n(s),A\varphi^2-2\varphi A\varphi)ds. \tag{11}$$

In order to compute the evolution equation of $(U_n(t),\varphi)$, we need to add all the parts coming from the particles' motion and all the parts coming from the particles' branching, which gives

$$\begin{aligned}(U_n(t),\varphi)&=(U_n(0),\varphi)+\int_0^t(U_n(s),A\varphi)ds+S_n^{\varphi}(t)\\&\quad+\sum_{i=1}^{[nt]}\left(\left(U_n\left(\frac{i}{n}\right),\varphi\right)-\left(U_n\left(\frac{i}{n}-\right),\varphi\right)\right),\end{aligned}$$

where $\{(S_n^{\varphi}(t),\,\mathcal{F}_t),\, t\in[0,1]\}$ is the square integrable martingale

$$S_n^{\varphi}(t)\triangleq S_n^{\varphi,[nt]}(t)+\sum_{i=0}^{[nt]-1}S_n^{\varphi,i}\left(\frac{i+1}{n}\right)$$

with the quadratic variation presented in (6). We then split the term coming from the branching into a martingale part and a bounded variation part and obtain

$$\begin{aligned}(U_n(t),\varphi)&=(U_n(0),\varphi)+\int_0^t(U_n(s),A\varphi)ds+S_n^{\varphi}(t)+M_n^{\varphi}([nt])\\&\quad+\sum_{i=1}^{[nt]}\left(\bar{E}\left[\left(U_n\left(\frac{i}{n}\right),\varphi\right)|\mathcal{F}_{\frac{i}{n}-}\right]-\left(U_n\left(\frac{i}{n}-\right),\varphi\right)\right),\end{aligned} \tag{12}$$

and $\{(M_n^{\varphi}(l),\,\mathcal{F}_{\frac{l+1}{n}-}),\, l=0,1,\ldots,n\}$ is the square integrable martingale

$$M_n^{\varphi}(0)\triangleq 0,$$

$$M_n^\varphi(l) \triangleq \sum_{i=1}^{l}\left(\left(U_n\left(\frac{i}{n}\right),\varphi\right)-\tilde{E}\left[\left(U_n\left(\frac{i}{n}\right),\varphi\right)|\mathcal{F}_{\frac{i}{n}-}\right]\right)$$

with quadratic variation

$$\langle M_n^\varphi\rangle(l)=\sum_{i=1}^{l}\tilde{E}\left[\left(\left(U_n\left(\frac{i}{n}\right),\varphi\right)-\tilde{E}\left[\left(U_n\left(\frac{i}{n}\right),\varphi\right)|\mathcal{F}_{\frac{i}{n}-}\right]\right)^2|\mathcal{F}_{\frac{i}{n}-}\right]. \tag{13}$$

The remark now follows easily from (8), (9), (12), and (13). □

We introduce the notation

$$\lambda_n^i(V,r)\triangleq \exp\left(\int_{\frac{i}{n}}^{r} h^*(V(t))dY_t-\frac{1}{2}\int_{\frac{i}{n}}^{r} h^*h(V(t))dt\right),\qquad r\in\left[\frac{i}{n},\frac{i+1}{n}\right],$$

where V is the trajectory of a particle alive in the interval $[\frac{i}{n},\frac{i+1}{n}]$. Of course, $\mu_n^i(V)=\lambda_n^i(V,\frac{i+1}{n})$. Applying Ito's rule and exploiting the fact that $\langle Y^i,Y^j\rangle(t)=t\delta_{ij}$, we get from (5) that

$$(U_n(t),\varphi)=(U_n(0),\varphi)+\int_0^t(U_n(s),A\varphi)ds+S_n^\varphi(t)+M_n^\varphi([nt])$$

$$+\frac{1}{n}\int_0^{\frac{[nt]}{n}}\sum_{j=1}^{m_n(\frac{[sn]}{n})}\varphi\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\lambda_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s. \tag{14}$$

4. The tightness of the sequence. We will prove first that the sequence is tight in $D_{M_F'(E)}[0,1]$. For this, it is sufficient to prove (cf. [29]) that the processes $\{(U_n(s),\varphi_i),\ s\in[0,1]\}$ form a tight sequence, where $\{\varphi_i\}_{i\geq 0}$ is defined as follows: φ_0 is the constant function 1 and $\{\varphi_i;\ \ i>0\}$ is a dense set in $C_0(E)$ (we will take them to be in $\mathcal{D}(A_0)$). In order to prove that $\{(U_n(s),\varphi_i),\quad s\in[0,1]\}$ is a tight sequence for every $i\geq 0$, we use the following theorem (cf. [1]).

THEOREM 4.1 (see Aldous [1]). *Let $\{a_n\}$ be a sequence of real valued processes with cádlág paths such that*

(i) *$\{a_n(t)\}$ is tight on the line for each $t\in[0,1]$;*

(ii) *for any arbitrary sequence of stopping times $\{\tau_n\}_{n\geq 0}$ (with respect to the natural filtration of $\{a_n\}$) and any sequence $\{\delta_n\}_{n\geq 0}$ of positive real numbers with $\lim_{n\to\infty}\delta_n=0$, we have*

$$\lim_{n\to\infty}a_n(\tau_n+\delta_n)-a_n(\tau_n)=0\quad in\quad probability.$$

Then $\{a_n\}$ is tight.

Conditions (i) and (ii) follow from Propositions 4.1 and 4.2.

PROPOSITION 4.2. *For every $t\in[0,1]$, we have*

$$\lim_{k\to\infty}\sup_{n\geq 0}\tilde{P}\left(\sup_{0\leq s\leq t}(U_n(s),1)>k\right)=0. \tag{15}$$

Proof. Since

$$\tilde{P}\left(\sup_{0\le s\le t}(U_n(s),1)>k\right)\le\frac{\tilde{E}[(\sup_{0\le s\le t}(U_n(s),1))^2]}{k^2}, \tag{16}$$

it is enough to show that $\sup_{n\ge 0}\bar{E}[(\sup_{0\le s\le t}(U_n(s),1))^2]$ is finite. Let us denote by $\psi_n(t)\triangleq\tilde{E}[(\sup_{0\le s\le t}(U_n(s),1))^2]$. From (14) and the Cauchy–Schwarz inequality, we obtain

$$\psi_n(t)\le 3(U_n(0),1)^2+3\tilde{E}\left[\left(\sup_{0\le i\le[nt]}|M_n^1(i)|\right)^2\right]$$

$$+\frac{3}{n^2}\tilde{E}\left[\left(\sup_{0\le p\le\frac{[tn]}{n}}\left|\int_0^p\sum_{j=1}^{m_n(\frac{[sn]}{n})}\lambda_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s\right|\right)^2\right]. \tag{17}$$

The idea is to find an upper bound for each of the three terms from the right-hand side of inequality (17) of the form $\alpha+\beta\int_0^t\psi_n(s)ds$ and then use Gronwall's inequality.

<u>The first term</u>.

Since we start with n particles, we have

$$(U_n(0),1)^2=1,\qquad\forall n\ge 1. \tag{18}$$

<u>The second term</u>.

Doob's maximal inequality, the fact that $\nu_n^i\le\frac{1}{4}$, and Proposition 3.1 give us the following upper bound:

$$\tilde{E}\left[\left(\sup_{0\le i\le[nt]}|M_n^1(i)|\right)^2\right]\le 4\tilde{E}[(M_n^1([nt]))^2]=4\tilde{E}[\langle M_n^1\rangle([nt])]$$

$$=\frac{4}{n}\sum_{i=1}^{[nt]}\tilde{E}\left[\left(U\left(\frac{i}{n}-\right),\nu_n^i\right)\right]\le\frac{1}{n}\sum_{i=1}^{[nt]}\tilde{E}\left[\left(U\left(\frac{i}{n}-\right),1\right)\right]\le\frac{1}{n}\sum_{i=1}^{[nt]}\frac{\tilde{E}[m_n(t)]}{n}\le 1. \tag{19}$$

<u>The third term</u>.

Using a standard technique (based on Ito's rule, Burkholder's inequality, and Gronwall's inequality), we find

$$\tilde{E}\left[\left(\sum_{j=1}^{m_n(\frac{[sn]}{n})}\lambda_n^{[sn]}(V_n^j,s)\right)^2\right]\le e^{2\|h\|^2}\tilde{E}\left[\left(m_n\left(\frac{[sn]}{n}\right)\right)^2\right]. \tag{20}$$

Then we obtain the following upper bound on the third term of (17), using (20) and Burkholder's inequality:

$$e^{2\|h\|^2}K_2\|h\|^2\int_0^t\psi_n(s)ds, \tag{21}$$

where K_2 is a constant independent of n.

From (17), (18), (19), and (21) we obtain

$$\psi_n(t) \le 6 + 3e^{2\|h\|^2} K_2 \|h\|^2 \int_0^t \psi_n(s) ds.$$

Using once again the Gronwall inequality we find that $\psi_n(t) \le c(t)$, where

$$c(t) \stackrel{\text{def}}{=} 6e^{\frac{1}{2}e^{2\|h\|^2} K_2 \|h\|^2 t},\ t \in [0,1]$$

and also $\sup_{n\ge 0} \bar{E}[(\sup_{0\le s\le t}(U_n(s),1))^2] \le c(t)$. ∎

PROPOSITION 4.3. *For any arbitrary sequence of stopping times $\{\tau_n\}_{n\ge 0}$, any real positive sequence $\{\delta_n\}_{n\ge 0}$ with $\lim_{n\to\infty}\delta_n = 0$ and $\varphi \in \mathcal{D}(A_0)$, we have*

$$\lim_{n\to\infty} \tilde{E}[|(U_n(\tau_n+\delta_n),\varphi) - (U_n(\tau_n),\varphi)|^2] = 0. \tag{22}$$

Proof. Using (14) and the Cauchy–Schwarz inequality we get

$$\begin{aligned}
&\tilde{E}[|(U_n(\tau_n+\delta_n),\varphi) - (U_n(\tau_n),\varphi)|^2] \\
&\qquad \le 4\tilde{E}\left[\left(\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi)ds\right)^2\right] \\
&\qquad\quad + 4\tilde{E}[(S_n^\varphi(\tau_n+\delta_n) - S_n^\varphi(\tau_n))^2] \\
&\qquad\quad + 4\tilde{E}[(M_n^\varphi([n(\tau_n+\delta_n)]) - M_n^\varphi([n\tau_n]))^2] \\
&\qquad\quad + 4\tilde{E}\left[\left(\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}} \sum_{j=1}^{m_n(\frac{[sn]}{n})} \varphi\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\right.\right. \\
&\qquad\quad \left.\left. \times\, \lambda_n^{[sn]}(V_n^j,s) h^*(V_n^j(s)) dY_s\right)^2\right].
\end{aligned} \tag{23}$$

We have, consecutively,

$$\begin{aligned}
\tilde{E}\left[\left(\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi)ds\right)^2\right] &\le \delta_n \tilde{E}\left[\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s),A\varphi)^2 ds\right] \\
&\le \delta_n^2 \|A\varphi\|^2 c(1),
\end{aligned} \tag{24}$$

$$\begin{aligned}
\tilde{E}[(S_n^\varphi(\tau_n+\delta_n) - S_n^\varphi(\tau_n))^2] &= \tilde{E}[\langle S_n^\varphi\rangle(\tau_n+\delta_n) - \langle S_n^\varphi\rangle(\tau_n)] \\
&\le \frac{1}{n}\tilde{E}\left[\int_{\tau_n}^{\tau_n+\delta_n}(U_n(s), A\varphi^2 - \varphi A\varphi)ds\right] \\
&\le \frac{(c(1)+1)(\|A\varphi^2\| + 2\|\varphi\|\|A\varphi\|)}{2n}\delta_n,
\end{aligned} \tag{25}$$

$$\begin{aligned}
\tilde{E}[(M_n^\varphi([n(\tau_n+\delta_n)]) - M_n^\varphi([n\tau_n]))^2] &= \tilde{E}[\langle M_n^\varphi\rangle([n(\tau_n+\delta_n)]) - \langle M_n^\varphi\rangle([n\tau_n])] \\
&= \frac{1}{n}\tilde{E}\left[\sum_{[n\tau_n]}^{[n(\tau_n+\delta_n)]}\left(U_n\left(\frac{i+1}{n}-\right), \nu_n^{i+1}\varphi^2\right)\right] \\
&\le \frac{1}{8}(c(1)+1)\|\varphi\|^2\left(\delta_n + \frac{1}{n}\right),
\end{aligned} \tag{26}$$

$$\tilde{E}\left[\left(\frac{1}{n^2}\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}}\sum_{j=1}^{m_n(\frac{[sn]}{n})}\varphi\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\lambda_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s\right)^2\right]$$
$$=\tilde{E}\left[\int_{\frac{[n\tau_n]}{n}}^{\frac{[n(\tau_n+\delta_n]}{n}}\frac{1}{n^2}\left(\sum_{j=1}^{m_n(\frac{[sn]}{n})}\varphi\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\lambda_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))\right)^2 ds\right]$$
$$\le \|\varphi\|^2\|h\|^2 c(1)e^{2\|h\|^2}\delta_n, \tag{27}$$

where c is the function defined in the previous proposition. The inequalities (24), (25), (26), (27) imply that all the terms from the right-hand side of (**??**) tend to 0 as n goes to ∞; hence $\tilde{E}[|(U_n(\tau_n+\delta_n),\varphi)-(U_n(\tau_n),\varphi)|^2]$ tends to 0 as well. ∎

We show now that the sequence "stays mostly" within compact sets and hence it is tight also in $D_{M_F(E)}[0,1]$, assuming the following property is satisfied.

Property K. There exists a sequence of compact sets $K_k \in E$ and a sequence of functions $\{\varphi_k\}_{k>0} \subset \mathcal{D}(A_0)$, $\varphi_k : E \to [0,1]$ such that $\lim_{k\to\infty}\varphi_k = 1$ π_0-a.s., $\lim_{k\to\infty}\|A\varphi_k\| = 0$, and $\varphi_k\big|_{CK_{k+1}} = 0$ ($CA \overset{\triangle}{=} E - A$).

THEOREM 4.4. *Assuming that property* **K** *is satisfied, the sequence is tight over the space* $D_{M_F(E)}[0,1]$.

Proof. Since we already know that the sequence is tight over $D_{M_F'(E)}[0,1]$, we only need to prove that there exists a sequence of compact sets $K_k \in E$ such that, for every $\varepsilon > 0$,

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{P}\left(\sup_{t\in[0,1]}(U_n(t), I_{CK_k}) > \varepsilon\right) = 0 \tag{28}$$

(we denote by I_A the characteristic function of the set $A \in E$). We show that (28) is valid for the sequence of compact sets for which property **K** is valid. Using Chebyshev's inequality, it is enough to prove that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,1]}(U_n(t), I_{CK_k})^2\right] = 0, \tag{29}$$

which, in turn, is implied by

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,1]}(U_n(t), 1-\varphi_k)^2\right] = 0. \tag{30}$$

Using (14) with $\varphi := 1-\varphi_k$, we get that

$$(U_n(t), 1-\varphi_k) \tag{31}$$
$$=(U_n(0),1-\varphi_k) + \int_0^t (U_n(s), A(1-\varphi_k))ds + S_n^{1-\varphi_k}(t) + M_n^{1-\varphi_k}([nt])$$
$$+\frac{1}{n}\int_0^{\frac{[nt]}{n}}\sum_{j=1}^{m_n(\frac{[sn]}{n})}(1-\varphi_k)\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\lambda_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s.$$

By integrating (31) and using the fact that $A1 = 0$, we obtain

$$\tilde{E}[(U_n(t), 1-\varphi_k)] = (U_n(0), 1-\varphi_k) + \tilde{E}\left[\int_0^t (U_n(s), A(1-\varphi_k))ds\right]$$
$$\leq (U_n(0), 1-\varphi_k) + t\|A\varphi_k\|.$$

Hence

$$\lim_{k\to\infty}\lim_{n\to\infty} \tilde{E}[(U_n(t), 1-\varphi_k)] = \lim_{k\to\infty}((\pi_0, 1-\varphi_k) + t\|A\varphi_k\|) = 0. \tag{32}$$

As before, we have that

$$\lim_{k\to\infty}\limsup_{n\to\infty}(U_n(0), 1-\varphi_k)^2 = \lim_{k\to\infty}(\pi_0, 1-\varphi_k)^2 = 0, \tag{33}$$

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}\left(\int_0^t (U_n(s), A(1-\varphi_k))ds\right)^2\right] = \lim_{k\to\infty}\|A\varphi_k\|^2 c(1) = 0. \tag{34}$$

Using Burkholder–Davis–Gundy inequality,

$$\begin{aligned}\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}(S_n^{1-\varphi_k}(t))^2\right] &\leq \lim_{k\to\infty}\limsup_{n\to\infty}\frac{K_2}{n}\left(\tilde{E}\left[\int_0^T (U_n(r), A\varphi_k^2)\right]\right.\\ &\qquad\left.-\tilde{E}\left[\int_0^T (U_n(r), 2\varphi_k A\varphi_k)dr\right]\right)\\ &= 0\end{aligned} \tag{35}$$

and also

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}|M_n^{1-\varphi_k}(t)|\right] \leq \lim_{k\to\infty}\limsup_{n\to\infty}\frac{K_2}{4}\tilde{E}\left[\int_0^T (U_n(r), (1-\varphi_k)^2)dr\right].$$

Since $\varphi_k^2 \leq \varphi_k$, we have, using Fatou's lemma,

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}\left(M_n^{1-\varphi_k}(t)\right)^2\right] \leq \lim_{k\to\infty}\frac{K_2}{4}\int_0^T \limsup_{n\to\infty}\tilde{E}[(U_n(r), 1-\varphi_k)]dr. \tag{36}$$

From (32) and (36) we obtain that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}(M_n^{\varphi_k}(t))^2\right] = 0. \tag{37}$$

In order to control the last term in (31) we make use of the following result whose proof we omit here (it is similar to the proof of Proposition 4.5 from [10])

$$\begin{aligned}\lim_{n\to\infty}\tilde{E}\Bigg[\sup_{t\in[0,1]}\Bigg(&\frac{1}{n}\int_0^{\frac{[nt]}{n}}\sum_{j=1}^{m_n(\frac{[sn]}{n})}(1-\varphi_k)\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\\ &\times\lambda_n^{[sn]}(V_n^j, s)h^*(V_n^j(s))dY_s - \int_0^t (U_n(s), h^*(1-\varphi_k))dY_s\Bigg)^2\Bigg] = 0.\end{aligned} \tag{38}$$

Using (38) we obtain that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}\left(\frac{1}{n}\int_0^{\frac{[nt]}{n}}\sum_{j=1}^{N_n(\frac{[sn]}{n})}(1-\varphi_k)\left(V_n^j\left(\frac{[sn]+1}{n}\right)\right)\right.\right.$$
$$\left.\left.\times\lambda_n^{[sn]}(V_n^j,s)h^*(V_n^j(s))dY_s\right)^2\right]$$
$$=\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}\left[\sup_{t\in[0,T]}\left(\int_0^t(U_n(r),h^*(1-\varphi_k))dY_r\right)^2\right]$$
$$\le k\left\|h\right\|^2\lim_{k\to\infty}\limsup_{n\to\infty}^{2}\int_0^t\tilde{E}\left[(U_n(r),1-\varphi_k)^2\right]dr. \tag{39}$$

Combining (31), (33), (34), (35), (37), and (39) and using again Fatou's lemma, we obtain that there exists a constant K such that

$$\iota(T)\le K\int_0^T\iota(s)ds,$$

where $\iota(T)\triangleq\lim_{k\to\infty}\limsup_{n\to\infty}\tilde{E}[\sup_{t\in[0,T]}(U_n(t),1-\varphi_k)^2]$, which implies our claim using, once again, Gronwall's inequality. □

5. Convergence of the sequence U_n. In the following we will denote by $h(r)$ the function $x\to h(r,x)$. We consider first the following backward stochastic partial differential equation (SPDE):

$$d\psi_s=-A\psi_sds-h^*(s)\psi_s\bar{d}Y_s,\quad s\le t, \tag{40}$$
$$\psi_t=\varphi$$

which, written in the integral form, gives us

$$\psi_r=\psi_s-\int_s^rA\psi_pdp-\int_s^rh^*(p)\psi_p\bar{d}Y_p\ r,s\in[0,t]. \tag{41}$$

In (41), we took $\int_s^r h^*(p)\psi_p\bar{d}Y_p$ to be a backward Itô integral. Written in Stratonovitch form, (41) becomes

$$\psi_r=\psi_s-\int_s^pA\psi_pdp-\int_s^ph^*(p)\psi_p\circ dY_p+\frac{1}{2}\int_s^ph^*h(p)\psi_pdp. \tag{42}$$

We will assume the following condition.

Condition U. For all $t\in[0,1]$, there exists a countable set $\mathcal{M}=\{\varphi_k,k\ge1\}$, uniformly dense in $C_0(E)$, such that for all $\varphi\in\mathcal{M}\cup\{1\}$ the SPDE (40) has a solution $\psi_s\in\mathcal{D}(A)$ which satisfies

$$B\stackrel{\text{def}}{=}\tilde{E}\left[\sup_{s\in[0,t]}\|\psi_s\|^2\right]<\infty \tag{43}$$

and

$$C\stackrel{\text{def}}{=}\tilde{E}\left[\sup_{s\in[0,t]}\|A\psi_s^2-2\psi_sA\psi_s\|\right]<\infty. \tag{44}$$

See [9] for the necessary conditions on A and h, under which $\mathbf{U}$ holds.

In this section we prove the following results.

PROPOSITION 5.1. *Under the assumption* $\mathbf{U}$*, for all* $\varphi \in \mathcal{M} \cup \{1\}$ *and* $t \in [0,1]$ *we have*

$$\lim_{k\to\infty} \tilde{E}[((U_n(t),\varphi) - p_t(\varphi))^2] = 0. \tag{45}$$

Let d be the following metric on $M_F(E)$:

$$d(\mu,\nu) = \sum_{k=0}^{\infty} \frac{|(\mu,\varphi_k) - (\nu,\varphi_k)|}{2^k \|\varphi_k\|},$$

where $\varphi_0 = 1$. Then Proposition 5.1 will imply the following theorem.

THEOREM 5.2. *Under the assumption* $\mathbf{U}$*, for all* $t \in [0,1]$ *we have that*

$$\lim_{k\to\infty} \tilde{E}[d(U_n(t),p_t)] = 0.$$

Since the metric d generates the topology on $M_F(E)$, Theorem 5.2 gives us the following obvious corollary.

COROLLARY 5.3. *The sequence* $U_n(t)$ *is convergent in measure to* p_t *in* $M_F(E)$*, i.e., to the law of* $X(t)$ *given the observation* σ*-field* $\mathcal{Y}_t$*.*

It is this result that justifies the use of U_n in order to numerically approximate p. Theorem 5.2 (together with the tightness of the laws of the processes U_n) has another corollary of theoretical importance.

COROLLARY 5.4. *Under the assumptions* $\mathbf{K}$ *and* $\mathbf{U}$*, the sequence* U_n *converges in distribution to the measure valued process that represents the unnormalized conditional law of* X *given the observation* Y *in* $D_{M_F(E)}[0,1]$*.*

REMARK 5.5. *In Corollary* 5.4*, we look at* U_n *and* p *as having values in the space of right continuous with limits to the left,* $M_F(E)$*-valued paths. The previous section proves that the sequence is tight over the space* $D_{M_F(E)}[0,1]$*. This together with Theorem* 5.2 *ensures that* U_n *converges in distribution to* p *in* $D_{M_F(E)}[0,1]$*.*

We need to prove first the following lemma (the notations are those from section 3).

LEMMA 5.6. *For all* $i = 0,\ldots,n-1$*, we have*

$$\tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} \nu_n^i(V_n^j)\right] \le \frac{M(h)}{\sqrt{n}},$$

where $M(h)$ *is a constant depending only on* h*.*

Proof. We remark that, if we have an integer random variable with mean μ and minimal variance ν, then $\nu \le |\mu - 1|$. Hence

$$\begin{aligned}\tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} \nu_n^i(V_n^j)\right] &\le \tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} |\mu_n^i(V_n^j) - 1|\right] \\ &\le \tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})} E\left[|\mu_n^i(V_n^j) - 1| \,\Big|\, \mathcal{F}_{\frac{i}{n}}\right]\right]\end{aligned}$$

$$\leq \tilde{E}\left[\frac{1}{n}\sum_{j=1}^{m_n(\frac{i}{n})}\sqrt{E\left[\left(\left(\int_{\frac{i}{n}}^{\frac{i+1}{n}}\lambda_n^i(V_N^j,s)h^*(V_N^j(s))dY_s\right)^2\middle|\mathcal{F}_{\frac{i}{n}}\right]\right]}\right]$$
$$\leq \frac{M(h)}{\sqrt{n}}. \qquad \square$$

Now let $\{V_r, \mathcal{F}_r, r\in[s,t]\}$ be a process solution of the martingale problem associated with the infinitesimal generator A independent of $\mathcal{Y}_s^t \triangleq \sigma(Y_t - Y_r, r\in[s,t])$. From (42) we obtain that

$$\psi_r(V_r) = \psi_s(V_s) - \int_s^r h^*(p,V_p)\psi_p(V_p)\circ dY_p + \frac{1}{2}\int_s^r h^*h(p,V_p)\psi_p(V_p)dp + M_r^\psi,$$

where $\{M_r^\psi, \mathcal{F}_r\vee\mathcal{Y}, r\in[s,t]\}$ is a square integrable martingale (due to (43)) with quadratic variation

$$\langle M_r^\psi\rangle = \int_s^r A\psi_p^2(V_p) - 2\psi_p(V_p)A\psi_p(V_p)dp.$$

Let

$$\xi_r = \exp\left(\int_s^r h^*(p,V_p)\circ dY_p - \frac{1}{2}\int_s^r h^*h(p,V_p)dp\right),$$

then

$$\xi_r = 1 + \int_s^r \xi_p h^*(p,V_p)\circ dY_p - \frac{1}{2}\int_s^r \xi_p h^*h(p,V_p)dp$$

and thus

$$\psi_r(V_r)\xi_r = \psi_s(V_s) + \overline{M}_r^\psi, \tag{46}$$

where $\{\overline{M}_r^\psi, \mathcal{F}_r\vee\mathcal{Y}, r\in[s,t]\}$ is a square integrable martingale with quadratic variation

$$\langle\overline{M}_r^\psi\rangle = \int_s^r \xi_p^2(A\psi_p^2(V_p) - 2\psi_p(V_p)A\psi_p(V_p))dp.$$

Hence

$$\begin{aligned}\widetilde{E}[(\psi_r(V_r)\xi_r - \psi_s(V_s))^2] &= \int_s^r \widetilde{E}[\xi_p^2(A\psi_p^2(V_p) - 2\psi_p(V_p)A\psi_p(V_p))]dp\\ &\leq \int_s^r \widetilde{E}[\xi_p^2\|A\psi_p^2 - 2\psi_pA\psi_p\|]dp\\ &\leq (r-s)Ce^{\|h\|^2}. \end{aligned}\tag{47}$$

The last inequality is true since ξ_p and $A\psi_p^2 - 2\psi_pA\psi_p$ are independent and $\widetilde{E}[\xi_p^2]\leq e^{\|h\|^2}$. Armed now with the inequality (47) we can prove Proposition 5.1.

Proof of Proposition 5.1. Since $\overline{M}_r^\psi$ is a martingale with respect to the filtration $\mathcal{F}_r\vee\mathcal{Y}$, from (46), we get that, for $\varphi\in\mathcal{M}\cup\{1\}$,

$$p_t(\varphi) = \widetilde{E}[\varphi(X_t)\xi_t\,|\mathcal{Y}] = \widetilde{E}[\psi_t(X_t)\xi_t\,|\mathcal{Y}] = \widetilde{E}[\psi_0(X_0)\,|\mathcal{Y}] = (\pi_0,\psi_0). \tag{48}$$

In (48), we used the fact that $\psi_t = \varphi$. Also, since $U_n(0)$ converges weakly to π_0 and $\psi_0 \in \mathcal{D}(A) \subset C_b(E)$, $\tilde{P}$-a.s., we have that

$$\lim_{n\to\infty} (U_n(0), \psi_0) = (\pi_0, \psi_0),\ \widetilde{P} - \text{a.s.} \tag{49}$$

and, by using (43) and the Dominated Convergence theorem, we obtain

$$\lim_{n\to\infty} \widetilde{E}[((U_n(0), \psi_0) - (\pi_0, \psi_0))^2] = 0. \tag{50}$$

Hence in order to prove the proposition, it is enough to show that

$$\lim_{n\to\infty} \widetilde{E}[((U_n(t), \psi_t) - (U_n(0), \psi_0))^2] = 0. \tag{51}$$

We have the following identity:

(52)

$$\begin{aligned}(U_n(t), \psi_t) - (U_n(0),\psi_0) &= (U_n(t), \psi_t) - \left(U_n\left(\frac{[nt]}{n}\right), \psi_{\frac{[nt]}{n}}\right)\\
&+\sum_{i=1}^{\frac{[nt]}{n}} \left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right) - \widetilde{E}\left[\left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right)\Big|\mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right]\\
&+\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right)\Big|\mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right] - \left(U_n\left(\frac{i-1}{n}\right), \psi_{\frac{i-1}{n}}\right).\end{aligned}$$

We show that all the terms from the right-hand side of (52) converge to 0 in $L^2(\Omega)$. For the first term we have the following:

$$\begin{aligned}(U_n(t), \psi_t) - \left(U_n\left(\frac{[nt]}{n}\right), \psi_{\frac{[nt]}{n}}\right) &= \frac{1}{n}\sum_{j=1}^{m_n(t)} \varphi(V_n^j(t))(1-\lambda_n^{[nt]}(V_n^j, t))\\
&+\frac{1}{n}\sum_{j=1}^{m_n(t)} \psi_t(V_n^j(t))\lambda_n^{[nt]}(V_n^j, t) - \psi_{\frac{[nt]}{n}}\left(V_n^j\left(\frac{[nt]}{n}\right)\right).\end{aligned}$$

Now using (47) and the fact the φ is bounded, we get the following upper bound for the L_2 norm of the first term:

$$\frac{2c(1)e^{\|h\|^2}(2\|\varphi\|^2\|h\|^2 + C)}{n}. \tag{53}$$

For the second term we have the following identity:

$$\begin{aligned}&\widetilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}} \left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right) - \widetilde{E}\left[\left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right)\Big|\mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right]\right)^2\right] \qquad (54)\\
&= \widetilde{E}\left[\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\left(\left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right) - \widetilde{E}\left[\left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right)\Big|\mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right]\right)^2 \Big|\mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right]\right]\\
&= \frac{1}{n^2}\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\sum_{j=1}^{m_n(\frac{i-1}{n})} \psi_{\frac{i}{n}}\left(V_n^j\left(\frac{i}{n}\right)\right)^2 \nu_n^{i-1}(V_n^j)\right] \le \frac{BM(h)}{\sqrt{n}},\end{aligned}$$

where we used Lemma 5.6 to obtain the last inequality. Lastly, for the third term we have the identity

$$\widetilde{E}\left[\left(\sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\left(U_n\left(\frac{i}{n}\right), \psi_{\frac{i}{n}}\right) \Big| \mathcal{F}_{\frac{i}{n}-} \vee \mathcal{Y}\right] - \left(U_n\left(\frac{i-1}{n}\right), \psi_{\frac{i-1}{n}}\right)\right)^2\right]$$
$$= \frac{1}{n^2} \sum_{i=1}^{\frac{[nt]}{n}} \widetilde{E}\left[\sum_{j=1}^{m_n(\frac{i-1}{n})} \widetilde{E}\left[\left(\psi_{\frac{i}{n}}\left(V_n^j\left(\frac{i}{n}\right)\right) \xi_n^{i-1}\left(V_n^j, \frac{i}{n}\right)\right.\right.\right.$$
$$\left.\left.\left. - \psi_{\frac{i-1}{n}}\left(V_n^j\left(\frac{i-1}{n}\right)\right)\right)^2 \Big| \mathcal{F}_{\frac{i-1}{n}} \vee \mathcal{Y}\right]\right]$$

and one finds the following upper bound for the L_2 norm of the third term

$$\frac{Ce^{\|h\|^2}}{n}. \tag{55}$$

From (52), (53), (54), and (55) we get the required limit (51) and with it the proof of the proposition.

Proof of Theorem 5.2. We have, for arbitrary $m > 0$, the following upper bound on $d(U_n, p_t)$:

$$d(U_n(t), p_t) = \sum_{k=0}^{\infty} \frac{|(U_n(t), \varphi_k) - p_t(\varphi_k)|}{2^k \|\varphi_k\|}$$
$$\leq \sum_{k=0}^{m} \frac{|(U_n(t), \varphi_k) - p_t(\varphi_k)|}{2^k \|\varphi_k\|} + \sum_{k=m}^{\infty} \frac{|(U_n(t), \varphi_k)| + |p_t(\varphi_k)|}{2^k \|\varphi_k\|}$$
$$\leq \sum_{k=0}^{m} \frac{|(U_n(t), \varphi_k) - p_t(\varphi_k)|}{2^k \|\varphi_k\|} + \frac{(U_n(t), 1) + p_t(1)}{2^{m-1}}. \tag{56}$$

From Proposition 5.1 we have that $\lim_{k\to\infty} \tilde{E}[|(U_n(t), \varphi_k) - p_t(\varphi_k)|] = 0$ for all $k \geq 0$; hence (56) implies that $\limsup_{n\to\infty} \tilde{E}[d(U_n(t), p_t)] \leq \frac{1}{2^{m-2}}$ (we know that $E[(U_n(t), 1)] = E[p_t(1)] = 1$). Since $m > 0$ was arbitrary we have our claim.

Proof of Corollary 5.3. From Theorem 5.2, we have that $\lim_{n\to\infty} d(U_n(t), p_t) = 0$ in (probability) measure and since the metric d is giving the topology on $M_F(E)$, $U_n(t)$ converges in measure to p_t in $M_F(E)$.

Proof of Corollary 5.4. Since for any fixed t, $U_n(t)$ converges in measure to $p(t)$, for all finite sequences $0 \leq t_1 \leq t_2 \leq \cdots \leq t_k \leq 1$ $(U_n(t_1), U_n(t_2), \ldots, U_n(t_k))$ converges in measure to $(p_{t_1}, p_{t_2}, \ldots, p_{t_k})$ and hence also in distribution, i.e., the finite-dimensional distributions of U_n converge to the finite-dimensional distributions of p. This, together with the tightness of the sequence, implies that the processes U_n are convergent in distribution to p.

6. Implementation. We have produced computer programs to simulate the branching particle system defined at the beginning of section 3. The method was implemented for the case where the d_1-dimensional signal process, $X(t)$, is given by the stochastic differential equation (SDE)

$$dX_t = f(X)dt + g(X)dW_1(t). \tag{57}$$

The generator A is then

$$A = \sum_{i=1}^{d_1} f_i \frac{\partial}{\partial x_1} + \frac{1}{2} \sum_{i,j=1}^{d_1} \sigma_{ij} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j},$$

where σ_{ij} is entry (i,j) of the matrix $\sigma = gg^T$. Let the d_2-dimensional observation process, $Y(t)$, be given by

$$dY_t = h(X_t)dt + dW_2(t),$$

where $W_1(t)$ and $W_2(t)$ are independent (multidimensional) standard Brownian motions.

Following section 3, we take the final time as $T = 1$ and divide that time into n generations. At time $t = 0$, we generate n particles randomly with distribution $\pi(0)$. Then each generation is divided into two stages. In the first stage, from time $\frac{i}{n}$ to time $\frac{i+1}{n}$, each particle moves along a trajectory, determined by numerical solution of (57), using independent simulations of a Brownian path for each particle. In the second stage, at time $t = \frac{i+1}{n}$, each particle is replaced by a number of offspring, with the mean number of offspring being determined by the trajectory of the particle during stage one, as given by equation (4). For a general final time T, we take the number of generations to be nT (presuming that T is chosen so that nT is an integer).

Various interesting questions about the numerical implementation are already apparent from this description. In stage one, how should the trajectories of the particles be simulated? The usual way to solve an SDE such as (57) numerically is to choose a discretization scheme and to divide the interval over which the SDE is to be solved into a number of time steps (see, e.g., [27], [19]). Which discretization scheme should be chosen and how should the time step compare with the size of the generation? What errors can be tolerated in the calculation of the trajectories? In stage two, the only loss of accuracy (at least, if we presume that we have a perfect source of uniform random numbers) is in the calculation of the mean number of offspring for each particle. How should the integrals in (4) be calculated? The error here has two sources: not only have we replaced the real trajectory, $V(t)$, with an approximate trajectory, $\tilde{V}(t)$, but we have to approximate the integrals by finite sums.

We consider the question of approximating the trajectories of the particles first. We decided to use a discretization scheme giving the best possible order of accuracy, given the SDE to be solved, from among schemes that use only increments of the Brownian path and not path integrals. When the SDE (57) is of dimension one or satisfies the commutativity condition

$$\sum_{k=1}^{n} \left(\frac{\partial g_q^j}{\partial x^k} g_p^k - \frac{\partial g_p^j}{\partial x^k} g_q^k \right) = 0, \tag{58}$$

$\forall j = 1, \ldots, d_1$, $\forall p, q = 1, \ldots, r_1$, then the best order of accuracy is usually $O(h)$, where h is the chosen time step. This order can be obtained by using the Milshtein [26] or Heun schemes [8]. In the general case of an SDE in more than one dimension and not satisfying (58), the best obtainable accuracy is $O(\sqrt{h})$ (see [7]), which can be obtained using the Euler–Maruyama scheme. If we chose to generate not only increments of the Brownian path, but also stochastic path integrals, we could, at least in some cases, improve on the order of accuracy. This would allow us to reduce the number of time steps and hence the number of function evaluations required in each generation when wishing to obtain a fixed accuracy. However, for calculation of the integrals in (4), it may be preferable to evaluate the approximate trajectory at more points rather than fewer. After various trials, we decided that more than one time step is needed in each generation, but not more than about eight.

Now let us consider the question of how to approximate the integrals in (4), needed for determining the mean number of offspring for each particle at the end of each generation. The simplest method is clearly,

$$\sum_{j=0}^{m-1} h^*(V(t_j))\Delta Y(j) - \frac{1}{2}\sum_{j=0}^{m-1} h^*h(V(t_j))\Delta t, \tag{59}$$

where $\Delta t = \frac{1}{mn}$, $t_j = \frac{i}{n} + j\Delta t$, and $\Delta Y(j) = Y(t_{j+1}) - Y(t_j)$. Since the integrals involved are stochastic, use of the trapezoidal rule

$$\frac{1}{2}\sum_{j=0}^{m-1}\left[h^*(V(t_j)) + h^*(V(t_{j+1}))\right]\Delta Y(j) - \frac{1}{4}\sum_{j=0}^{m-1}\left[h^*h(V(t_j)) + h^*h(V(t_{j+1}))\right]\Delta t \tag{60}$$

does not give a higher-order approximation than (59), however it does improve the leading coefficient of the error significantly and we therefore decided to use it. Note that the sums in (60) converge to the same values as those in (59), with no correction term required, due to the independence of the processes V and Y.

7. Numerical examples.

7.1. Example 1. Our first example consists of a one-dimensional signal, $x(t)$, and a one-dimensional observation, $y(t)$, given by

$$dx_t = -\alpha x \; dt + \sigma dw_1(t),$$
$$dy_t = \arctan(x) dt + dw_2(t),$$

where $w_1(t)$ and $w_2(t)$ are independent one-dimensional standard Brownian motions. The parameter values used for the figures below are $\alpha = 1$ and $\sigma = 0.25$. The distribution of $x(0)$ was taken as normal with mean 1 and variance 0.25 and filtering was carried out from $t = 0$ until $t = 5$.

In Figure 1, we show the historical process for a simulation with 20 particles at the initial time. The past is shown only for particles alive at the final time. The signal is shown in Figure 2, along with the quartiles of the distribution of particles, for a simulation starting with 160 particles. In both pictures the simulation time has been divided into 160 generations and the time step used to calculate the trajectories of the particles is $h = 2^{-8}$. Figure 3 compares the expected mean of the signal calculated by numerical solution of the Zakai equation on the one hand and by the branching particle system on the other hand. The curve corresponding to the Zakai equation is lower at the final time than the other. The conditional densities of the signal at various times as calculated by solution of the Zakai equation are shown in Figure 4. The graphs progress with time from the right to left of the picture.

7.2. Example 2. In the second example we have a six-dimensional signal, representing the position and velocity of a tennis ball, and a four-dimensional observation, consisting of angles measured by observers at two different positions. We suppose that the tennis ball is thrown at time $t_0 = 0$ from an initial position and with an initial velocity, both drawn from normal distributions. The observers do not start

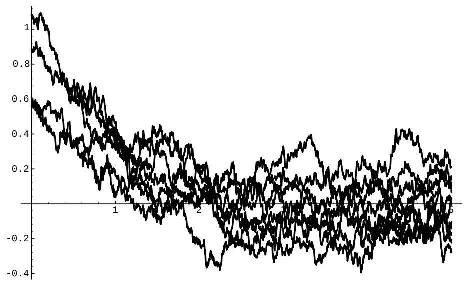

FIG. 1. *The historical process.*

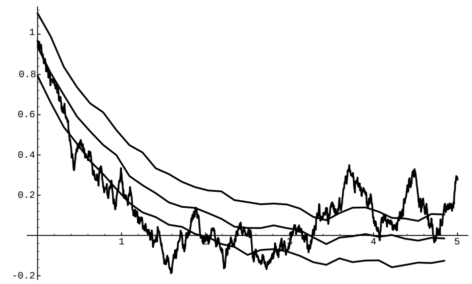

FIG. 2. *The signal and quartiles.*

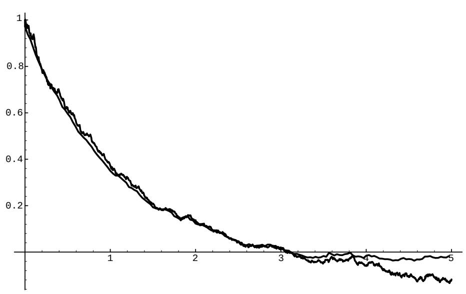

FIG. 3. *The conditional mean.*

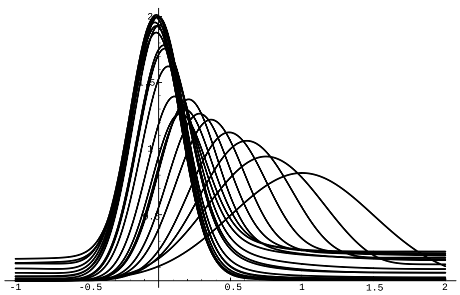

FIG. 4. *The conditional densities.*

measuring until time $t_1 > t_0$. The velocity of the ball is subject to white noise supposed to represent wind. The observations also include noise. The signal $(x(t), v(t))$, where $x, v \in \mathbb{R}^3$ are the position and velocity of the tennis ball, is therefore given by

$$dx(t) = v\ dt,$$
$$dv(t) = -Av\ dt + B\ dw(t),$$

where $w(t)$ is standard Brownian motion in $\mathbb{R}^3$. A and B are diagonal matrices with constant entries and $A(1,1) = A(2,2) = \epsilon$, $A(3,3) = g + \epsilon$, where g is the gravitational constant. The observation vector, $y(t) \in \mathbb{R}^4$ is defined by

$$dy_1(t) = \alpha_1 \left(\arctan \frac{x_3 - p_3}{x_1 - p_1} \right) dt + d\bar{w}_1(t),$$
$$dy_2(t) = \alpha_2 \left(\arctan \frac{x_3 - p_3}{x_2 - p_2} \right) dt + d\bar{w}_2(t),$$
$$dy_3(t) = \alpha_3 \left(\arctan \frac{x_3 - q_3}{x_1 - q_1} \right) dt + d\bar{w}_3(t),$$

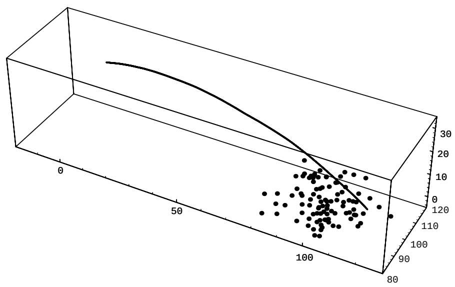

FIG. 5. *Time* 0.

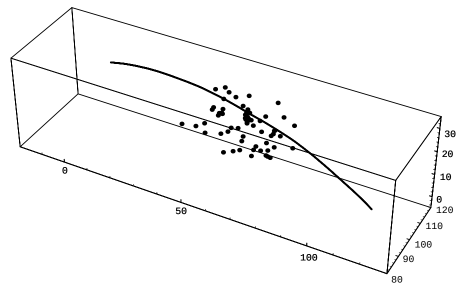

FIG. 6. *Time* 1.

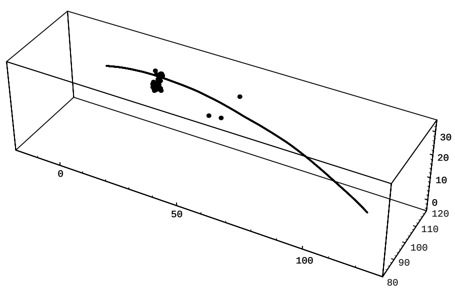

FIG. 7. *Time* 2.

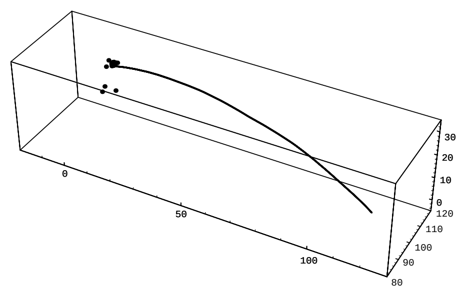

FIG. 8. *Time* 3.

$$dy_4(t) = \alpha_4 \left(\arctan \frac{x_3 - q_3}{x_2 - q_2} \right) dt + d\bar{w}_4(t),$$

where $p, q \in \mathbb{R}^3$ are the positions of the two observers and the four standard Brownian motions $\bar{w}_i$, $i = 1, \ldots, 4$ are all independent.

For the figures below, the parameter values chosen are $\epsilon = 0.01$, $B(i,i) = 3$, $\alpha_j = 1$, $i = 1, \ldots, 3$, $j = 1, \ldots, 4$. We started observations at $t_1 = 0.5$ and simulated until $T = 3$. The mean and standard deviation of the initial values are

	Mean	S.D.
x_1	100	10
x_2	100	10
x_3	0	0
v_1	-30	5
v_2	0	5
v_3	25	0

.

The number of generations per unit time and the time step used for simulating the trajectories of the particles are the same as in the first example. The pictures in Figures 4–8 show the cloud of particles at four different points in time. The signal has also been plotted in each picture.

Acknowledgments. The authors would like to sympathize with and thank the anonymous referee who read the paper very carefully and patiently (over more than one paper) protested at the point where the authors skated over the distinction between vague and weak convergence. We finally stopped looking for soft arguments and did the proof. Of course, the extra information simplified the whole paper.

The authors would also like to thank another referee for valuable references.

REFERENCES

[1] D. Aldous, *Stopping times and tightness*, Ann. Prob., 6, (1978), pp. 335–340.

[2] V. E. Beneš, *Exact finite-dimensional filters for certain diffusions with nonlinear drift*, Stochastics, 5 (1981), pp. 65–92.

[3] A. Bensoussan, *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, 1992.

[4] A. Bensoussan, R. Glowinski, and A. Rascanu, *Approximation of the Zakai equation by splitting up method*, SIAM J. Control Optim., 28 (1990), pp. 1420–1431.

[5] A. Bensoussan, R. Glowinski, and A. Rascanu, *Approximation of some stochastic differential equations by the splitting up method*, Appl. Math. Optim., 25 (1992), pp. 81–106.

[6] D. Brigo, B. Hanzon, and F. Le Gland, *A Differential Geometric Approach to Nonlinear Filtering: The Projection Filter*, preprint no. 914, IRISA, France.

[7] J. M. C. Clark and R. J. Cameron, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, in Stochastic Differential Systems, Lecture Notes in Control and Information Sciences 25, B. Grigelionis, ed., Springer-Verlag, Berlin, 1980, pp. 162–171.

[8] J. M. C. Clark, *The discretization of stochastic differential equations: A primer*, in Road Vehicle Systems and Related Mathematics, Proc. Second DMV-GAMM Workshop Turino 1987, Series Mathematical Methods in Technology, H. Neunzert, ed., Teubner, Stuttgart, and Kluwer, Academic Publishers, Amsterdam, 1988, pp. 163–179.

[9] D. Crisan, *Uniform Upper Bounds for the Solution of the Zakai Equation*, preprint.

[10] D. Crisan and T. Lyons, *Nonlinear filtering and measure valued processes*, Prob. Theory Related Fields, 109 (1997), pp. 217–244.

[11] D. Crisan and T. Lyons, *A Particle Approximation of the Solution of the Kushner–Stratonovitch Equation*, Prob. Theory Related Fields, preprint.

[12] F. E. Daum, *New exact nonlinear filters*, in Bayesian Analysis of Time Series and Dynamic Models, J. C. Spall, ed., New York, 1988, pp. 199–226.

[13] P. Del Moral, *Non-linear filtering: Interacting particle resolution*, Markov Process. Related Fields, 2 (1996), pp. 555–579.

[14] S. N. Ethier and T. G. Kurtz, *Markov Processes, Characterization and Convergence*, Wiley Series in Probability and Statistics, John Wiley, New York, 1986.

[15] P. Florchingen and F. Le Gland, *Time discretization of the Zakai equation for diffusion processes observed in correlated noise*, Stochastics Stochastics Rep., 35 (1991), pp. 233–256.

[16] M. Fujisaki, G. Kallianpur, and H. Kunita, *Stochastic differential equations for the nonlinear filtering problem*, Osaka J. Math., 9 (1972), pp. 19–40.

[17] N. Gordon, D. J. Salmond, and A. F. M. Smith, *Novel approach to nonlinear/non-Gaussian bayesian state estimation*, in IEE Proceedings F, 140 (1993), pp. 107–113.

[18] G. Kallianpur and R. L. Karandikar, *The nonlinear filtering problem for the unbounded case*, Stochastic Process. Appl., 18 (1984), pp. 57–66.

[19] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Appl. Math. 23, Springer-Verlag, Berlin, 1992.

[20] F. Le Gland, *Splitting-up approximation for the SPDEs and SDEs with application to nonlinear filtering*, in Stochastic Partial Differential Equations and their Applications, Lecture Notes in Control and Inform. Sci. 176, Springer-Verlag, New York, 1992.

[21] R. S. Liptser and A. N. Shiryayev, *The Statistics of Random Processes* I. *General Theory*, Springer-Verlag, New York, 1977.

[22] S. Lototsky, R. Mikulevicius, and B. L. Rozovskii, *Nonlinear filtering revisited: A spectral approach*, SIAM J. Control Optim., 35 (1997), pp. 435–461.

[23] S. Lototsky and B. L. Rozovskii, *Recursive multiple Wiener integral expansion for nonlinear filtering of diffusion processes*, in Stochastic Processes and Functional Analysis, Lecture Notes in Pure and Appl. Math. 186, Marcel Dekker, New York, pp. 199–208.

[24] T. G. Kurtz and D. L. Ocone, *Unique characterisation of conditional distributions in nonlinear filtering*, Ann. Probab., 16 (1988), pp. 80–107.

[25] H. J. Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Mathematics in Science and Engineering 129, Academic Press, New York, 1977.

[26] G. N. Milshtein, *Approximate integration of stochastic differential equations*, Theory Probab. Appl., 19 (1974), pp. 557–562.

[27] E. Pardoux and D. Talay, *Discretization and simulation of stochastic differential equations*, Acta Appl. Math., 3 (1985), pp. 23–47.

[28] E. Pardoux, *Filtrage non linéaire et equations aux dérivées partielles stochastiques associées*, Ecole d'Eté de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math. 1464, Springer-Verlag, Berlin, 1991.

[29] S. Roelly-Coppoletta, *A criterion of convergence of measure-valued processes: Application to measure valued branching processes*, Stochastics, 17 (1986), pp. 43–65.

[30] G.C. Schmidt, *Designing nonlinear filters based on Daum's theory*, J. Guidance, Control Dynam., 16 (1993) pp. 371–376.

[31] M. Sun and R. Glowinski, *Pathwise approximation and simulation for the Zakai filtering equation through operator splitting*, Calcolo, 30 (1993), pp. 219–239.

[32] M. Yor, *Sur les théories du filtrage et de la prédiction*, in Séminaire de Probabilités XI, Lecture Notes in Math. 581, Springer-Verlag, New York, 1977, pp. 257–297.