Numerical solutions for a class of SPDEs with application to filtering

Thomas G. Kurtz[1] *and Jie Xiong*[2]

Abstract

A simulation scheme for a class of nonlinear stochastic partial differential equations is proposed and error bounds for the scheme are derived. The scheme is based on the fact that the solutions of the SPDEs can be represented by the weighted empirical measure of an infinite system of interacting particles. There are two sources of error in the scheme, one due to finite sampling of the infinite collection of particles and the other due to the Euler scheme used in the simulation of the individual particle motions. The error bounds take into account both sources of error. The results can be applied to nonlinear filtering problems.

MSC 2000 subject classifications: Primary 60H35, 60H15; Secondary 60F25, 60G35, 93E11.
Keywords: Stochastic partial differential equations, nonlinear filtering, Euler scheme, simulation, interacting infinite particle system

1. Introduction

Let $\mathcal{M}(\mathbb{R}^d)$ be the collection of all finite signed measures on $\mathbb{R}^d$, and let U be a Polish space, $\mathcal{B}(U)$ the Borel subsets of U, and μ a σ-finite Borel measure on U. Let $\mathcal{A}(U) = \{A \in \mathcal{B}(U) : \mu(A) < \infty\}$. For $1 \le i, j \le d$, let a_{ij}, b_i, d be functions on $\mathbb{R}^d \times \mathcal{M}(\mathbb{R}^d)$ and let α_i, β be functions on $\mathbb{R}^d \times \mathcal{M}(\mathbb{R}^d) \times U$. We are interested in numerical approximation for the measure-valued process V governed by the following nonlinear stochastic partial differential equation (SPDE) written in the weak form: for each

[1]Research supported in part by NSF grant DMS 96-26116
[2]This research was carried out while the second author was on leave from the University of Tennessee visiting the University of Wisconsin - Madison and the Fields Institute. Financial support from these institutes and the hospitality of the latter two are appreciated. Support was also provided by NSF grant DMS 94-24340

$\phi \in C_b^2(\mathbb{R}^d)$,

$$\begin{aligned}
&\langle \phi, V(t)\rangle \qquad (1.1)\\
&= \langle \phi, V(0)\rangle + \int_0^t \langle d(\cdot, V(s))\phi + L(V(s))\phi, V(s)\rangle\, ds \\
&\quad + \int_{U\times[0,t]} \big\langle \beta(\cdot, V(s), u)\phi + \alpha^T(\cdot, V(s), u)\nabla\phi, V(s)\big\rangle\, W(duds),
\end{aligned}$$

where for any $v \in \mathcal{M}(\mathbb{R}^d)$, $L(v)$ is a second-order differential operator

$$L(v)\phi(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x,v)\partial_{x_i}\partial_{x_j}\phi(x) + \sum_{i=1}^d b_i(x,v)\partial_{x_i}\phi(x),$$

and W is Gaussian white noise with

$$\mathbb{E}[W(A,t)W(B,t)] = \mu(A\cap B)t, \qquad \forall A, B \in \mathcal{A}(U).$$

Under appropriate conditions, we proved in [18] that V is the weighted empirical measure process of the following interacting system of diffusions:

$$\begin{aligned}
X_i(t) &= X_i(0) + \int_0^t \sigma(X_i(s), V(s))dB_i(s) \qquad (1.2)\\
&\quad + \int_0^t c(X_i(s), V(s))ds \\
&\quad + \int_{U\times[0,t]} \alpha(X_i(s), V(s), u)W(duds)
\end{aligned}$$

$$\begin{aligned}
A_i(t) &= A_i(0) + \int_0^t A_i(s)\gamma^T(X_i(s), V(s))dB_i(s) \qquad (1.3)\\
&\quad + \int_0^t A_i(s)d(X_i(s), V(s))ds \\
&\quad + \int_{U\times[0,t]} A_i(s)\beta(X_i(s), V(s), u)W(duds)
\end{aligned}$$

and

$$V(t) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n A_i(t)\delta_{X_i(t)}, \qquad (1.4)$$

where the B_i are independent standard $\mathbb{R}^d$-valued Brownian motions and σ_{ij}, c_i, γ_i, $1 \le i, j \le d$, are functions on $\mathbb{R}^d \times \mathcal{M}(\mathbb{R}^d)$ such that

$$a(x,v) = \sigma(x,v)\sigma^T(x,v) + \int_U \alpha(x,v,u)\alpha^T(x,v,u)\mu(du)$$

and

$$b(x,v)=c(x,v)+\sigma(x,v)\gamma(x,v)+\int_U\beta(x,v,u)\alpha(x,v,u)\mu(du).$$

It will be useful to note that $Z_i=\log A_i$ satisfies

$$\begin{aligned}Z_i(t) &= Z_i(0)+\int_0^t\gamma^T(X_i(s),V(s))dB_i(s)+\int_0^t d(X_i(s),V(s))ds \\ &\quad +\int_{U\times[0,t]}\beta(X_i(s),V(s),u)W(duds) \qquad (1.5)\\ &\quad -\frac{1}{2}\int_0^t\left(|\gamma(X_i(s),V(s))|^2+\int_U\beta(X_i(s),V(s),u)^2\mu(du)\right)ds.\end{aligned}$$

As in the classical Monte Carlo approximation considered, for example, in Milstein [25], Kloeden and Platen [15], and Kurtz and Protter [17], there are two sources of error in the numerical solution of the SPDE: The sampling error due to the fact that only finitely many particles are used in the approximation and the bias introduced by the approximation of the motion of each particle. For simplicity of notation, we consider the two sources of error separately. First, we study the following finite particle system:

$$\begin{aligned}X_i^n(t) &= X_i(0)+\int_0^t\sigma(X_i^n(s),V^n(s))dB_i(s) \qquad (1.6)\\ &\quad +\int_0^t c(X_i^n(s),V^n(s))ds\\ &\quad +\int_{U\times[0,t]}\alpha(X_i^n(s),V^n(s),u)W(duds)\end{aligned}$$

$$\begin{aligned}A_i^n(t) &= A_i(0)+\int_0^t A_i^n(s)\gamma^T(X_i^n(s),V^n(s))dB_i(s) \qquad (1.7)\\ &\quad +\int_0^t A_i^n(s)d(X_i^n(s),V^n(s))ds\\ &\quad +\int_{U\times[0,t]}A_i^n(s)\beta(X_i^n(s),V^n(s),u)W(duds),\end{aligned}$$

for $i=1,2,\cdots,n$, and

$$V^n(t)=\frac{1}{n}\sum_{i=1}^n A_i^n(t)\delta_{X_i^n(t)}. \qquad (1.8)$$

In Theorem 2.3 and Corollary 2.4, we give a bound on the error in estimating $V(t)$ by $V^n(t)$.

Next, we consider the approximation of the finite particle system (1.6-1.8). For $\delta > 0$, let $\{U_j^\delta : 1 \le j \le k(\delta)\}$ be a partition of U and for each j, let $u_j^\delta \in U_j^\delta$. We apply an Euler scheme to the finite particle system (1.6-1.8). The Euler step for X_i^n is given by

$$\begin{aligned}
&X_i^{n,\delta}((k+1)\delta) \qquad (1.9)\\
&\quad = X_i^{n,\delta}(k\delta) + \sigma(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta))(B_i((k+1)\delta) - B_i(k\delta))\\
&\qquad + c(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta))\delta\\
&\qquad + \sum_j \alpha(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta), u_j^\delta) W(U_j^\delta \times (k\delta, (k+1)\delta]).
\end{aligned}$$

If we used a similar Euler approximation for A_i^n, we would run the risk of the sign changing. Consequently, we approximate $Z_i^n = \log A_i^n$ instead giving

$$\begin{aligned}
&Z_i^{n,\delta}((k+1)\delta) \qquad (1.10)\\
&\quad = Z_i^{n,\delta}(k\delta) + \gamma^T(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta))(B_i((k+1)\delta) - B_i(k\delta))\\
&\qquad + \sum_j \beta(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta), u_j^\delta) W(U_j^\delta \times (k\delta, (k+1)\delta])\\
&\qquad + d(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta))\delta\\
&\qquad - \frac{1}{2}\Big(|\gamma(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta))|^2\\
&\qquad\qquad + \sum_j \beta(X_i^{n,\delta}(k\delta), V^{n,\delta}(k\delta), u_j^\delta)^2 \mu(U_j^\delta)\Big)\delta.
\end{aligned}$$

Of course

$$V^{n,\delta}(k\delta) = \frac{1}{n}\sum_{i=1}^n e^{Z_i^{n,\delta}(k\delta)} \delta_{X_i^{n,\delta}(k\delta)}.$$

Note that the random inputs are all independent Gaussian so that the scheme is implementable.

Define $\xi_\delta : U \to U$ by

$$\xi_\delta(u) = u_j^\delta, \quad u \in U_j^\delta,\ 1 \le j \le k(\delta), \qquad (1.11)$$

and set $\eta_\delta(s)=[s/\delta]\delta$. Then the solution of

$$\begin{aligned} X_i^{n,\delta}(t) &= X_i(0)+\int_0^t \sigma(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)))dB_i(s) \qquad (1.12)\\ &\quad +\int_0^t c(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)))ds\\ &\quad +\int_{U\times[0,t]} \alpha(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)),\xi_\delta(u))W(duds) \end{aligned}$$

$$\begin{aligned} A_i^{n,\delta}(t) &= A_i(0)+\int_0^t A_i^{n,\delta}(s)\gamma^T(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)))dB_i(s) \qquad (1.13)\\ &\quad +\int_0^t A_i^{n,\delta}(s)d(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)))ds\\ &\quad +\int_{U\times[0,t]} A_i^{n,\delta}(s)\beta(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)),\xi_\delta(u))W(duds) \end{aligned}$$

for $i=1,2,\cdots,n$, and

$$V^{n,\delta}(t)=\frac{1}{n}\sum_{i=1}^n A_i^{n,\delta}(t)\delta_{X_i^{n,\delta}(t)}, \qquad (1.14)$$

agrees with the Euler recursion at times that are multiples of δ. In Theorem 3.3 and Corollary 3.4, we give a bound on the error in estimating $V^n(t)$ by $V^{n,\delta}(t)$. Finally, in Theorem 4.1, we combine both estimates to obtain an error estimate for the approximation of $V(t)$ by $V^{n,\delta}(t)$. If $\delta=O(n^{-1})$, then the error is $O(1/\sqrt{n})$.

1.1. Application to filtering equations

One of the applications of the present work is to the numerical solution of the nonlinear filtering problem. To motivate our approach and for the convenience of the reader, we briefly introduce nonlinear filtering theory. We refer the reader to Kallianpur [14] and Liptser and Shiryayev [22] for a detailed treatment.

On a stochastic basis $(\Omega,\mathcal{F},\mathcal{F}_t,P)$, let X be the d-dimensional signal process governed by the following stochastic differential equation (SDE):

$$\begin{aligned} X(t)=X(0)+\int_0^t c(X(s))ds+\int_0^t \sigma(X(s))dB(s) \qquad (1.15)\\ +\int_{U\times[0,t]} \alpha(X(s),u)W(duds), \end{aligned}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, and $\alpha : \mathbb{R}^d \times U \to \mathbb{R}^d$, are measurable, W is the Gaussian white noise given in the model (1.1), and B is a d-dimensional Wiener process, independent of W. Let $h : \mathbb{R}^d \times U \to \mathbb{R}$ be a measurable map such that

$$\sup_{x\in\mathbb{R}^d} \int_{U\times[0,T]} |h(x,u)|^2 \mu(du) < \infty, \tag{1.16}$$

and let Y be the random measure on $U \times [0,T]$ given by

$$Y(A\times[0,t]) = \int_{A\times[0,t]} h(X(s),u)\mu(du)ds + W(A\times[0,t]), \qquad \forall A \in \mathcal{A}(U),\ t\in[0,T]. \tag{1.17}$$

We want to estimate the conditional distribution of X given observations of Y, that is, we want to compute the random probability measure π_t determined by

$$\pi_t f = \mathbb{E}[f(X(t))|\mathcal{F}_t^Y], \tag{1.18}$$

where $\mathcal{F}_t^Y = \sigma\{Y(A\times[0,s]) : A \in \mathcal{B}(U),\ 0 \le s \le t\}$ is the information available up to time t.

Remark 1.1 *If $U \subset \mathbb{R}^d$, then Y models spatial observations. If U is a space containing only m points, then Y and W can be regarded as m-dimensional processes and (1.17) becomes the classical observation model.*

Under (1.16), we can assume that there exists a *reference measure* on $(\Omega, \mathcal{F}, \mathcal{F}_t)$ such that for each $t > 0$, $P\ |_{\mathcal{F}_t} << Q\ |_{\mathcal{F}_t}$, the Radon Nikodym derivative $\frac{dP}{dQ}$ on $\mathcal{F}_t$ is given by

$$A(t) = e^{\left(\int_{U\times[0,T]} h(X(s),u)Y(duds) - \frac{1}{2}\int_{U\times[0,T]} |h(X(s),u)|^2\mu(du)ds\right)},$$

and under Q, Y is Gaussian white noise with covariance measure μ and is independent of B.

By the Kallianpur-Striebel formula, we have

$$\pi_t f = \frac{\mu_t f}{\mu_t 1}$$

where

$$\mu_t f = \mathbb{E}\left[f(X(t))A(t)|\mathcal{F}_t^Y\right].$$

The random measure μ_t solves the following Zakai equation which is a special case of (1.1):

$$\begin{aligned}\mu_t f &= \pi_0 f + \int_0^t \mu_s(\mathcal{L}f)ds \qquad (1.19)\\ &\quad + \int_0^t \mu_s(\nabla f\cdot\alpha(\cdot,u)+h(\cdot,u)f)Y(duds), \quad \forall f\in C_b^2(\mathbb{R}^d),\end{aligned}$$

where

$$\mathcal{L}f(x)=\frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\partial_{x_i}\partial_{x_j}f(x)+\sum_{i=1}^d c_i(x)\partial_{x_i}f(x)$$

with

$$a_{ij}(x)=\sum_{k=1}^d\sigma_{ik}(x)\sigma_{kj}(x)+\int_U\alpha_i(x,u)\alpha_j(x,u)\mu(du).$$

In the case of U finite, the uniqueness of the solution to (1.19) has been discussed by various authors (for example, Szpirglas [28], Fujisaki, Kallianpur, and Kunita [12], Kurtz and Ocone [16], Rozovskii [27] and Bhatt, Kallianpur and Karandikar [1]. Under the boundedness and Lipschitz conditions we assume here, uniqueness for the general case follows by the results of [18].

Note that

$$A(t)=1+\int_{U\times[0,t]}A(s)h(X(s),u)Y(duds).$$

To approximate μ_t, we consider the following particle system

$$\begin{aligned}X_j(t)=X_j(0)+\int_0^t b(X_j(s))ds+\int_0^t\sigma(X_j(s))dB_j(s)\\ +\int_{U\times[0,t]}\alpha(X_j(s),u)Y(duds)\end{aligned}$$

and

$$A_j(t)=1+\int_{U\times[0,t]}A_j(s)h(X_j(s),u)Y(duds), \qquad (1.20)$$

$j=1,2,\cdots,n$, where $b(x)=c(x)-\int\alpha(x,u)h(x,u)\mu(du)$, and $B_1,\cdots,B_n$ are independent Brownian motions, independent of Y under Q. Define

$$\mu_t^n=\frac{1}{n}\sum_{j=1}^n A_j(t)\delta_{X_j(t)}. \qquad (1.21)$$

We will show that $\mu_t^n \to \mu_t$ and study the convergence rate as a special case of the approximation problem for the empirical measure process V discussed above.

Numerical solution of the filtering problem has been studied extensively in the classical setting (U finite), although much of the work has been done under the assumption that the observation noise is independent of the signal. Kushner [19, 20, 21] develops approximation methods based on replacing the signal process by a finite state Markov chain that approximates the signal. In the simplest cases, this method is equivalent to a finite difference approximation in the filtering equation. Picard [26] considers a time discretization of the Zakai equation involving the replacement of the signal by a discrete-time process and discrete-time approximations of the Radon-Nikodym derivative in the Kallinapur-Striebel formula. The error in the approximation is $O(\delta)$, where δ is the time step. The approximations still involve integrals against process distributions, and Picard suggests a Monte Carlo scheme to implement the approximation. Di Masi, Pratelli, and Runggaldier [8] consider a similar time discretization, but they also introduce a signal approximation that reduces the problem to a finite dimensional computation somewhat similar to the approach taken by Kushner. Lototsky and Rozovskii [23] and Lototsky, Mikulevicius and Rozovskii [24] derive algorithms based on a Wiener chaos decomposition. This point of view is also explored by Budhiraja and Kallianpur [2]. Hu, Kallianpur and Xiong [13] considered a Wong-Zakai type approximation. An error bound of the order of $\sqrt{\delta}$ was obtained, where δ is the size of the discretization time step.

Florchinger and Le Gland ([9], [10]) consider a time-discretization of the Zakai equation for diffusion processes observed in correlated noise based on a split-up approximation and a Trotter-like product formula. The error estimate is also of the order of $\sqrt{\delta}$. In [11], a particle approximation is formulated similar to the one considered here. Del Moral [7] considers a particle approximation for a model with independent observation noise that discounts past information. His results give convergence uniform in time but without a rate.

Crisan and Lyons [6] and Crisan, Gaines and Lyons [5] derive an approximation for the independent noise problem based on an interacting, branching particle system. For a closely related method, Crisan, Del Moral, and Lyons [3] give an error bound of order $n^{-1/4}$, where n is essentially the number of particles. Recently, Crisan, Del Moral and Lyons [3] considered the numerical solution for the filtering problem with discrete time parameter. The error bound they derive is of the order of $\frac{1}{\sqrt{n}}$. These branching models attempt to reduce the variance of the approximation by avoiding the weights A_i in the empirical measure process. Roughly, the schemes kill particles that would have small A_i and replicate particles that would have large A_i.

Simulation results by various investigators support the argument that some kind of branching or resampling improves the accuracy of particle approximations. The results of the present paper demonstrate that the error of simple (non-branching) Monte Carlo integration of the Kallianpur-Striebel formula is of the same (or better) *order* as the branching/resampling methods. Branching/resampling can only improve the error by reducing the conditional variance of the empirical measure, and since at the time of a branching or killing event, the conditional variance will typically increase, considerable care needs to be taken to ensure that branching/resampling does not make the error worse.

1.2. Organization of paper

In section 2, we estimate the error resulting from replacing the infinite particle system by a finite particle system, that is, the error due to *sampling.* In section 3, we consider the Euler scheme for the finite system and estimate the error due to the *time discretization* and to the *space discretization* needed to approximate the Gaussian white noise ($W(duds)$) integrals. Finally, in Section 4, we combine sampling and the Euler scheme to obtain an approximation of V and its error bound.

2. Sampling error

In this section, we bound the error caused by replacing V by a finite empirical measure V^n.

Recall that for $\nu_1, \nu_2 \in \mathcal{M}_+(\mathbb{R}^d)$, the Wasserstein metric is given by

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

where

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

We will be dealing with measures of the form $\nu^k = \frac{1}{n} \sum_{i=1}^n a_i^k \delta_{x_i^k}$, and it is useful to note that in this case

$$\rho(\nu_1, \nu_2) \leq \frac{1}{n} \sum_{i=1}^n a_i^1 \vee a_i^2 (|x_i^1 - x_i^2| + |\log a_i^1 - \log a_i^2|). \qquad (2.1)$$

For simplicity of notation, we restrict our attention to $\mathcal{M}_+(\mathbb{R}^d)$-valued processes. To this end, we make the following assumption:

(I) $\{(A_i(0), X_i(0))\}$ is an *iid* sequence which is independent of $\{B_i\}$ and W. $A_1(0) \geq 0$ a.s. and

$$\mathbb{E}A_1(0)^2 + \mathbb{E}|X_1(0)|^2 < \infty.$$

The following assumptions were made in [18] for the existence and uniqueness of the solution of the SPDE (1.1).

(S1) There exists a constant K such that for each $x \in \mathbb{R}^d$, $\nu \in \mathcal{M}_+(\mathbb{R}^d)$,

$$\begin{aligned}|\sigma(x,\nu)|^2 + |c(x,\nu)|^2 + \int_U |\alpha(x,\nu,u)|^2\mu(du) & \\ +|\gamma(x,\nu)|^2 + |d(x,\nu)|^2 + \int_U \beta(x,\nu,u)^2\mu(du) &\leq K^2.\end{aligned}$$

(S2) For each $x_1, x_2 \in \mathbb{R}^d$, $\nu_1, \nu_2 \in \mathcal{M}_+(\mathbb{R}^d)$

$$\begin{aligned}&|\sigma(x_1,\nu_1) - \sigma(x_2,\nu_2)|^2 + |c(x_1,\nu_1) - c(x_2,\nu_2)|^2 \\ &+|\gamma(x_1,\nu_1) - \gamma(x_1,\nu_1)|^2 + \int_U |\alpha(x_1,\nu_1,u) - \alpha(x_2,\nu_2,u)|^2\mu(du) \\ &+|d(x_1,\nu_1) - d(x_2,\nu_2)|^2 + \int_U |\beta(x_1,\nu_1,u) - \beta(x_2,\nu_2,u)|^2\mu(du) \\ &\quad \leq K^2(|x_1 - x_2|^2 + \rho(\nu_1,\nu_2)^2).\end{aligned}$$

With reference to (1.13), let

$$\begin{aligned}M_i^n(t) &= \int_0^t \gamma^T(X_i^n(s), V^n(s))dB_i(s) \\ &\qquad + \int_{U\times[0,t]} \beta(X_i^n(s), V^n(s), u)W(duds).\end{aligned}$$

Then $M_i^n(t)$ is a martingale and

$$\begin{aligned}\langle M_i^n\rangle_t &= \int_0^t |\gamma(X_i^n(s), V^n(s))|^2 ds \\ &\qquad + \int_{U\times[0,t]} \beta(X_i^n(s), V^n(s), u)^2\mu(du)ds \\ &\leq K^2 t.\end{aligned}$$

An application of Itô's formula shows that the solution of (1.13) is given by

$$A_i^n(t) = A_i(0)\exp\left(M_i^n(t) - \frac{1}{2}\langle M_i^n\rangle_t + \int_0^t d(X_i^n(s), V^n(s))ds\right). \quad (2.2)$$

Proposition 2.1 *Suppose that Assumptions (I) and (S1) hold. Then for each n*

$$\mathbb{E} \sup_{0\le s\le T} |A_i^n(s)|^2 < 4\mathbb{E}|A_i(0)|^2 e^{(K^2+2K)t}, \tag{2.3}$$

and the same bound holds with A_i^n replace by A_i.

Proof. By (2.2),

$$\begin{aligned} A_i^n(t) &= A_i(0)\exp\left(M_i^n(t) - \frac{1}{2}\langle M_i^n\rangle_t + \int_0^t d(X_i^n(s), V^n(s))ds\right) \\ &\le A_i(0)\exp\left(M_i^n(t) - \frac{1}{2}\langle M_i^n\rangle_t\right) e^{Kt}, \end{aligned}$$

and $A_i(0)\exp\left(M_i^n(t) - \frac{1}{2}\langle M_i^n\rangle_t\right)$ is a square integrable martingale with

$$\mathbb{E}A_i(0)^2\exp\left(2M_i^n(t) - \langle M_i^n\rangle_t\right) \le \mathbb{E}A_i(0)^2 e^{K^2 t}.$$

Consequently, (2.3) follows by Doob's inequality. □

We make the following additional assumption:

(S3) There exists a constant K such that for any *iid* sequence (ξ_i, η_i), $i = 1, 2, \cdots$ and $x \in \mathbb{R}^d$, we have

$$\mathbb{E}\left|\sigma\left(x, \frac{1}{n}\sum_{i=1}^n \xi_i\delta_{\eta_i}\right) - \sigma(x,\mu)\right|^2 \le \frac{K^2\mathbb{E}\xi_1^2}{n},$$

where $\mu(\cdot) = \mathbb{E}\left[\xi_1 1_{\eta_1\in\cdot}\right]$, and the same inequality holds for the other coefficients.

Remark 2.2 *i) If $\sigma(x,\nu)$ does not depend on ν, then (S3) holds.*
ii) If $\sigma(x,\nu) = \int_{\mathbb{R}^d}\sigma_1(x,y)\nu(dy)$ and $|\sigma_1(x,y)| \le K$, for all $x, y \in \mathbb{R}^d$, then (S3) holds.
iii) If $\sigma(x,\nu) = h(x, \langle\psi_1,\nu\rangle, \ldots, \langle\psi_m,\nu\rangle)$, $\psi_1, \ldots, \psi_m \in \bar{C}(\mathbb{R}^d)$, and there exists K such that

$$|h(x, z_1, \ldots, z_m) - h(x, y_1, \ldots, y_m)| \le K\sum_{i=1}^m |z_i - y_i|,$$

then (S3) holds.

Theorem 2.3 *Assume (I) and (S1)-(S3). For* $T > 0$,

$$\mathbb{E}\left(\sup_{t\le T\wedge\eta_m^n}|X_i^n(t)-X_i(t)|^2+\sup_{t\le T\wedge\eta_m^n}|Z_i^n(t)-Z_i(t)|^2\right)\le\frac{c_1(T,m)}{n},\quad(2.4)$$

where $c_1(T,m)$ *is a constant,*

$$\eta_m^n=\inf\left\{t:\frac{1}{n}\sum_{i=1}^n A_i^n(t)^2>m^2\text{ or }\lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^k A_i(t)^2>m^2\right\},$$

and

$$\mathbb{P}\{\eta_m^n<T\}\le\frac{8e^{(K^2+K)T}\mathbb{E}A_i(0)^2}{m^2}.\quad(2.5)$$

Proof. By Doob's inequality and Hölder's inequality, for $t\le T$,

$$\begin{aligned}&\mathbb{E}\sup_{r\le t\wedge\eta_m^n}|X_i^n(r)-X_i(r)|^2 &(2.6)\\ &\le 12\mathbb{E}\int_0^t|\sigma(X_i^n(s),V^n(s))-\sigma(X_i(s),V(s))|^2 1_{s\le\eta_m^n}ds\\ &\quad+3T\mathbb{E}\int_0^t|c(X_i^n(s),V^n(s))-c(X_i(s),V(s))|^2 1_{s\le\eta_m^n}ds\\ &\quad+12\mathbb{E}\int_0^t\int_U|\alpha(X_i^n(s),V^n(s),u)\\ &\qquad\qquad-\alpha(X_i(s),V(s),u)|^2\mu(du)1_{s\le\eta_m^n}ds.\end{aligned}$$

Let

$$\tilde V^n(t)=\frac{1}{n}\sum_{i=1}^n A_i(t)\delta_{X_i(t)}\qquad\text{and}\qquad\tilde V_i^n(t)=\frac{1}{n-1}\sum_{j=1,\ j\ne i}^n A_j(t)\delta_{X_j(t)}.$$

Then

$$\begin{aligned}&\mathbb{E}|\sigma(X_i^n(s),V^n(s))-\sigma(X_i(s),V(s))|^2 1_{s\le\eta_m^n} &(2.7)\\ &\le 3\mathbb{E}|\sigma(X_i^n(s),V^n(s))-\sigma(X_i(s),\tilde V^n(s))|^2 1_{s\le\eta_m^n}\\ &\quad+3\mathbb{E}|\sigma(X_i(s),\tilde V^n(s))-\sigma(X_i(s),\tilde V_i^n(s))|^2 1_{s\le\eta_m^n}\\ &\quad+3\mathbb{E}|\sigma(X_i(s),\tilde V_i^n(s))-\sigma(X_i(s),V(s))|^2 1_{s\le\eta_m^n}\\ &\le 3K^2\mathbb{E}\left(|X_i^n(s)-X_i(s)|^2+\rho(V^n(s),\tilde V^n(s))^2\right)1_{s\le\eta_m^n}\\ &\quad+3K^2\mathbb{E}\rho(\tilde V^n(s),\tilde V_i^n(s))^2 1_{s\le\eta_m^n}\\ &\quad+3\mathbb{E}\mathbb{E}[|\sigma(X_i(s),\tilde V_i^n(s))-\sigma(X_i(s),V(s))|^2|W,X_i].\end{aligned}$$

By (2.1),

$$
\begin{aligned}
&\rho(V^n(s),\tilde V^n(s)) \qquad\qquad (2.8)\\
&\le \frac{1}{n}\sum_{j=1}^n A_j^n(s)\vee A_j(s)(|X_j^n(s)-X_j(s)|+|Z_j^n(s)-Z_j(s)|)\\
&\le \sqrt{\frac{1}{n}\sum_{j=1}^n (A_j^n(s)\vee A_n(s))^2}\Big((\frac{1}{n}\sum_{j=1}^n |X_j^n(s)-X_j(s)|^2)^{\frac12}\\
&\qquad +(\frac{1}{n}\sum_{j=1}^n |Z_j^n(s)-Z_j(s)|^2)^{\frac12}\Big).
\end{aligned}
$$

A simple calculation gives

$$\rho(\tilde V^n(s),\tilde V_i^n(s)) \le \frac{1}{n}A_i(s)+\frac{1}{n(n-1)}\sum_{j=1}^n A_j(s).$$

Let

$$f_m^n(t)=\mathbb{E}\sup_{r\le t\wedge\eta_m^n}\left|X_i^n(r)-X_i(r)\right|^2,\quad g_m^n(t)=\mathbb{E}\sup_{r\le t\wedge\eta_m^n}\left|Z_i^n(r)-Z_i(r)\right|^2.$$

By (2.8) and the definition of η_m^n,

$$\mathbb{E}\sup_{r\le t\wedge\eta_m^n}\rho(V^n(t),\tilde V^n(t))^2\le 4m^2(f_m^n(t)+g_m^n(t)).$$

Then for the right side of (2.7),

$$\text{1st term}\le 3K^2\left((1+4m^2)f_m^n(s)+4m^2g_m^n(t)\right)$$

and

$$\text{2nd term}\le 6K^2\left(\frac{1}{n^2}\mathbb{E}|A_i(s)|^2+\frac{1}{(n-1)^2}m^2\right).$$

Since, conditioned on (W,X_i), (A_j,X_j), $j\neq i$, are *iid*, by (S3), we have

$$\text{3rd term}\le 3\mathbb{E}\left(\frac{K^2}{n-1}\mathbb{E}\left(A_1(s)^2|W,X_i\right)\right)=\frac{3K^2}{n-1}\mathbb{E}A_1(s)^2.$$

Hence, the first term on the right side of (2.6) is dominated by

$$
\begin{aligned}
&12\int_0^t 3K^2\left((1+4m^2)f_m^n(s)+4m^2g_m^n(s)\right)ds\\
&\quad+12\left[6K^2\left(\frac{1}{n^2}\sup_{s\le T}\mathbb{E}A_i(s)^2+\frac{1}{n-1}m^2\right)+\frac{3K^2}{n-1}\sup_{s\le T}\mathbb{E}A_1(s)^2\right]T.
\end{aligned}
$$

Similar estimates hold for the other terms on the right side of (2.6). Therefore, there exist constants $c_2(T,m)$ and $c_3(T,m)$ such that

$$f_m^n(t) \leq c_2(T,m)\int_0^t (f_m^n(s)+g_m^n(s))ds + \frac{c_3(T,m)}{n}. \tag{2.9}$$

Similar arguments give

$$g_m^n(t) \leq c_4(T,m)\int_0^t (f_m^n(s)+g_m^n(s))ds + \frac{c_5(T,m)}{n}. \tag{2.10}$$

Therefore

$$f_m^n(t)+g_m^n(t) \leq (c_2+c_4)\int_0^t (f_m^n(s)+g_m^n(s))ds + \frac{c_3+c_5}{n},$$

and by Gronwall's inequality, we have

$$f_m^n(t)+g_m^n(t) \leq \frac{c_3+c_5}{n}e^{(c_2+c_4)t},$$

giving (2.4).

(2.5) follows from (2.3). □

For a bounded Lipschitz function f, define

$$\|f\|_L = \sup_{x\in\mathbb{R}^d}|f(x)| + \sup_{x,y\in\mathbb{R}^d}\frac{|f(x)-f(y)|}{|x-y|}.$$

Corollary 2.4 *Assume (I) and (S1)-(S3). For each bounded Lipschitz function f and each $t\geq 0$,*

$$\mathbb{E}|V^n(t)f - V(t)f|1_{t\leq\eta_m^n} \leq \frac{c_6(t,m)\|f\|_L}{\sqrt{n}}. \tag{2.11}$$

Proof. Noting that $(a\vee b)^2 \leq a^2+b^2$, (2.1), Hölder's inequality, and

Theorem 2.3 give

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T\wedge\eta_m^n}|V^n(t)f-\tilde V^n(t)f| \qquad (2.12)\\
&\le \|f\|_L\mathbb{E}\sup_{t\le T\wedge\eta_m^n}\rho(V^n(t),\tilde V^n(t))\\
&\le \|f\|_L(\mathbb{E}\sup_{t\le T\wedge\eta_m^n}\frac{1}{n}\sum_{j=1}^n(A_j(t)\vee A_j^n(t))^2)^{1/2}\\
&\qquad(\mathbb{E}\sup_{t\le T\wedge\eta_m^n}\frac{1}{n}\sum_{j=1}^n|X_j^n(t)-X_j(t)|^2)^{1/2}\\
&\quad+\|f\|_L(\mathbb{E}\sup_{t\le T\wedge\eta_m^n}\frac{1}{n}\sum_{j=1}^n(A_j(t)\vee A_j^n(t))^2)^{1/2}\\
&\qquad(\mathbb{E}\sup_{t\le T\wedge\eta_m^n}\frac{1}{n}\sum_{j=1}^n|Z_j^n(t)-Z_j(t)|^2)^{1/2}\\
&\le \frac{c_7(T,m)\|f\|_L}{\sqrt{n}}.
\end{aligned}$$

In addition,

$$\begin{aligned}
\mathbb{E}|\tilde V^n(t)f-V(t)f| &\le \sqrt{\mathbb{E}|\tilde V^n(t)f-V(t)f|^2}\\
&= \sqrt{\frac{1}{n}\mathbb{E}\left(A_1(t)f(X_1(t))-\mathbb{E}(A_1(t)f(X_1(t))|W)\right)^2}\\
&\le \frac{c_8(t)\|f\|_L}{\sqrt{n}}, \qquad (2.13)
\end{aligned}$$

and (2.11) follows from (2.12) and (2.13). □

The rate of convergence given by (2.11) can be viewed as convergence in a metric for $\mathcal{M}_+(\mathbb{R}^d)$ under which convergence is equivalent to weak convergence. The metric we define is similar to that used in [6] and [13]. Let $\{f_k\}$ be a dense subset of $C_b(\mathbb{R}^d)$ such that $\|f_k\|_L<\infty$, for each k. Define

$$\tilde\rho(\nu_1,\nu_2)=\sum_{k=1}^\infty\frac{|\langle\nu_1,f_k\rangle-\langle\nu_2,f_k\rangle|}{2^k\|f_k\|_L}.$$

Note that $\tilde\rho(\nu_1,\nu_2)\le\rho(\nu_1,\nu_2)$, for all $\nu_1,\nu_2\in\mathcal{M}_+(\mathbb{R}^d)$. The estimate in (2.11) implies the following:

Corollary 2.5 *For each $t\ge 0$,*

$$\mathbb{E}\tilde\rho(V^n(t),V(t))1_{t\le\eta_m^n}\le\frac{c_6(t,m)}{\sqrt{n}}. \qquad (2.14)$$

Proof. (2.14) is a direct consequence of (2.11) and the definition of $\tilde{\rho}$. □

2.1. Application to filtering equations

To verify (I), (S1) and (S2) for the filtering problem, we make the following assumptions:

(I′) $\{X_i(0)\}$ is an *iid* sequence which is independent of $\{B_i\}$ and Y, and

$$\mathbb{E}|X_1(0)|^2 < \infty.$$

(F1) There exists a constant K such that for each $x \in \mathbb{R}^d$ and $\nu \in \mathcal{M}^+(\mathbb{R}^d)$,

$$|b(x)|^2 + |\sigma(x)|^2 + \int_U |h(x,u)|^2 \mu(du) + \int_U |\alpha(x,\nu,u)|^2 \mu(du) \le K^2.$$

(F2) For each $x_1, x_2 \in \mathbb{R}^d$, $\nu_1, \nu_2 \in \mathcal{M}^+(\mathbb{R}^d)$,

$$\begin{aligned}
&|b(x_1) - b(x_2)|^2 + |\sigma(x_1) - \sigma(x_2)|^2 \\
&+ \int_U |\alpha(x_1,\nu_1,u) - \alpha(x_2,\nu_2,u)|^2 \mu(du) \\
&\qquad + \int_U |h(x_1,u) - h(x_2,u)|^2 \mu(du) \\
&\qquad\qquad \le K^2 |x_1 - x_2|^2.
\end{aligned}$$

Corollary 2.6 *Assume (I′), (F1) and (F2). Under both the model measure P and the reference measure Q, for each bounded, Lipschitz function f and each $t \ge 0$,*

$$\mathbb{E}|\mu_t^n f - \mu_t f| \le \frac{c_9(t)\|f\|_L}{\sqrt{n}}, \tag{2.15}$$

and hence,

$$\mathbb{E}\tilde{\rho}(\mu_t^n, \mu_t) \le \frac{c_9(t)}{\sqrt{n}}. \tag{2.16}$$

Proof. Since the coefficients do not depend on μ_t and μ_t^n, we do not need to introduce the stopping time η_m^n in the analysis of (2.7), and hence, in the statement of the final estimate. Under (I′), (F1) and (F2), it is clear that (I), (S1) and (S2) hold for the filtering problem. Since $X_i^n(t) = X_i(t)$, $\mu_t^n = \tilde{\mu}_t^n$, for the reference measure Q under which Y is Gaussian white noise, (2.15) follows from (2.13).

Let $dP = A(t)dQ$, so that under P, Y is the observation process satisfying (1.17) with W being Gaussian white noise. Let $\tilde{\mu}_t^n$ be defined as $\tilde{V}^n(t)$ is in the proof of Theorem 2.3. Then as in (2.12),

$$\begin{aligned} &\mathbb{E}^P \sup_{t\le T} |\mu_t^n f - \tilde{\mu}_t^n f| \qquad (2.17)\\ &\le \|f\|_L \mathbb{E}^P \sup_{t\le T} \rho(\mu_t^n, \tilde{\mu}_t^n)\\ &= \|f\|_L \mathbb{E}^Q A(t) \sup_{t\le T} \rho(\mu_t^n, \tilde{\mu}_t^n)\\ &\le \|f\|_L (\mathbb{E}^Q \sup_{t\le T} \frac{1}{n}\sum_{j=1}^n A(t)^2 (A_j(t) \vee A_j^n(t))^2)^{1/2}\\ &\qquad (\mathbb{E}^Q \sup_{t\le T} \frac{1}{n}\sum_{j=1}^n |X_j^n(t) - X_j(t)|^2)^{1/2}\\ &\quad + \|f\|_L (\mathbb{E}^Q \sup_{t\le T} \frac{1}{n}\sum_{j=1}^n A(t)^2 (A_j(t) \vee A_j^n(t))^2)^{1/2}\\ &\qquad (\mathbb{E}^Q \sup_{t\le T} \frac{1}{n}\sum_{j=1}^n |Z_j^n(t) - Z_j(t)|^2)^{1/2}\\ &\le \frac{c_7^\mu(T,m)\|f\|_L}{\sqrt{n}}. \end{aligned}$$

Note that

$$\begin{aligned} &\mathbb{E}^Q \sup_{t\le T} \frac{1}{n}\sum_{j=1}^n A(t)^2 (A_j(t) \vee A_j^n(t))^2 \qquad (2.18)\\ &\le \sqrt{\mathbb{E}^Q \sup_{t\le T} A(t)^4 \mathbb{E}^Q \sup_{t\le T} A_j(t)^4} + \sqrt{\mathbb{E}^Q \sup_{t\le T} A(t)^4 \mathbb{E}^Q \sup_{t\le T} A_j^n(t)^4} \end{aligned}$$

and the fourth moments exist by the same argument employed in the proof of Proposition 2.1. Finally, as in (2.13),

$$\begin{aligned} \mathbb{E}^P |\tilde{\mu}_t^n f - \mu_t f| = \mathbb{E}^Q A(t) |\tilde{\mu}_t^n f - \mu_t f| &\le \sqrt{\mathbb{E}^Q A(t)^2 \mathbb{E}^Q |\mu_t^n f - \mu_t f|^2}\\ &\le \frac{\tilde{c}_9(T)\|f\|_L}{\sqrt{n}}, \end{aligned}$$

and combining (2.17) and (2.18), we have an estimate of the form (2.15).

In both cases, (2.16) is a direct consequence of (2.15) and the definition of $\tilde{\rho}$. □

3. Euler scheme

In this section, we consider an error bound for the Euler scheme (1.12-1.14) for the finite particle system (1.6-1.8). Combined with the results of the previous section, we obtain an error bound for a numerical method for the SPDE (1.1). Throughout this section we assume that $0 < \delta \leq 1$.

Recalling the definition of ξ_δ in (1.11), we need the following assumptions:

(S4) There exists a constant K such that for each $x \in \mathbb{R}^d$, $\nu \in \mathcal{M}_+(\mathbb{R}^d)$,

$$\int_U |\alpha(x,\nu,u) - \alpha(x,\nu,\xi_\delta(u))|^2 \mu(du) \leq K^2\delta \tag{3.1}$$

and

$$\int_U |\beta(x,\nu,u) - \beta(x,\nu,\xi_\delta(u))|^2 \mu(du) \leq K^2\delta. \tag{3.2}$$

Remark 3.1 *If α and β are Lipschitz functions and*

$$\int_U d_U(u,\xi_\delta(u))^2 \mu(du) \leq K_1^2 \delta, \tag{3.3}$$

then (S4) holds.

Example 3.2 *i) Let $U = [0,1)$ and let μ be Lebesgue measure. Take $k(\delta) = \left[\delta^{-1/2}\right]$ and*

$$U_j^\delta = [(j-1)\sqrt{\delta}, (j\sqrt{\delta}) \wedge 1), \qquad j = 1, 2, \cdots, k(\delta).$$

Then (3.3) holds.

ii) Let $U = \mathbb{R}$ and let μ be the standard Gaussian measure. Take $k(\delta) = 2\left[\delta^{-1}\right] + 2$ and

$$U_j^\delta = \begin{cases} [(j-1)\sqrt{\delta}, j\sqrt{\delta}) & j = 1, 2, \cdots, \left[\delta^{-1}\right], \\ \Big[(\left[\delta^{-1}\right] - j)\sqrt{\delta}, (\left[\delta^{-1}\right] - j + 1)\sqrt{\delta}\Big), & j = \left[\delta^{-1}\right] + 1, \cdots, 2\left[\delta^{-1}\right], \\ \Big[\left[\delta^{-1}\right]\sqrt{\delta}, \infty\Big), & j = 2\left[\delta^{-1}\right] + 1 \\ \Big(-\infty, -\left[\delta^{-1}\right]\sqrt{\delta}\Big), & j = 2\left[\delta^{-1}\right] + 2. \end{cases}$$

Then (3.3) holds.

Theorem 3.3 *Assume (I) and (S1)-(S4). For each $T > 0$,*

$$\mathbb{E}\Big(\sup_{t \leq T\wedge\eta_m^{n,\delta}} |X_i^{n,\delta}(t) - X_i^n(t)|^2 + \sup_{t \leq T\wedge\eta_m^{n,\delta}} |Z_i^{n,\delta}(t) - Z_i^n(t)|^2\Big) \leq c_{10}(T,m)\delta, \tag{3.4}$$

where $c_{10}(T,m)$ *is a constant,*

$$\eta_m^{n,\delta} = \inf\left\{t : \frac{1}{n}\sum_{i=1}^n A_i^n(t)^2 > m^2 \text{ or } \frac{1}{n}\sum_{i=1}^n A_i^{n,\delta}(t)^2 > m^2\right\},$$

and

$$P\{\eta_m^{n,\delta} < T\} \leq \frac{8e^{(K^2+K)T}\mathbb{E}A_i(0)^2}{m^2}.$$

Proof. Since

$$\begin{aligned}
&X_i^{n,\delta}(t) - X_i^n(t)\\
&\quad= \int_0^t \Big(\sigma(X_i^{n,\delta}(\eta_\delta(s)), V^{n,\delta}(\eta_\delta(s))) - \sigma(X_i^n(s), V^n(s))\Big)\, dB_i(s)\\
&\qquad+ \int_0^t \Big(c(X_i^{n,\delta}(\eta_\delta(s)), V^{n,\delta}(\eta_\delta(s))) - c(X_i^n(s), V^n(s))\Big)\, ds\\
&\qquad+ \int_{U\times[0,t]} \Big(\alpha(X_i^{n,\delta}(\eta_\delta(s)), V^{n,\delta}(\eta_\delta(s)), \xi_\delta(u))\\
&\qquad\qquad\qquad -\alpha(X_i^n(s), V^n(s), u)\Big) W(duds),
\end{aligned}$$

by Doob's inequality and Hölder's inequality,

$$
\begin{aligned}
&\mathbb{E} \sup_{t\le T\wedge\eta_m^{n,\delta}} |X_i^{n,\delta}(t)-X_i^n(t)|^2 \qquad (3.5)\\
&\le 24\int_0^T \mathbb{E}\Big|\sigma(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)))\\
&\qquad\qquad -\sigma(X_i^{n,\delta}(s),V^{n,\delta}(s))\Big|^2 1_{s<\eta_m^{n,\delta}}ds\\
&\quad +6T\int_0^T \mathbb{E}\Big|c(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)))\\
&\qquad\qquad -c(X_i^{n,\delta}(s),V^{n,\delta}(s))\Big|^2 1_{s<\eta_m^{n,\delta}}ds\\
&\quad +24\int_{U\times[0,T]} \mathbb{E}\Big|\alpha(X_i^{n,\delta}(\eta_\delta(s)),V^{n,\delta}(\eta_\delta(s)),\xi_\delta(u))\\
&\qquad\qquad -\alpha(X_i^{n,\delta}(s),V^{n,\delta}(s),u)\Big|^2 \mu(du)1_{s<\eta_m^{n,\delta}}ds\\
&\quad +24\int_0^T \mathbb{E}\Big|\sigma(X_i^{n,\delta}(s),V^{n,\delta}(s))-\sigma(X_i^n(s),V^n(s))\Big|^2 1_{s<\eta_m^{n,\delta}}ds\\
&\quad +6T\int_0^T \mathbb{E}\Big|c(X_i^{n,\delta}(s),V^{n,\delta}(s))-c(X_i^n(s),V^n(s))\Big|^2 1_{s<\eta_m^{n,\delta}}ds\\
&\quad +24\int_{U\times[0,T]} \mathbb{E}\Big|\alpha(X_i^{n,\delta}(s),V^{n,\delta}(s),u)\\
&\qquad\qquad -\alpha(X_i^n(s),V^n(s),u)\Big|^2 \mu(du)1_{s<\eta_m^{n,\delta}}ds,
\end{aligned}
$$

with a similar inequality holding for $\mathbb{E}\sup_{t\le T\wedge\eta_m^{n,\delta}}|Z_i^{n,\delta}(t)-Z_i^n(t)|^2$.

Note that

$$
\begin{aligned}
&\mathbb{E}|X_i^{n,\delta}(t)-X_i^{n,\delta}(\eta_\delta(t))|^2 \qquad (3.6)\\
&\le 3(t-\eta_\delta(t))\mathbb{E}|\sigma(X_i^{n,\delta}(\eta_\delta(t)),V^{n,\delta}(\eta_\delta(t)))|^2\\
&\quad +3(t-\eta_\delta(t))^2\mathbb{E}|c(X_i^{n,\delta}(\eta_\delta(t)),V^{n,\delta}(\eta_\delta(t)))|^2\\
&\quad +3(t-\eta_\delta(t))\mathbb{E}\int_U |\alpha(X_i^{n,\delta}(\eta_\delta(t)),V^{n,\delta}(\eta_\delta(t)),u)|^2\mu(du)\\
&\le K^2\left(6(t-\eta_\delta(t))+3(t-\eta_\delta(t))^2\right)
\end{aligned}
$$

and that a similar bound holds for $\mathbb{E}|Z_i^{n,\delta}(t)-Z_i^{n,\delta}(\eta_\delta(t))|^2$. Since by (2.1),

$$
\begin{aligned}
&\rho(V^{n,\delta}(t),V^{n,\delta}(\eta_\delta(t)))\\
&\le \frac{1}{n}\sum_{i=1}^n A_i^{n,\delta}(t)\vee A_i^{n,\delta}(\eta_\delta(t))(|X_i^{n,\delta}(t)-X_i^{n,\delta}(\eta_\delta(t))|\\
&\qquad\qquad +|Z_i^{n,\delta}(t)-Z_i^{n,\delta}(\eta_\delta(t))|),
\end{aligned}
$$

we have

$$\begin{aligned}
&\mathbb{E}\rho^2(V^{n,\delta}(t), V^{n,\delta}(\eta_\delta(t)))1_{t<\eta_m^{n,\delta}} \\
&\leq 4m^2\mathbb{E}(\frac{1}{n}\sum_{i=1}^n |X_i^{n,\delta}(t) - X_i^{n,\delta}(\eta_\delta(t))| + |Z_i^{n,\delta}(t) - Z_i^{n,\delta}(\eta_\delta(t))|)^2 \\
&\leq c_{11}(m)(t - \eta_\delta(t)).
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathbb{E}\rho(V^{n,\delta}(t), V^n(t))^2 1_{t<\eta_m^{n,\delta}} \\
&\leq 4m^2\Big(\mathbb{E}\frac{1}{n}\sum_{i=1}^n |X_i^{n,\delta}(t) - X_i^n(t)|^2 1_{t<\eta_m^{n,\delta}} \\
&\qquad + \mathbb{E}\frac{1}{n}\sum_{i=1}^n |Z_i^{n,\delta}(t) - Z_i^n(t)|^2 1_{t<\eta_m^{n,\delta}}\Big).
\end{aligned}$$

Define

$$\begin{aligned}
a^\delta(t) &\equiv \mathbb{E}\sup_{s\leq t\wedge\eta_m^{n,\delta}} |X_i^{n,\delta}(s) - X_i^n(s)|^2 \\
b^\delta(t) &\equiv \mathbb{E}\sup_{s\leq t\wedge\eta_m^{n,\delta}} |Z_i^{n,\delta}(s) - Z_i^n(s)|^2.
\end{aligned}$$

Then by (3.5), the Lipshitz conditions on the coefficients, and the estimates above,

$$a^\delta(t) \leq c_{12}(m)\delta t + c_{13}(m)\delta t^2 + (c_{14}(m) + c_{15}(m)t)\int_0^t (a^\delta(s) + b^\delta(s))ds \ . \tag{3.7}$$

A similar inequality holds for $b^\delta(t)$, and (3.4) follows by Gronwall's inequality. □

The proof of the following corollary is similar to that of Corollary 2.4.

Corollary 3.4 *Assume (I) and (S1)-(S3). For each bounded, Lipschitz function f and $T > 0$,*

$$\mathbb{E}\sup_{t\leq T\wedge\eta_m^{n,\delta}} |V^{n,\delta}(t)f - V^n(t)f| \leq c_{16}(T,m)\sqrt{\delta}\|f\|_L,$$

and hence,

$$\mathbb{E}\sup_{t\leq T\wedge\eta_m^{n,\delta}} \tilde{\rho}(V^{n,\delta}(t), V^n(t)) \leq c_{16}(T,m)\sqrt{\delta}. \tag{3.8}$$

3.1. Application to filtering equations

Let X_i^δ, A_i^δ and $\mu^{n,\delta}$ be the Euler scheme for the system (1.20-1.21) given by formulae similar to (1.12-1.14). Let ξ_δ be as in Section 1. We need the following additional assumption:

(F3) There exists a constant K such that for each $x \in \mathbb{R}^d$,

$$\int_U |h(x,u) - h(x,\xi_\delta(u))|^2 \mu(du) + \int_U |\alpha(x,\nu,u) - \alpha(x,\nu,\xi_\delta(u))|^2 \mu(du) \leq K^2 \delta.$$

Corollary 3.5 *Assume (I′), (F1)-(F3). For each $T > 0$,*

$$\left(\mathbb{E}\sup_{t\leq T}|X_i^\delta(t) - X_i(t)|^2\right)^{1/2} + \left(\mathbb{E}\sup_{t\leq T}|Z_i^\delta(t) - Z_i(t)|^2\right)^{1/2} \leq c_{17}(T)\sqrt{\delta}. \tag{3.9}$$

Proof. Since the coefficients do not depend on the empirical measure, $X_i^{n,\delta}(t) = X_i^\delta(t)$ and $X_i^n(t) = X_i(t)$ in the filtering case. It also follows we can take $m = \infty$ in the definitions of a_δ and b_δ and in (3.7), and we obtain (3.9). □

Corollary 3.6 *Assume (I′), (F1)-(F3). Under both the model measure P and the reference measure Q, for each bounded, Lipschitz function f and each $T > 0$,*

$$\mathbb{E}\sup_{t\leq T}|\mu_t^{n,\delta} f - \mu_t^n f| \leq c_{18}(T)\sqrt{\delta}\|f\|_L, \tag{3.10}$$

and hence,

$$\mathbb{E}\sup_{t\leq T}\tilde{\rho}(\mu_t^{n,\delta}, \mu_t^n) \leq c_{18}(T)\sqrt{\delta}. \tag{3.11}$$

Proof. Under Q,

$$\begin{aligned}
&\mathbb{E}^Q \sup_{t\leq T}|\mu_t^{n,\delta} f - \mu_t^n f| \\
&\quad\leq \mathbb{E}^Q \sup_{t\leq T}\frac{1}{n}\sum_{i=1}^n \Big[A_i^\delta(t) \vee A_i(t)| \Big(|f(X_i^\delta(t))||Z_i^\delta(t) - Z_i(t)| \\
&\qquad\qquad + |f(X_i^\delta(t)) - f(X_i(t))|\Big)\Big] \\
&\quad\leq \|f\|_L (\mathbb{E}^Q(\sup_{t\leq T} A_i(t)^2 + \sup_{t\leq T} A_i^\delta(t)^2))^{\frac{1}{2}} \\
&\qquad [(\mathbb{E}^Q \sup_{t\leq T}|X_i^\delta(t) - X_i(t)|^2)^{\frac{1}{2}} + (\mathbb{E}^Q \sup_{t\leq T}|Z_i^\delta(t) - Z_i(t)|^2)^{\frac{1}{2}}] \\
&\quad\leq 2\|f\|_L \exp\left(K^2 T\right) c_{17}(T)\sqrt{\delta},
\end{aligned}$$

and (3.10) follows. The analogous result for P follows as in (2.17).

(3.11) is a direct consequence of (3.10) and the definition of $\tilde{\rho}$. $\square$

4. Overall error estimate

Finally, we combine the estimates of the sampling error and the discretization error to obtain the following:

Theorem 4.1 *a) Let $\bar{V}^n(t) = V^{n,1/n}(t)$ and $\bar{\eta}^n_m = \eta^n_m \wedge \eta^{n,1/n}_m$. Assume (I) and (S1)-(S3). For each bounded, Lipschitz function f and each $t \geq 0$,*

$$\mathbb{E}|\bar{V}^n(t)f - V(t)f|1_{t\leq\bar{\eta}^n_m} \leq \frac{c_{19}(t,m)}{\sqrt{n}}$$

and hence

$$\mathbb{E}\tilde{\rho}(\bar{V}^n(t), V(t))1_{t\leq\bar{\eta}^n_m} \leq \frac{c_{19}(t,m)}{\sqrt{n}}. \tag{4.1}$$

As a consequence, for each fixed t, the sequence $\{\sqrt{n}\tilde{\rho}(\bar{V}^n(t), V(t))\}_{n\geq 1}$ is stochastically bounded, i.e., for each $\epsilon > 0$, there exists $M > 0$, such that for all n,

$$\mathbb{P}\left(\sqrt{n}\tilde{\rho}(\bar{V}^n(t), V(t)) > M\right) < \epsilon. \tag{4.2}$$

b) For the filtering problem, let $\bar{\mu}^n_t = \mu^{n,1/n}_t$. Assume (I'), (F1) - (F3). Under both the model measure P and the reference measure Q, for each bounded, Lipschitz function f and each $t \geq 0$,

$$\mathbb{E}|\bar{\mu}^n_t f - \mu_t f| \leq \frac{c_{20}(t)}{\sqrt{n}},$$

and hence

$$\mathbb{E}\tilde{\rho}(\bar{\mu}^n_t, \mu_t) \leq \frac{c_{20}(t)}{\sqrt{n}}. \tag{4.3}$$

Remark 4.2 *Note that the estimates here are for fixed t, while most of the intermediate estimates involved a supremum* inside *the expectation. The only point in the development where we have not been able to make the estimates with the supremum inside the expectation is in (2.13).*

Proof. (4.1) follows from (2.14) and (3.8) with $\delta = \frac{1}{n}$. (4.3) follows from (2.16) and (3.11) with $\delta = \frac{1}{n}$.

To obtain the stochastic boundedness, observe that

$$\begin{aligned}
&\mathbb{P}\left(\sqrt{n}\tilde{\rho}(\bar{V}^n(t), V(t)) > M\right)\\
&\leq \mathbb{P}\left(\sqrt{n}\tilde{\rho}(V^{n,\frac{1}{n}}(t), V^n(t)) > \frac{M}{2}\right) + \mathbb{P}\left(\sqrt{n}\tilde{\rho}(V^n(t), V(t)) > \frac{M}{2}\right)\\
&\leq \mathbb{P}\left(\sqrt{n}\tilde{\rho}(V^{n,\frac{1}{n}}(t\wedge\eta_m^{n,\frac{1}{n}}), V^n(t\wedge\eta_m^{n,\frac{1}{n}})) > \frac{M}{2}\right) + \mathbb{P}\left(\eta_m^{n,\frac{1}{n}} < t\right)\\
&\quad +\mathbb{P}\left(\sqrt{n}\tilde{\rho}(V^n(t\wedge\eta_m^n), V(t\wedge\eta_m^n)) > \frac{M}{2}\right) + \mathbb{P}\left(\eta_m^n < t\right)\\
&\leq \frac{2}{M}\sup_{1\leq n<\infty}\sqrt{n}\mathbb{E}\tilde{\rho}(V^{n,\frac{1}{n}}(t\wedge\eta_m^{n,\frac{1}{n}}), V^n(t\wedge\eta_m^{n,\frac{1}{n}}))\\
&\quad +\frac{1}{m^2}\sup_{1\leq n<\infty}\mathbb{E}\sup_{0\leq s\leq t}|A_i^{n,\frac{1}{n}}(s)|^2\\
&\quad +\frac{2}{M}\sup_{1\leq n<\infty}\sqrt{n}\mathbb{E}\tilde{\rho}(V^n(t\wedge\eta_m^n), V(t\wedge\eta_m^n))\\
&\quad +\frac{1}{m^2}\sup_{1\leq n\leq\infty}\mathbb{E}\sup_{0\leq s\leq t}|A_i^n(s)|^2
\end{aligned}$$

Since for each fixed m, each of the suprema over n is finite, the right side can be made less than $\epsilon > 0$ by first making m large enough so that the second and fourth terms are each less than $\epsilon/4$ and then making M large enough so that the first and third terms are each less than $\epsilon/4$. □

References

[1] Bhatt, A. G.; Kallianpur, G. and Karandikar, R. L. (1995). Uniqueness and robustness of solution of measure-valued equations of nonlinear filtering, *Ann. Probab.* 23, 1895-1938.

[2] Budhiraja, Amarjit and Kallianpur, G. (1996). Approximations to the solution of the Zakai equation using multiple Wiener and Stratonovitch integral expansions. *Stochastics Stochastics Rep.* 56, 271-315.

[3] Crisan, Dan; Del Moral, Pierre and Lyons, Terry (1999). Discrete filtering using branching and interacting particle systems. *Markov Process. Related Fields* 5, 293-318.

[4] Crisan, Dan; Del Moral, Pieree and Lyons, Terry (1999). Interacting particle approximations of the Kushner Stratonovich equation. *Adv. Appl. Probab.* (to appear)

[5] Crisan, Dan; Gaines, Jessica; and Lyons, Terry (1998). Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Math.* 58, 1568-1590.

[6] Crisan, Dan and Lyons, Terry (1999), A particle approximation to the solution of the Kushner-Stratonovitch equation. *Probab. Theory and Related Fields* (to appear)

[7] Del Moral, Pierre (1995). Non-linear filtering using random particles. *Theory Probab. Appl.* **40**, 690-701.

[8] Di Masi, G. B.; Pratelli, M.; and Runggaldier, W. J. (1985). An approximation for the nonlinear filtering problem with error bound. *Stochastics.* 14, 247-271.

[9] Florchinger. Patrick and Le Gland, François (1991). Time-discretization of the Zakai equation for diffusion processes observed in correlated noise. *Stochastics Stochastics Rep.* 35, 233-256.

[10] Florchinger, Patrick and Le Gland, François (1990). Time-discretization of the Zakai equation for diffusion processes observed in correlated noise. In: *Analysis and optimization of systems (Antibes, 1990)*, 228-237, *Lecture Notes in Control and Inform. Sci.,* 144, Springer, Berlin.

[11] Florchinger, Patrick and Le Gland, François (1992). Particle approximations for first order stochastic partial differential equations. *Applied stochastic analysis (New Brunswick, NJ*, 1991), 121-133, *Lecture Notes in Control and Inform. Sci.*, **177** Springer, Berlin.

[12] Fujisaki, Masatoshi; Kallianpur, G.; Kunita, Hiroshi (1972). Stochastic differential equations for the non linear filtering problem. *Osaka J. Math.* 9, 19-40.

[13] Hu, Y.; Kallianpur, G.; and Xiong, Jie An approximation for Zakai equation. *Applied Mathematics and Optimization.* (to appear)

[14] Kallianpur, G. (1980). *Stochastic Filtering Theory.* Springer-Verlag, Berlin.

[15] Kloeden, Peter E. and Platen, Eckhard (1992) *Numerical solution of stochastic differential equations. Applications of Mathematics, 23.* Springer-Verlag, Berlin.

[16] Kurtz, Thomas G. and Ocone, Daniel L. (1988). Unique characterization of conditional distributions in nonlinear filtering. *Ann. Probab.* 16, 80-107.

[17] Kurtz, Thomas G. and Protter, Philip Weak error estimates for simulation schemes for SDEs. (preprint)

[18] Kurtz, Thomas G. and Xiong, Jie (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.* 83, 103-126.

[19] Kushner, Harold J. (1977). *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations.* Academic Press, New York.

[20] Kushner, Harold J. (1979). A robust discrete state approximation to the optimal nonlinear filter for a diffusion. *Stochastics Stochastics Rep.* 3, 75-83.

[21] Kushner, Harold J. (1997). Robustness and convergence of approximations to nonlinear filters for jump-diffusions. *Comput. Appl. Math., V.* 16, 153-183.

[22] Liptser, R. Sh. and Shiryayev, A. N. (1977). *Statistics of Random Processes I.* Springer-Verlag, Berlin.

[23] Lototsky, Sergey and Rozovskii, Boris L. (1997). Recursive multiple Wiener integral expansion for nonlinear filtering of diffusion processes. In: *Stochastic processes and functional analysis (Riverside, CA, 1994)*, 199-208, *Lecture Notes in Pure and Appl. Math.,* 186, Dekker, New York.

[24] Lototsky, Sergey; Mikulevicius, R. and Rozovskii, Boris L. (1997). Nonlinear filtering revisited: a spectral approach. *SIAM J. Control Optim.* 35, 435-461.

[25] Milstein, G. N. (199*5). Numerical integration of stochastic differential equations. Mathematics and its Applications, 313.* Kluwer Academic Publishers Group, Dordrecht.

[26] Picard, J. (1984). Approximation of nonlinear filtering problems and order of convergence. *Filtering and Control of Random Processes. Lecture Notes Control Inf. Sci.* 61, Springer, New York.

[27] Rozovskii, Boris L. (1991). A simple proof of uniqueness for Kushner and Zakai equations. *Stochastic analysis*, 449-458, Academic Press, Boston, MA.

[28] Szpirglas, J. (1978). Sur l'équivalence d'équations différentielles stochastiques à valeurs mesures intervenant dans le filtrage markovien non linéaire. *Ann. Inst. H. Poincaré Sect. B (N.S.)* 14, 33-59.

Thomas G. Kurtz
Departments of Mathematics and Statistics
University of Wisconsin - Madison
480 Lincoln Drive
Madison, WI 53706-1388
kurtz@math.wisc.edu

Jie Xiong
Department of Mathematics
University of Tennessee
Knoxville, TN 37996-1300
jxiong@math.utk.edu