Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT
Stochastic Processes and their Applications 110 (2004) 275–294
stochastic processes and their applications
www.elsevier.com/locate/spa

Markov chain approximations to filtering equations for reflecting diffusion processes

Michael A. Kouritzin, Hongwei Long, Wei Sun*

Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada T6G 2G1

Received 29 September 2001; received in revised form 28 October 2003; accepted 28 October 2003

Abstract

Herein, we consider direct Markov chain approximations to the Duncan–Mortensen–Zakai equations for nonlinear filtering problems on regular, bounded domains. For clarity of presentation, we restrict our attention to reflecting diffusion signals with symmetrizable generators. Our Markov chains are constructed by employing a wide band observation noise approximation, dividing the signal state space into cells, and utilizing an empirical measure process estimation. The upshot of our approximation is an efficient, effective algorithm for implementing such filtering problems. We prove that our approximations converge to the desired conditional distribution of the signal given the observation. Moreover, we use simulations to compare computational efficiency of this new method to the previously developed branching particle filter and interacting particle filter methods. This Markov chain method is demonstrated to outperform the two-particle filter methods on our simulated test problem, which is motivated by the fish farming industry.

MSC: 60G35; 60H15; 60J27

Keywords: Nonlinear filtering; Reflecting diffusion; Duncan–Mortensen–Zakai equation; Markov chain approximation; Law of large numbers; Particle filters

1. Introduction

The requirement of finding workable approximate solutions to the filtering distributions, which are not resolved by exact methods like the Kalman filter, is key to many engineering disciplines. In this regard, many authors e.g. Kushner (1977, 1979) and Di Masi and Runggaldier (1981, 1982), have utilized Markov chain approximations

* Corresponding author.
E-mail address: wsun@math.ualberta.ca (W. Sun).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2003.10.013

of the signal process which, when combined with Clark's robust filter (1978), reduce these approximation difficulties to solving a system of ordinary differential equations (ODEs) parameterized by the observation process or, when combined with the Kallianpur–Striebel formula, provide a method of calculating a discretized approximation to the conditional distribution of the original signal given the observations. Recently, there have been many works devoted to applying various particle methods to construct approximate solutions to the celebrated Duncan–Mortensen–Zakai equation. Among them, we would like to mention Del Moral and collaborators' adaptive interacting particle (AIP) filter (see Del Moral (1996) for one of the earlier works and Del Moral and Miclo (2000) for an excellent and complete account) and the improved refining interacting particle filter (see Del Moral et al., 2001). Similarly, Crisan, Lyons and collaborators introduced the adaptive branching particle filter (see Crisan and Lyons, 1997; Crisan et al., 1998, 1999), that was improved by Kouritzin and collaborators' refining branching particle (RBP) filter (see Ballantyne et al., 2000). In this work, we take the new method of approximating the Duncan–Mortensen–Zakai equation directly. Our method utilizes the so-called stochastic particle Markov chain approximation introduced in the context of ODEs by Kurtz (1971), of partial differential equations (PDEs) by Arnold and Theodosopulu (1980), and of stochastic partial differential equations (SPDEs) by Kouritzin and Long (2002). Blount (1991, 1994) and Kotelenez (1986, 1988) have also made fundamental contributions to the analysis of such stochastic particle approximations.

We consider the low observable filtering problem of detecting and tracking a target buried in high-amplitude synthetic observation noise. Motivated by fish farming applications, we constrain our target to live within the closure $\bar{D}$ of a d-dimensional rectangular region $D=(0,L_1)\times(0,L_2)\times\cdots\times(0,L_d)$, undergoing reflections at the boundary ∂D of this region. Without loss of generality, we assume that L_i is a positive integer for each $1\leqslant i\leqslant d$. We suppose that $\{a_{ij}\}_{i,j=1}^d$, $\rho:\bar{D}\to\mathbb{R}$ are functions satisfying the following conditions:

(i) $a_{ij}(\cdot)=a_{ji}(\cdot)\in C^3(\bar{D})$, the space of three times continuously differentiable functions on $\bar{D}$, for all $1\leqslant i,j\leqslant d$. Moreover,

$$\lambda^{-1}|\xi|^2\leqslant\sum_{i,j=1}^d a_{ij}(x)\xi_i\xi_j\leqslant\lambda|\xi|^2,\quad \forall\xi\in\mathbb{R}^d \text{ and } x\in\bar{D} \tag{1.1}$$

for some $\lambda>0$ not depending on x.

(ii) $\rho(\cdot)\in C^3(\bar{D})$ such that $\inf_{x\in\bar{D}}\rho(x)>0$.

We denote by $H^{i,2}(D)$ the $(i,2)$-Sobolev space on D for $i=1,2$ and consider on $H:=L^2(\bar{D};\rho^2\,\mathrm{d}x)$ the symmetric bilinear form

$$\begin{cases}\mathscr{E}(u,v)=\dfrac{1}{2}\displaystyle\int_{\bar{D}}\langle a(x)\nabla u(x),\nabla v(x)\rangle\rho^2(x)\,\mathrm{d}x, & u,v\in\mathscr{D}(\mathscr{E}),\\ \mathscr{D}(\mathscr{E})=H^{1,2}(D).\end{cases}$$

One can check that $\mathscr{E}$ is a regular Dirichlet form satisfying the local property so it is associated with a strong Markov diffusion process $(C_{\bar{D}}[0,\infty),(x_t)_{t\geqslant 0},(P_x)_{x\in\bar{D}})$, where $(x_t)_{t\geqslant 0}$ is the coordinate process (cf. Fukushima et al., 1994, Theorems 7.2.1 and 7.2.2). In particular, if $(p_t)_{t>0}$ and $(T_t)_{t>0}$ denote the semigroups associated with $(x_t)_{t\geqslant 0}$ and $\mathscr{E}$, respectively, then $p_t f = T_t f$ dx-a.e. for any $f\in L^\infty(\bar{D};\rho^2\,\mathrm{d}x)$ and $t>0$. Let $U(x)=(U_1(x),\ldots,U_d(x))$ be the unit inward normal at $x\in\partial D$. We define the conormal vector field γ by $\gamma_i(x) := \sum_{j=1}^d a_{ij}(x)U_j(x),\ x\in\partial D$. We denote $\partial_i=\partial/\partial x_i$ for $1\leqslant i\leqslant d$. Then, the generator associated with $\mathscr{E}$ is the self-adjoint operator

$$\begin{cases} \mathscr{L}f = \dfrac{1}{2\rho^2}\displaystyle\sum_{j=1}^d\left(\partial_j\sum_{i=1}^d \rho^2 a_{ij}\partial_i f\right), \quad f\in\mathscr{D}(\mathscr{L}),\\ \mathscr{D}(\mathscr{L}) = \{f\in H^{2,2}(D) : \langle\gamma,\nabla f\rangle|_{\partial D}=0\}. \end{cases}$$

The Markov family $((x_t)_{t\geqslant 0},(P_x)_{x\in\bar{D}})$ is thus governed by $\mathscr{L}$ inside the domain D with reflections at the boundary in the direction $\gamma(x)$. In fact, the actual "signal" $((x_t)_{t\geqslant 0},(P_x)_{x\in\bar{D}})$ satisfies the Skorohod stochastic differential equation (see Freidlin, 1985, Section 1.6)

$$\begin{cases} \mathrm{d}x_t = \sigma(x_t)\mathrm{d}v_t + b(x_t)\mathrm{d}t + \chi_{\partial D}(x_t)\gamma(x_t)\mathrm{d}\eta_t,\\ x_0 = x, \quad \eta_0 = 0, \end{cases}$$

where $\sigma\in C^3(\bar{D})$ satisfies $\sigma(x)\sigma^*(x)=(a_{ij}(x))$, $b := a\nabla\ln\rho + (\frac{1}{2}\sum_{j=1}^d \partial_j a_{ij})$, χ is the indicator function, v_t is a standard d-dimensional Brownian motion, and η_t is the local time of x_t, which is an increasing continuous additive functional that increases only when $x_t\in\partial D$.

We let $(C_{\mathbb{R}}[0,\infty),(w_t)_{t\geqslant 0},P)$ be a standard Brownian motion, where $(w_t)_{t\geqslant 0}$ denotes the coordinate process. For $h\in C^1(\bar{D})$, we define

$$y_t := \int_0^t h(x_s)\,\mathrm{d}s + w_t, \quad t\geqslant 0. \tag{1.2}$$

The real valued process $(y_t)_{t\geqslant 0}$ is a "noisy observation" of the signal $(x_t)_{t\geqslant 0}$. Throughout this paper, we let $T>0$ be arbitrary and $\Omega := C_{\bar{D}}[0,T]\times C_{\mathbb{R}}[0,T]$. We use $\mathbf{P}$ to denote the measure on $(\Omega,\mathscr{B}(\Omega))$ that gives the joint distribution of the independent processes $(x_t)_{0\leqslant t\leqslant T}$ and $(w_t)_{0\leqslant t\leqslant T}$ with the initial condition $x_0=$ given random variable. For $t\in[0,T]$, we let $\mathscr{N}$ be the collection of all $\mathbf{P}$ null sets, $\mathscr{Y}_t$ be the σ-algebra $\sigma\{y_s, 0\leqslant s\leqslant t\}\vee\mathscr{N}$, and $\mathbf{E}$ be the expectation with respect to $\mathbf{P}$. For a real measurable function f on $\bar{D}$ satisfying $\mathbf{E}|f(x_t)|^2<\infty$ for all $0\leqslant t\leqslant T$, the filtering problem is to evaluate

$$\pi_t(f) := \mathbf{E}[f(x_t)\,|\,\mathscr{Y}_t], \quad t\in[0,T], \tag{1.3}$$

which is the least-square estimate of $f(x_t)$ given all the observations up to time t. For each $t\in[0,T]$, we call any version of $\pi_t(f)$ in (1.3) an optimal filter.

For $t\in[0,T]$, we define $\Lambda_t := \exp(\int_0^t h(x_s)\,\mathrm{d}y_s - \frac{1}{2}\int_0^t h^2(x_s)\,\mathrm{d}s)$. It is well known that the formula $\mathrm{d}\mathbf{P}^0/\mathrm{d}\mathbf{P} := \Lambda_T^{-1}$ defines a probability measure $\mathbf{P}^0$ under which

(i) the distribution of $\{x_t\}$ is the same as under $\mathbf{P}$,
(ii) $\{y_t, t\in[0,T]\}$ is a standard Brownian motion,
(iii) $\{x_t\}$ and $\{y_t\}$ are independent.

Let $\mathbf{E}^0$ denote the expectation with respect to $\mathbf{P}^0$. Then, one has the Kallianpur–Striebel formula

$$\pi_t(f) = \frac{\mathbf{E}^0[f(x_t)\Lambda_t|\mathscr{Y}_t]}{\mathbf{E}^0[\Lambda_t|\mathscr{Y}_t]} := \frac{\sigma_t(f)}{\sigma_t(1)}.$$

For any $f\in C^\infty(\bar{D})$ satisfying the boundary condition $\langle\gamma,\nabla f\rangle|_{\partial D}=0$, one has the weak form of the Duncan–Mortensen–Zakai equation (cf. Zakai, 1969)

$$\sigma_t(f) = \sigma_0(f) + \int_0^t \sigma_s(\mathscr{L}f)\,\mathrm{d}s + \int_0^t \sigma_s(hf)\,\mathrm{d}y_s \quad \text{a.s. } \mathbf{P}^0. \tag{1.4}$$

Under our ellipticity and smoothness condition, following the elegant arguments of Pardoux (1982), one finds that σ_t has a density $p(t,x)$ on H, which is the pathwise unique weak solution to the SPDE on H

$$\begin{cases} \mathrm{d}p(t,x) = \mathscr{L}p(t,x)\,\mathrm{d}t + h(x)p(t,x)\,\mathrm{d}y_t, & t>0,\ x\in D,\\ p(0,x) = p_0(x), & x\in\bar{D},\\ \langle\gamma(x),\nabla p(t,x)\rangle|_{x\in\partial D} = 0, & t>0, \end{cases}$$

where the density function p_0 of the distribution of x_0 is assumed to be in H.

We define $(\hat{H},\|\cdot\|) := L^2(\bar{D};\mathrm{d}x)$. Using the unitary map I from $\hat{H}$ to $H: I\circ f = f/\rho$, $\forall f\in\hat{H}$, we get the image Dirichlet form on $\hat{H}$

$$\begin{cases} \hat{\mathscr{E}}(u,v) = \mathscr{E}(I\circ u, I\circ v), & u,v\in\mathscr{D}(\hat{\mathscr{E}}),\\ \mathscr{D}(\hat{\mathscr{E}}) = \{u: I\circ u\in H^{1,2}(D)\}. \end{cases}$$

Let $(\hat{T}_t)_{t>0}$ and $\hat{\mathscr{L}}$ denote the semigroup and generator associated with $(\hat{\mathscr{E}},\mathscr{D}(\hat{\mathscr{E}}))$ respectively, then one can check that

$$\hat{T}_t f = I^{-1}\circ T_t(I\circ f) \quad \mathrm{d}x\text{-a.e.}, \quad \forall f\in\hat{H},\ t>0$$

and

$$\begin{cases} \hat{\mathscr{L}}f = I^{-1}\circ\mathscr{L}(I\circ f)\\ \quad = \dfrac{1}{2}\displaystyle\sum_{j=1}^d\left(\partial_j\sum_{i=1}^d a_{ij}\partial_i f\right) - \frac{f}{2\rho}\sum_{j=1}^d\left(\partial_j\sum_{i=1}^d a_{ij}\partial_i\rho\right), & f\in\mathscr{D}(\hat{\mathscr{L}}),\\ \mathscr{D}(\hat{\mathscr{L}}) = \{f: I\circ f\in\mathscr{D}(\mathscr{L})\}. \end{cases}$$

We define $\hat{p}(t,x) := \rho(x)p(t,x)$. Then, $\hat{p}(t,x)$ is the pathwise unique weak solution to the SPDE on $\hat{H}$

$$\begin{cases} \mathrm{d}\hat{p}(t,x) = \hat{\mathscr{L}}\hat{p}(t,x)\,\mathrm{d}t + h(x)\hat{p}(t,x)\,\mathrm{d}y_t, & t>0,\ x\in D,\\ \hat{p}(0,x) = \hat{p}_0(x) := \rho(x)p_0(x), & x\in\bar{D},\\ \left\langle \gamma(x), \nabla\dfrac{\hat{p}(t,x)}{\rho(x)}\right\rangle|_{x\in\partial D} = 0, & t>0. \end{cases}$$

Motivated by Kushner and Huang (1985), we let $z(t)$ be a stationary, zero mean, bounded and right continuous ϕ-mixing process with $\phi(\cdot)$ satisfying $\int_0^\infty \phi^{1/2}(t)\,\mathrm{d}t < \infty$. Also, we let $\int_{-\infty}^{\infty} \mathbf{E}z(s)z(0)\,\mathrm{d}s = 1$. For $n\in\mathbb{N}$, we define an approximate observation process y^n by

$$y_t^n := \int_0^t (h(x_s) + \sqrt{n}z_{ns})\,\mathrm{d}s, \quad t \geqslant 0, \tag{1.5}$$

$\dot{y}_t^n := (\mathrm{d}/\mathrm{d}t)y_t^n$, and $\hat{p}^n$ by the pathwise unique weak solution to the random PDE on $\hat{H}$

$$\begin{cases} \dfrac{\partial\hat{p}^n(t,x)}{\partial t} = \hat{\mathscr{L}}\hat{p}^n(t,x) - \frac{1}{2}h^2(x)\hat{p}^n(t,x) + h(x)\hat{p}^n(t,x)\dot{y}_t^n, & t>0,\ x\in D,\\ \hat{p}^n(0,x) = \hat{p}_0(x), & x\in\bar{D},\\ \langle\gamma(x), \nabla\dfrac{\hat{p}^n(t,x)}{\rho(x)}\rangle|_{x\in\partial D} = 0, & t>0. \end{cases} \tag{1.6}$$

Then, similar to Kushner and Huang (1985, Theorem 2), one can show that

$$\sup_{n,t\leqslant T} \mathbf{E}\|\hat{p}^n(t)\|^2 < \infty. \tag{1.7}$$

Let $\hat{H}_{\mathrm{w}}$ denote $\hat{H}$ endowed with the weak topology. For any $M\in\mathbb{N}$, we define $S_M := \{f\in\hat{H}\colon \|f\| < M\}$, let d_M be the usual metric for the weak topology of S_M, and set $d = \sum_{M=1}^\infty 2^{-M}d_M$. Owing to (1.7), convergence for $\hat{p}^n$ in the metric d is actually convergence in the weak topology. Let $C_{\hat{H}_{\mathrm{w}}}[0,T]$ denote the $\hat{H}_{\mathrm{w}}$-valued continuous functions on $[0,T]$. One can prove along the lines of Kushner and Huang (1985, Theorem 6) that $\{\hat{p}^n\}$ converges to $\hat{p}$ in distribution in $C_{\hat{H}_{\mathrm{w}}}[0,T]$. We will give a computer workable approximation to $\hat{p}^n$ and prove convergence for the approximation.

We denote $v := (1/2\rho)\sum_{j=1}^d \left(\partial_j \sum_{i=1}^d a_{ij}\partial_i\rho\right)$, $\alpha := \sup_{x\in\bar{D}}|v(x)|$, and define a semigroup $(\tilde{T}_t)_{t>0}$ on $\hat{H}$ by $\tilde{T}_t := \mathrm{e}^{-\alpha t}\hat{T}_t$ for all $t>0$. Then, the generator associated with $(\tilde{T}_t)_{t>0}$ is $\tilde{\mathscr{L}} := \hat{\mathscr{L}} - \alpha$. $\tilde{\mathscr{L}}$ is the evolution generator on the unweighted space $\hat{H}$ where we perform our analysis. Here we choose $\tilde{\mathscr{L}}$, rather than $\mathscr{L}$ or $\hat{\mathscr{L}}$, to be the evolution generator since it is easier to construct explicit Markov chain approximations for $\tilde{\mathscr{L}}$ (cf. (2.2) below) and employ the useful symmetric closed-form technique to prove the convergence for the approximations. Using variation of constants, the solution of (1.6) can be put in the form

$$\hat{p}^n(t) = \tilde{T}(t)\hat{p}_0^n + \int_0^t \tilde{T}(t-s)\left(\left(h\dot{y}^n(s) + \alpha - \frac{1}{2}h^2\right)\hat{p}^n(s)\right) \mathrm{d}s. \tag{1.8}$$

In the sequel, we discuss Markov chain approximations to the integral equation (1.8) for each fixed $n \in \mathbb{N}$ and thereby initiate analysis of a novel Markov chain filter. To ease notation, we omit the superscript n, understanding that all results are for $\hat{p}^n$ and y^n.

In Section 2, we discuss the construction of Markov chain approximations to Eq. (1.8). Then, in Section 3, we state and prove the quenched and annealed laws of large numbers which establish the convergence of our Markov chains to the solution of (1.8) in mean squares sense. Note that although we consider the continuous observation model (1.2) and employ the wide band observation noise approximation (1.5) in Sections 2 and 3, similar results continue to hold if we employ some other observation approximations, e.g. a polygonal approximation (cf. Hu et al., 2002), to the Duncan–Mortensen–Zakai equation (1.4). Finally, in Section 4, we compare computational efficiency of this Markov chain method with the previously mentioned AIP filter and RBP filter methods for the fish tracking problem. We consider a discrete observation model (cf. (4.1) below) and use simulation results to show that this Markov chain method outperforms the two-particle filter methods on our test problem.

2. Construction of Markov chain

The Markov chain approximation discussed in this paper is motivated by the stochastic particle models of chemical reaction with diffusion studied by Arnold and Theodosopulu (1980), Kotelenez (1986, 1988), Blount (1991, 1994), and Kouritzin and Long (2002). In their models, the operator $\tilde{\mathscr{L}}$ is replaced by a less general operator like the Laplacian. Blount (1991, 1994), and Kouritzin and Long (2002), proved that a sequence of Markov chain approximations converges to the solution of their models weakly (in the distribution convergence sense uniformly in time). In our model, we employ the symmetric closed form technique to get the convergence in mean square of Markov chain approximations for our more general class of operators.

Before defining the stochastic particle models, we prepare some preliminaries concerning the discretization of the operator $\tilde{\mathscr{L}}$. Now, for $N \in \mathbb{N}$, we let $D_N := \{k = (k_1, \ldots, k_d) \in \mathbb{N}^d\colon\ 1 \leqslant k_i \leqslant L_i N \text{ for each } 1 \leqslant i \leqslant d\}$ and divide $[0, L_1) \times \cdots \times [0, L_d)$ into $L_1 N \times \cdots \times L_d N$ cells of size $1/N^d$:

$$I_k^N := \left[\frac{k_1 - 1}{N}, \frac{k_1}{N}\right) \times \cdots \times \left[\frac{k_d - 1}{N}, \frac{k_d}{N}\right), \quad k \in D_N.$$

We define $\hat{H}^N := \mathbb{R}^{L_1 N \times \cdots \times L_d N}$ and endow $\hat{H}^N$ with the inner product

$$\langle \varphi, \psi \rangle_N := \frac{1}{N^d} \sum_{k \in D_N} \varphi_k \psi_k, \quad \forall \varphi = (\varphi_k)_{k \in D_N}, \quad \psi = (\psi_k)_{k \in D_N} \in \hat{H}^N.$$

Then, $(\hat{H}^N, \langle\cdot,\cdot\rangle_N)$ is a Hilbert space with norm denoted by $\|\cdot\|_N$. For $k \in D_N$, we define $a_{ij}^N(k) := N^d \int_{I_k^N} a_{ij}(x)\,\mathrm{d}x$ for $1 \leqslant i,j \leqslant d$, $v_k^N := N^d \int_{I_k^N} v(x)\,\mathrm{d}x$ and $h_k^N := N^d \int_{I_k^N} h(x)\,\mathrm{d}x$. For $1 \leqslant i \leqslant d$ we denote $e_i = (0,\dots,0,1,0,\dots,0)$ with 1 in the ith coordinate, and set $S_i\varphi_k := \varphi_{k+e_i}$ if $k,\ k+e_i \in D_N$, and $S_{-i}\varphi_k := \varphi_{k-e_i}$ if $k,\ k-e_i \in D_N$.

We consider on $\hat{H}^N$ the symmetric closed form $(\tilde{\mathscr{E}}^N, \mathscr{D}(\tilde{\mathscr{E}}^N))$

$$\begin{cases} \tilde{\mathscr{E}}^N(u,v) = \dfrac{1}{N^d} \displaystyle\sum_{k \in D_N^0} \left[\frac{N^2}{2} \sum_{i,j=1}^d a_{ij}^N(k)(S_i - I)u_k \cdot (S_j - I)v_k + (v_k^N + \alpha)u_k v_k \right], \\ \qquad\qquad u,v \in \mathscr{D}(\tilde{\mathscr{E}}^N), \\ \mathscr{D}(\tilde{\mathscr{E}}^N) = \hat{H}^N, \end{cases}$$

where $D_N^0 := \{k \in D_N : 2 \leqslant k_i \leqslant L_i N - 1 \text{ for each } 1 \leqslant i \leqslant d\}$. Define

$$\begin{cases} \tilde{\mathscr{E}}(u,v) := \hat{\mathscr{E}}(u,v) + \alpha(u,v), \quad u,v \in \mathscr{D}(\tilde{\mathscr{E}}), \\ \mathscr{D}(\tilde{\mathscr{E}}) = \mathscr{D}(\hat{\mathscr{E}}). \end{cases}$$

Then, $\tilde{\mathscr{E}}^N$ can be thought as a discretized version of $\tilde{\mathscr{E}}$. For $\varphi = (\varphi_k)_{k \in D_N} \in \hat{H}^N$, we define

$$(\tilde{\mathscr{L}}^N \varphi)_k := \frac{N^2}{2} \sum_{i,j=1}^d [a_{ij}^N(k)(S_i - I)\varphi_k 1_{k \in D_N^0} - (S_{-j}(a_{ij}^N(k)(S_i - I)\varphi_k))1_{k \in D_N^0 + e_j}]$$
$$-(v_k^N + \alpha)\varphi_k 1_{k \in D_N^0}.$$

Then, $\tilde{\mathscr{L}}^N$ is a symmetric bounded linear operator on $\hat{H}^N$ associated with $\tilde{\mathscr{E}}^N$.

For $t \geqslant 0$ we define $\tilde{T}_t^N = \exp(t\tilde{\mathscr{L}}^N)$, which is a symmetric strongly continuous semigroup of linear operators on $\hat{H}^N$. We introduce the projective mapping $P^N : \hat{H} \to \hat{H}^N$, $(P^N f)_k := N^d \int_{I_k^N} f(x)\,\mathrm{d}x$, $\forall f \in \hat{H}$. Denote $\mathscr{C}(\mathscr{L}) := \{f \in C^\infty(\bar{D}){:}\ \langle \gamma, \nabla f\rangle|_{\partial D} = 0\}$. One can check that $\mathscr{C}(\mathscr{L})$ is a core of $\mathscr{L}$ by virtue of Ethier and Kurtz (1986, Theorem 8.1.5). Denote $\mathscr{C}(\tilde{\mathscr{L}}) := \{\rho f : f \in \mathscr{C}(\mathscr{L})\}$ so $\mathscr{C}(\tilde{\mathscr{L}})$ is a core of $\tilde{\mathscr{L}}$. Note that $\|\tilde{\mathscr{L}}^N P^N f - P^N \tilde{\mathscr{L}} f\|_N \to 0$ for each $f \in \mathscr{C}(\tilde{\mathscr{L}})$ as $N \to \infty$. It follows by Trotter–Kato theorem that for each $f \in \hat{H}$, $\|\tilde{T}_t^N P^N f - P^N \tilde{T}_t f\|_N \to 0$ for all $t \geqslant 0$, uniformly on $[0,T]$. Hence, $\sup_{N,t \leqslant T} \|\tilde{T}_t^N\|_N < \infty$ by the principle of uniform boundedness.

Let $\{X_{+,N}^{k,y}(t), X_{-,N}^{k,y}(t); X_{+,N}^{k,(1,1)}(t), X_{-,N}^{k,(1,1)}(t); \dots; X_{+,N}^{k,(d,d)}(t), X_{-,N}^{k,(d,d)}(t), k \in D_N\}$ be independent Poisson processes on some probability space $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathbf{P}})$ for each $N \in \mathbb{N}$ (for a practical means of construction of these processes we refer the reader to Kouritzin and Long (2002)). We define from $(\Omega, \mathscr{F}, \mathbf{P})$ and $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathbf{P}})$ the product probability space $(\Omega_0, \mathscr{F}_0, \mathbf{P}_0) = (\Omega \times \bar{\Omega}, \mathscr{F} \otimes \bar{\mathscr{F}}, \mathbf{P} \times \bar{\mathbf{P}})$. Let $l = l(N)$ be a function such that $l(N) \to \infty$ as $N \to \infty$ and

$$n_k^N(0) = \left\lfloor l N^d \int_{I_k^N} \hat{p}_0(x)\,\mathrm{d}x \right\rfloor. \tag{2.1}$$

Hereafter, $\lfloor r \rfloor$ denotes the greatest integer not more than a real number r. Then, motivated by Ethier and Kurtz (1986, pp. 326–327) and Kouritzin and Long (2002), we let

$$
\begin{aligned}
n_k^N(t) = n_k^N(0) &+ X_{+,N}^{k,y}\left(\int_0^t \left[\left(h_k^N \dot{y}_s - \frac{1}{2}(h_k^N)^2 - v_k^N\right) n_k^N(s)\right]^+ \mathrm{d}s\right) \\
&- X_{-,N}^{k,y}\left(\int_0^t \left[\left(h_k^N \dot{y}_s - \frac{1}{2}(h_k^N)^2 - v_k^N\right) n_k^N(s)\right]^- \mathrm{d}s\right) \\
&+ \sum_{i,j=1}^d \left[X_{+,N}^{k,(i,j)}\left(\int_0^t \delta_{i,j,N}^{k,+}(n^N(s))\,\mathrm{d}s\right) - X_{-,N}^{k,(i,j)}\left(\int_0^t \delta_{i,j,N}^{k,-}(n^N(s))\,\mathrm{d}s\right)\right] \\
&- \sum_{i,j=1}^d \left[X_{+,N}^{k-e_j,(i,j)}\left(\int_0^t \delta_{i,j,N}^{k-e_j,+}(n^N(s))\,\mathrm{d}s\right)\right. \\
&\left.- X_{-,N}^{k-e_j,(i,j)}\left(\int_0^t \delta_{i,j,N}^{k-e_j,-}(n^N(s))\,\mathrm{d}s\right)\right] 1_{k\in(D_N+e_j)},
\end{aligned} \tag{2.2}
$$

where $\delta_{i,j,N}^{k,+}(n^N)$, $\delta_{i,j,N}^{k,-}(n^N)$ denote, respectively, the positive, negative parts of

$$
\delta_{i,j,N}^k(n^N) = \begin{cases} \dfrac{N^2}{2} a_{ij}^N(k)(n_{k+e_i}^N - n_k^N), & k \in D_N^0, \\ 0, & \text{otherwise.} \end{cases}
$$

Eq. (2.2) provides a very explicit and powerful construction of our Markov chain approximations to Eq. (1.8), and can be implemented directly on a computer.

By Ethier and Kurtz (1986, Appendixes, Theorem 8.1), there exists $\tilde{\mathbf{P}} : \Omega \times \bar{\mathscr{F}} \to [0,1]$ such that for each $\omega \in \Omega$, $\tilde{\mathbf{P}}^\omega(\cdot) := \tilde{\mathbf{P}}(\omega, \cdot)$ is a probability measure on $\bar{\mathscr{F}}$, for each $B \in \bar{\mathscr{F}}$, $\omega \to \tilde{\mathbf{P}}^\omega(B)$ is $\mathscr{F}$-measurable, and $\mathbf{P}_0(\mathrm{d}\omega_0) = \tilde{\mathbf{P}}^\omega(\mathrm{d}\bar{\omega})\mathbf{P}(\mathrm{d}\omega)$, $\omega_0 = (\omega, \bar{\omega})$. Note that $\tilde{\mathbf{P}}^\omega$ is the probability measure for the quenched results. However, to use the quenched results within the annealed ones we need to know that $\omega \to \tilde{\mathbf{P}}^\omega(B)$ is measurable for each $B \in \bar{\mathscr{F}}$.

For $k \in D_N$, we have

$$
\begin{aligned}
n_k^N(t) = n_k^N(0) &+ \int_0^t \left(h_k^N \dot{y}_s + \alpha - \frac{1}{2}(h_k^N)^2\right) n_k^N(s)\,\mathrm{d}s + \int_0^t \tilde{\mathscr{L}}^N n_k^N(s)\,\mathrm{d}s \\
&+ Z_{k,y}^N(t) + \sum_{i,j=1}^d Z_{k,(i,j)}^N(t) - \sum_{i,j=1}^d Z_{k-e_j,(i,j)}^N(t) 1_{k\in(D_N+e_j)},
\end{aligned} \tag{2.3}
$$

where

$$Z_{k,y}^N(t) = X_{+,N}^{k,y}\left(\int_0^t \left[\left(h_k^N \dot{y}_s - \frac{1}{2}(h_k^N)^2 - v_k^N\right) n_k^N(s)\right]^+ \mathrm{d}s\right)$$

$$- X_{-,N}^{k,y}\left(\int_0^t \left[\left(h_k^N \dot{y}_s - \frac{1}{2}(h_k^N)^2 - v_k^N\right) n_k^N(s)\right]^- \mathrm{d}s\right)$$

$$- \int_0^t \left(h_k^N \dot{y}_s - \frac{1}{2}(h_k^N)^2 - v_k^N\right) n_k^N(s)\,\mathrm{d}s$$

and

$$Z_{k,(i,j)}^N(t) = X_{+,N}^{k,(i,j)}\left(\int_0^t \delta_{i,j,N}^{k,+}(n^N(s))\,\mathrm{d}s\right) - X_{-,N}^{k,(i,j)}\left(\int_0^t \delta_{i,j,N}^{k,-}(n^N(s))\,\mathrm{d}s\right)$$

$$- \int_0^t \delta_{i,j,N}^{k}(n^N(s))\,\mathrm{d}s, \qquad 1 \leqslant i,j \leqslant d.$$

To get the density in each cell, we divide $n^N(t)$ by l and, consequently, the description of the stochastic particle model can be given by

$$\hat{p}^{l,N}(t,x) = \sum_{k \in D_N} \frac{n_k^N(t)}{l} 1_k^N(x),$$

where $1_k^N(\cdot)$ is the indicator function on I_k^N. Now, we set

$$Z_y^{l,N}(t) := \sum_{k \in D_N} l^{-1} Z_y^{l,N}(t) 1_k^N,$$

and

$$Z_A^{l,N}(t) := \sum_{i,j=1}^d l^{-1} \sum_{k \in D_N^0 \cup (D_N^0 + e_j)} (Z_{k,(i,j)}^N(t) - Z_{k-e_j,(i,j)}^N(t)) 1_k^N.$$

Let $\mathscr{G}_t^N$ denote the σ-algebra generated by the observations up to time t, $n^N(0)$, the time changed Poisson processes used to construct n^N, and the collection of all $\mathbf{P}_0$ null sets. Then, similar to Kouritzin and Long (2002, Lemma 2.5), one can show that both $Z_y^{l,N}(t)$ and $Z_A^{l,N}(t)$ are L^2-martingales with respect to $\mathscr{G}_t^N$ under probability measure $\tilde{P}^\omega$.

To ease notation, for $f \in \hat{H}$ and $t \geqslant 0$, we denote

$$\tilde{\mathscr{L}}_N f := \sum_{k \in D_N} (\tilde{\mathscr{L}}^N P^N f)_k 1_k^N, \qquad \tilde{T}_N(t) f := \sum_{k \in D_N} (\tilde{T}^N(t) P^N f)_k 1_k^N.$$

We define $h_N(x) := \sum_{k \in D_N} h_k^N 1_k^N(x)$. Then, it follows from (2.3) that

$$\hat{p}^{l,N}(t) = \hat{p}^{l,N}(0) + \int_0^t \left(h_N \dot{y}_s + \alpha - \frac{1}{2}(h_N)^2\right) \hat{p}^{l,N}(s)\,\mathrm{d}s$$

$$+ \int_0^t \tilde{\mathscr{L}}_N \hat{p}^{l,N}(s)\,\mathrm{d}s + Z_y^{l,N}(t) + Z_A^{l,N}(t).$$

Using variation of constants, $\hat{p}^{l,N}(t)=\hat{p}^{l,N}(t,\omega_0)$ can be written as

$$\hat{p}^{l,N}(t)=\tilde{T}_N(t)\hat{p}^{l,N}(0)+\int_0^t \tilde{T}_N(t-s)\left(\left(h_N\dot{y}(s)+\alpha-\frac{1}{2}(h_N)^2\right)\hat{p}^{l,N}(s)\right)\mathrm{d}s$$
$$+\int_0^t \tilde{T}_N(t-s)\,\mathrm{d}Z_y^{l,N}(s)+\int_0^t \tilde{T}_N(t-s)\,\mathrm{d}Z_A^{l,N}(s). \tag{2.4}$$

3. Law of large numbers

Following Section 2, all results in this section are stated for y^n with fixed $n\in\mathbb{N}$. Throughout this section, we assume that $(N,l(N))$ is any sequence satisfying $l(N)\to\infty$ as $N\to\infty$. Then, our dependence on (l,N) is reduced to dependence only on N and we will write $\hat{p}^N$ for $\hat{p}^{l(N),N}$. For $f:\bar{D}\to\mathbb{R}$ we define $\|f\|_\infty:=\sup_{x\in\bar{D}}|f(x)|$. Now we have the following *quenched* law of large numbers:

Theorem 3.1. *Suppose that* $\|\hat{p}_0\|_\infty<\infty$. *Then, for each fixed* $\omega\in\Omega$,

$$\sup_{t\leqslant T}\tilde{\mathbf{E}}^\omega\|\hat{p}^N(t)-\hat{p}(t)\|^2\to 0 \textit{ as } N\to\infty.$$

Before proving Theorem 3.1, we prepare some preliminary lemmas. For convenience, we denote $f_N:=\sum_{k\in D_N}(P^Nf)_k 1_k^N$ for $f\in\hat{H}$ and $N\in\mathbb{N}$.

Lemma 3.2. *For any* $f\in\hat{H}$, *we have*

$$\tilde{\mathbf{E}}^\omega[\langle Z_y^N(t),f\rangle^2]\leqslant\frac{1}{N^d l}\tilde{\mathbf{E}}^\omega\int_0^t\left\langle\left|\left(h_N\dot{y}(s)-\frac{1}{2}(h_N)^2-v_N\right)\hat{p}^N(s)\right|,f_N^2\right\rangle\mathrm{d}s,$$

and for some constant $C>0$

$$\tilde{\mathbf{E}}^\omega[\langle Z_A^N(t),f\rangle^2]\leqslant\frac{C}{N^d l}\tilde{\mathscr{E}}^N(P^Nf_N,P^Nf_N)\int_0^t\|\tilde{\mathbf{E}}^\omega|\hat{p}^N(s)|\|_\infty\,\mathrm{d}s.$$

Proof. Inasmuch as the proofs of two parts follow the same steps, we just show the second part. By independence, the fact $[X_{\cdot\wedge\tau}]_t=[X]_{t\wedge\tau}$ for stopping time τ, the bilinear property of quadratic variation, and the fact that $\langle Z_A^N(t),f\rangle$ is an L^2-martingale, we get

$$\tilde{\mathbf{E}}^\omega(\langle Z_A^N(t),f\rangle^2)$$
$$=\tilde{\mathbf{E}}^\omega\left\{\left[\sum_{i,j=1}^d l^{-1}\left(\sum_{k\in D_N^0\cup(D_N^0+e_j)}(Z_{k,(i,j)}^N(t)-Z_{k-e_j,(i,j)}^N(t))\right)\langle 1_k^N,f\rangle\right]_t\right\}$$
$$=\sum_{i,j=1}^d\sum_{k\in D_N^0}l^{-2}(\langle 1_k^N,f\rangle-\langle 1_{k+e_j}^N,f\rangle)^2\tilde{\mathbf{E}}^\omega[Z_{k,(i,j)}^N]_t$$

$$= \sum_{i,j=1}^{d} \sum_{k \in D_N^0} l^{-2} \frac{N^2}{2} (\langle 1_k^N, f \rangle - \langle 1_{k+e_j}^N, f \rangle)^2 \int_0^t \tilde{\mathbf{E}}^\omega |a_{ij}^N(k)(n_{k+e_i}^N(s) - n_k^N(s))| \, ds$$

$$\leqslant \frac{C}{N^d l} \tilde{\mathscr{E}}^N(P^N f_N, P^N f_N) \int_0^t \|\tilde{\mathbf{E}}^\omega |\hat{p}^N(s)|\|_\infty \, ds. \qquad \square$$

Next, we need to estimate the moments of $\hat{p}^N(t)$. Similar to Kouritzin and Long (2002, Lemma 3.4), we have the following lemma.

Lemma 3.3. *Suppose that $\|\hat{p}_0\|_\infty < \infty$. Then, for each fixed $\omega \in \Omega$,*

$$\sup_{s \leqslant t} \|\tilde{\mathbf{E}}^\omega |\hat{p}^N(s)|\|_\infty \leqslant D(t, l, \omega) < \infty,$$

where $D(\cdot)$ can be chosen to be increasing in t, decreasing in l and measurable in ω.

Proof. Setting $\hat{p}^N(t) = \sum_{k \in D_N} \hat{p}^N(t,k) 1_k^N$ and $\xi_k^N = N^{d/2} 1_k^N$. We obtain from (2.4) that

$$|\hat{p}^N(t,k)| \leqslant \langle \tilde{T}_N(t) \hat{p}^N(0), \xi_k^N \rangle N^{d/2}$$
$$+ \left\langle \int_0^t \tilde{T}_N(t-s) \left| \left(h_N \dot{y}(s) + \alpha - \frac{1}{2}(h_N)^2 \right) \hat{p}^N(s) \right| ds, \xi_k^N \right\rangle N^{d/2}$$
$$+ \left| \left\langle \int_0^t \tilde{T}_N(t-s) \, dZ_y^N(s), \xi_k^N \right\rangle \right| N^{d/2}$$
$$+ \left| \left\langle \int_0^t \tilde{T}_N(t-s) \, dZ_A^N(s), \xi_k^N \right\rangle \right| N^{d/2}.$$

By the symmetry and the uniform bound on $(\tilde{T}_N(s))_{s \leqslant t}$, one finds that

$$\tilde{\mathbf{E}}^\omega \langle \tilde{T}_N(t) \hat{p}^N(0), \xi_k^N \rangle N^{d/2} \leqslant D_1(t) \|\tilde{\mathbf{E}}^\omega \hat{p}^N(0)\|_\infty \tag{3.1}$$

and

$$\tilde{\mathbf{E}}^\omega \left\langle \int_0^t \tilde{T}_N(t-s) \left| \left(h_N \dot{y}(s) + \alpha - \frac{1}{2}(h_N)^2 \right) \hat{p}^N(s) \right| ds, \xi_k^N \right\rangle N^{d/2}$$
$$\leqslant D_2(t, \omega) \int_0^t \|\tilde{\mathbf{E}}^\omega |\hat{p}^N(s)|\|_\infty \, ds. \tag{3.2}$$

Now, following the arguments in the proof of Kotelenez (1988, Lemma 3.2), for fixed $t > 0$ and $J \in \{y, A\}$, we define L^2-martingales by

$$L_J^N(s,k) = \begin{cases} \left\langle \int_0^s \tilde{T}_N(t-v) \, dZ_J^N(v), \xi_k^N \right\rangle N^{d/2}, & s \leqslant t, \\ L_J^N(t,k), & s > t. \end{cases}$$

Similar to Lemma 3.2, we get by independence, the bilinear property of quadratic variation and the martingale property that

$$\tilde{\mathbf{E}}^{\omega}[L_y^N(\cdot,k)]_s \leqslant \frac{1}{l}\tilde{\mathbf{E}}^{\omega}\int_0^s \left\langle \left| \left(h_N\dot{y}(v) - \frac{1}{2}(h_N)^2 - v_N \right) \hat{p}^N(v) \right|, (\tilde{T}_N(t-v)\xi_k^N)^2 \right\rangle \mathrm{d}v \tag{3.3}$$

and

$$\tilde{\mathbf{E}}^{\omega}[L_A^N(\cdot,k)]_s \leqslant \frac{C}{l}\sup_{v\leqslant s}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(v)|\|_\infty \int_0^s \tilde{\mathscr{E}}^N(\tilde{T}^N(t-v)P^N\xi_k^N, \tilde{T}^N(t-v)P^N\xi_k^N)\,\mathrm{d}v. \tag{3.4}$$

Then, by (3.3), Jensen's inequality and the fact $\sup_{N,s\leqslant t}\|\tilde{T}_s^N\|_N < \infty$, we get

$$\tilde{\mathbf{E}}^{\omega}|L_y^N(t,k)| \leqslant D_3(t,\omega)l^{-1/2}\left(\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty\right)^{1/2}. \tag{3.5}$$

By (3.4), (1.1) and Jensen's inequality, we get

$$\begin{aligned}
&\tilde{\mathbf{E}}^{\omega}|L_A^N(t,k)| \\
&\quad\leqslant \left(\frac{C}{l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \int_0^t \tilde{\mathscr{E}}^N(\tilde{T}^N(t-s)P^N\xi_k^N, \tilde{T}^N(t-s)P^N\xi_k^N)\,\mathrm{d}s\right)^{1/2} \\
&\quad= \left(\frac{C}{l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \int_0^t \left\langle -\tilde{\mathscr{L}}^N\tilde{T}^N(t-s)P^N\xi_k^N, \tilde{T}^N(t-s)P^N\xi_k^N)\,\mathrm{d}s\right\rangle_N\right)^{1/2} \\
&\quad= \left(\frac{C}{l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \cdot \frac{\|P^N\xi_k^N\|_N^2 - \|\tilde{T}^N(t)P^N\xi_k^N\|_N^2}{2}\right)^{1/2} \\
&\quad\leqslant \left(\frac{C}{2l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty\right)^{\frac{1}{2}}.
\end{aligned} \tag{3.6}$$

Combining (3.1), (3.2), (3.5) and (3.6), one finds that

$$\begin{aligned}
\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \leqslant D_4(t,\omega)\Bigg(&\|\tilde{\mathbf{E}}^{\omega}\hat{p}^N(0)\|_\infty + \int_0^t \sup_{v\leqslant s}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(v)|\|_\infty\,\mathrm{d}s \\
&+ l^{-1/2}\left(\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty\right)^{1/2}\Bigg).
\end{aligned}$$

By the assumption $\|\hat{p}_0\|_\infty < \infty$ and (2.1), it is easy to see that $\sup_N \|\tilde{\mathbf{E}}^\omega \hat{p}^N(0)\|_\infty < \infty$. Therefore, by Gronwall's inequality and the inequality $a^{1/2}b^{1/2} \leqslant \frac{1}{2}(a+b)$, we conclude that

$$\sup_{s \leqslant t} \|\tilde{\mathbf{E}}^\omega |\hat{p}^N(s)|\|_\infty \leqslant D(t,l,\omega) < \infty,$$

where $D(\cdot)$ can be chosen as desired. □

Finally we are in a position to prove Theorem 3.1.

Proof of Theorem 3.1. For convenience, we put

$$V_J^N(t) = \int_0^t \tilde{T}_N(t-s)\,\mathrm{d}Z_J^N(s), \quad J = y, A$$

and find by variation of constants that

$$V_J^N(t) = \int_0^t \tilde{\mathscr{L}}_N V_J^N(s)\,\mathrm{d}s + Z_J^N(t), \quad J \in \{y, A\}.$$

For each $N \in \mathbb{N}$, we let $\{(\lambda_p^N, \phi_p^N)\}$ be the eigenvalues and eigenfunctions of $\tilde{\mathscr{L}}_N$. Denote $\langle V_J^N, \phi_p^N \rangle := V_{J,p}^N$ and $\langle Z_J^N, \phi_p^N \rangle := Z_{J,p}^N$ for $J = y, A$. Then, by Itô's rule, we get

$$V_{J,p}^N(t) = \int_0^t \lambda_p^N V_{J,p}^N(s)\,\mathrm{d}s + Z_{J,p}^N(t)$$

and

$$[V_{J,p}^N(t)]^2 = 2\lambda_p^N \int_0^t [V_{J,p}^N(s)]^2\,\mathrm{d}s + 2\int_0^t V_{J,p}^N(s-)\,\mathrm{d}Z_{J,p}^N(s) + [Z_{J,p}^N]_t. \tag{3.7}$$

Using (3.7), the fact $\lambda_p^N \leqslant 0$ and Lemma 3.2 with $f = \phi_p^N$, we get

$$\begin{aligned}\tilde{\mathbf{E}}^\omega[V_{y,p}^N(t)|^2 &\leqslant \frac{1}{N^d l}\tilde{\mathbf{E}}^\omega \int_0^t \left\langle \left| \left(h_N \dot{y}(s) - \frac{1}{2}(h_N)^2 - v_N \right) \hat{p}^N(s) \right|, (\phi_p^N)^2 \right\rangle \mathrm{d}s \\ &\leqslant \frac{K_1(t,\omega)}{N^d l} \int_0^t \|\tilde{\mathbf{E}}^\omega |\hat{p}^N(s)|\|_\infty\,\mathrm{d}s.\end{aligned}$$

Then, $\tilde{\mathbf{E}}^\omega \|V_y^N(t)\|^2 \leqslant (K_1(t,\omega)L_1 \cdots L_d/l) \int_0^t \|\tilde{\mathbf{E}}^\omega |\hat{p}^N(s)|\|_\infty\,\mathrm{d}s$. Thus, by Lemma 3.3 and the assumption $\|\hat{p}_0\|_\infty < \infty$, we conclude that

$$\sup_{t \leqslant T} \tilde{\mathbf{E}}^\omega \|V_y^N(t)\|^2 \to 0 \quad \text{as } N \to \infty. \tag{3.8}$$

By (3.7), Lemmas 3.2 and 3.3, we get

$$\begin{aligned}
&\tilde{\mathbf{E}}^{\omega}|V_{A,p}^N(t)|^2\\
&\quad= \int_0^t \exp\{2\lambda_p^N(t-s)\}\mathrm{d}\tilde{\mathbf{E}}^{\omega}[Z_{J,p}^N]_s\\
&\quad\leqslant \int_0^t \frac{C}{N^d l}\,\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \tilde{\mathscr{E}}^N(P^N\phi_p^N, P^N\phi_p^N)\exp\{2\lambda_p^N(t-s)\}\,\mathrm{d}s\\
&\quad\leqslant \frac{C}{N^d l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \int_0^t \tilde{\mathscr{E}}^N(\tilde{T}^N(t-s)P^N\phi_p^N, \tilde{T}^N(t-s)P^N\phi_p^N)\,\mathrm{d}s\\
&\quad= \frac{C}{N^d l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty \int_0^t \left\langle -\tilde{\mathscr{L}}^N\tilde{T}^N(t-s)P^N\phi_p^N, \tilde{T}^N(t-s)P^N\phi_p^N\right\rangle_N\,\mathrm{d}s\\
&\quad= \frac{C}{N^d l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty\cdot\frac{\|P^N\phi_p^N\|_N^2 - \|\tilde{T}^N(t)P^N\phi_p^N\|_N^2}{2}\\
&\quad\leqslant \frac{C}{2N^d l}\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty.
\end{aligned}$$

Then, $\tilde{\mathbf{E}}^{\omega}\|V_A^N(t)\|^2 \leqslant (CL_1\cdots L_d/2l)\sup_{s\leqslant t}\|\tilde{\mathbf{E}}^{\omega}|\hat{p}^N(s)|\|_\infty$ and thus

$$\sup_{t\leqslant T}\tilde{\mathbf{E}}^{\omega}\|V_A^N(t)\|^2 \to 0 \quad \text{as } N\to\infty. \tag{3.9}$$

By (1.8) and (2.4), we get

$$\begin{aligned}
\|\hat{p}^N(t)-\hat{p}(t)\|^2 &\leqslant 4\Bigg(\|\tilde{T}_N(t)\hat{p}^N(0)-\tilde{T}(t)\hat{p}(0)\|^2\\
&\quad+\int_0^t \Bigg\|\tilde{T}_N(t-s)\left(\left(h_N\dot{y}(s)+\alpha-\frac{1}{2}(h_N)^2\right)\hat{p}^N(s)\right)\\
&\quad-\tilde{T}(t-s)\left(\left(h\dot{y}(s)+\alpha-\frac{1}{2}h^2\right)\hat{p}(s)\right)\Bigg\|^2\,\mathrm{d}s\\
&\quad+\|V_y^N(t)\|^2+\|V_A^N(t)\|^2\Bigg)\\
&\leqslant 8\Bigg(\|\tilde{T}_N(t)[\hat{p}^N(0)-\hat{p}(0)]\|^2+\|[\tilde{T}_N(t)-\tilde{T}(t)]\hat{p}(0)\|^2\\
&\quad+\int_0^t \Bigg\|\tilde{T}_N(t-s)\left[\left(h_N\dot{y}(s)+\alpha-\frac{1}{2}(h_N)^2\right)\hat{p}^N(s)\right.\\
&\quad\left.-\left(h\dot{y}(s)+\alpha-\frac{1}{2}h^2\right)\hat{p}(s)\right]\Bigg\|^2\,\mathrm{d}s\\
&\quad+\int_0^t \left\|[\tilde{T}_N(t-s)-\tilde{T}(t-s)]\left(\left(h\dot{y}(s)+\alpha-\frac{1}{2}h^2\right)\hat{p}(s)\right)\right\|^2\,\mathrm{d}s\\
&\quad+\|V_y^N(t)\|^2+\|V_A^N(t)\|^2\Bigg).
\end{aligned}$$

By (1.8) and Gronwall's inequality, $\sup_{t\leqslant T}\|\hat{p}(t)\|^2 \leqslant K_2(\omega) < \infty$ for each fixed ω. Note that for each $f\in\hat{H}$, $\tilde{T}_N(t)f \to \tilde{T}(t)f$ for all $t\geqslant 0$, uniformly on $[0,T]$, and $\sup_{N,t\leqslant T}\|\tilde{T}_N(t)\| < \infty$. By (3.8) and (3.9), dominated convergence theorem and Gronwall's inequality, we conclude that

$$\sup_{t\leqslant T}\tilde{\mathbf{E}}^{\omega}\|\hat{p}^N(t)-\hat{p}(t)\|^2 \to 0 \quad \text{as } N\to\infty. \qquad \square$$

Using the boundedness assumption of the process $z(t)$ and following the same procedure as used in Theorem 3.1, we have the following *annealed* law of large numbers.

Theorem 3.4. *Suppose that* $\|\hat{p}_0\|_\infty < \infty$. *Then*

$$\sup_{t\leqslant T}\mathbf{E}_0\|\hat{p}^N(t)-\hat{p}(t)\|^2 \to 0 \quad \textit{as } N\to\infty,$$

where $\mathbf{E}_0$ *is the expectation with respect to* $\mathbf{P}_0$.

Remark 3.5. Before ending this section, we would like to point out that our approximate Dirichlet form $\tilde{\mathscr{E}}^N$ in Section 2 yields a Markov Chain. So we are really almost filtering an approximate signal with an approximate observation. In this connection, we refer the interested reader to the recent paper by Bhatt et al. (1999). In the paper, Bhatt et al. showed that the filter depends continuously on the law of the signal. Note that our approximations are not filters so their analysis would not apply. Compared with their result, our Markov chain approximation is far more explicit and has different type of convergence.

4. Practical application: fish tracking problem

In this section, we use simulation results to compare computational efficiency of our Markov chain method to the particle filter methods. Motivated by the fish farming industry, the test problem is the tracking of a single fish in a tank with boundary reflections. For simplicity, we choose a two-dimensional fish motion described by the following Skorohod SDE

$$\mathrm{d}x_t = \beta\,\mathrm{d}v_t - \alpha\left(x_t - \frac{L}{2}\right)\mathrm{d}t + \chi_{\partial D}(x_t)\gamma(x_t)\,\mathrm{d}\eta_t,$$

where $L=(L_1,L_2)^{\mathrm{T}}$ is the size of the tank, α and β are parameters, and v_t, γ and η_t are defined as in Section 1. In the simulations, we take $\alpha=0.00005$, $\beta=0.02$ and simplify our example by selecting $L_1=L_2=1$.

The observation process consists of a discrete sequence of images arriving at observation times $\{t_k\}_{k=1}^{\infty}$, each observation being a 2-dimensional raster $\{y_{t_k}^{(i,j)}\}_{i,j=1,1}^{R,R}$. $y_{t_k}^{(i,j)}$ is the (i,j)th component of a raster depiction of the observation. We let $h^{(i,j)}(\cdot)$ be the

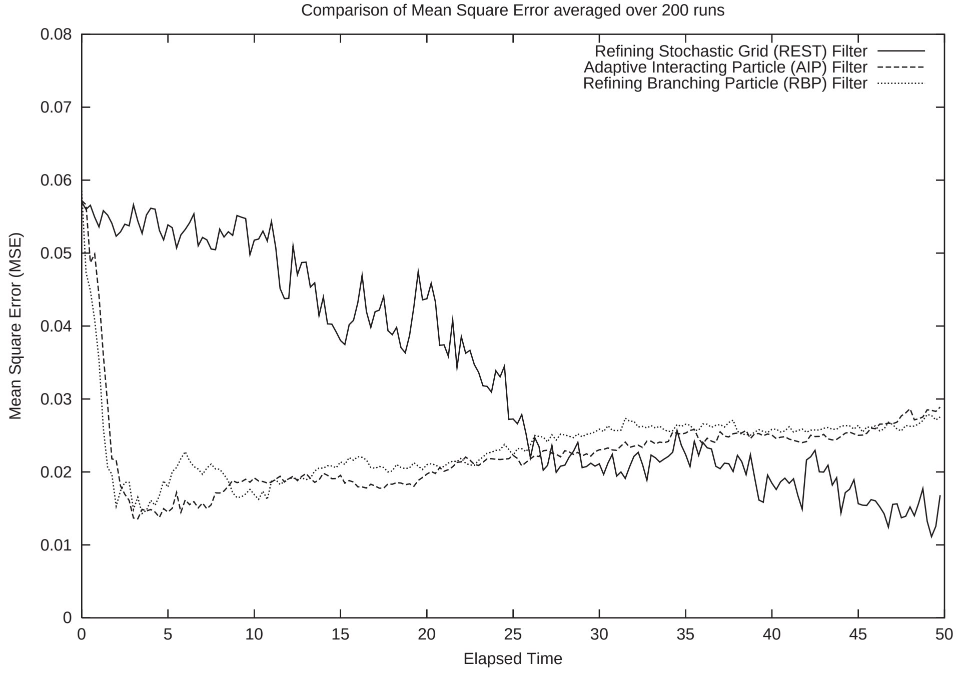

Fig. 1. REST filter initialized with 2500 particles, AIP filter initialized with 35 000 particles, RBP filter initialized with 35 000 particles.

indicator function

$$h^{(i,j)}(z_1,z_2) = 1_{\left[\frac{z_1 R}{L_1}-\frac{3}{2},\frac{z_1 R}{L_1}+\frac{3}{2}\right]\times\left[\frac{z_2 R}{L_2}-\frac{3}{2},\frac{z_2 R}{L_2}+\frac{3}{2}\right]}(i,j)$$

representing a 3×3 pixel square image of the fish and $w_k^{(i,j)}$ be pixel-by-pixel standard Gaussian noise. Then, an observation at time t_k is constructed by superimposing the square of the signal onto the raster and adding noise according to the formula

$$y_{t_k}^{(i,j)} = h^{(i,j)}(x_{t_k}) + w_k^{(i,j)}. \tag{4.1}$$

For our simulations, the length and width of the observation rasters, R, is 256. Observations are not preprocessed, the information from the raster pixel is used directly in the filter algorithm.

The observations given by (4.1) are taken at discrete times which makes the practical algorithm slightly different from that given in Section 2. However, we can follow the similar ideas presented in Section 2 to construct Markov chain approximations to the corresponding Duncan–Mortensen–Zakai equation with discrete time observations. In the following, we apply the so called refining stochastic grid (REST) filter, developed from the Markov chain method in Section 2, to do the

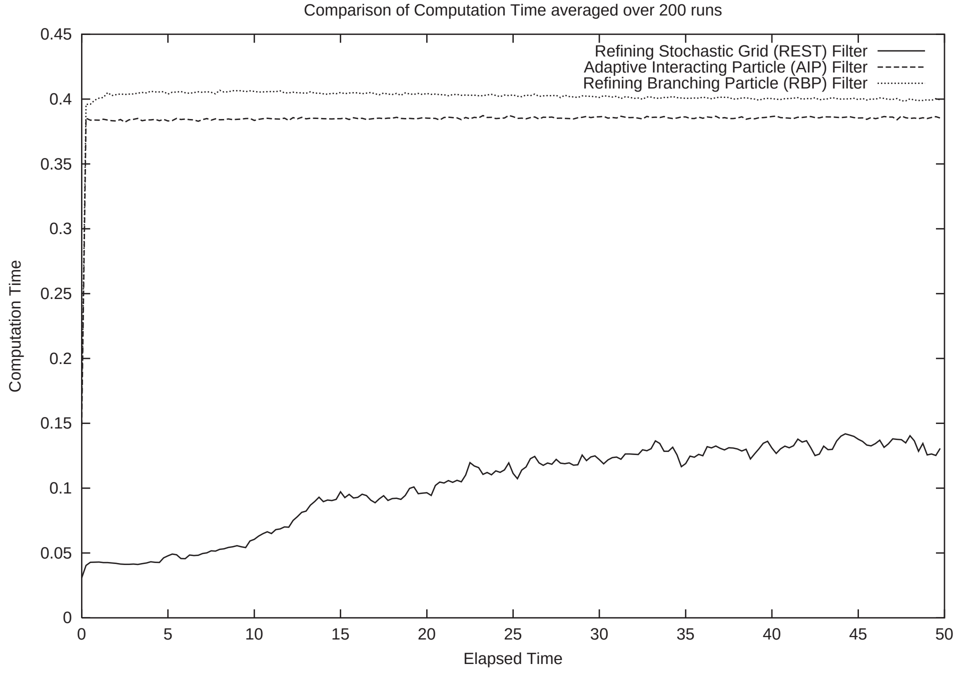

Fig. 2. REST filter initialized with 2500 particles, AIP filter initialized with 35 000 particles, RBP filter initialized with 35 000 particles.

simulations. This practical algorithm to implement our filter is reduced to an algorithm to implement a specific time-inhomogeneous Markov chain, which can be done using a single Poisson process and independent sequences of Bernoulli trials. The inhomogeneity is due to the observations themselves. The discretization of state space results in *particles* representing a small mass of the conditional distribution at particular grid points in the signal domain. These particles diffuse, drift, give birth, and die within the region. The particles contain information from the observations through observation-dependent births and deaths. We refer the interested readers to Ballantyne et al. (2002) for more details about this refining method.

Comparison data of the REST filter, AIP filter and RBP filter are presented in Figs. 1–4. 2 500 *particles* are initially used for the REST filter in Figs. 1–4. 35 000 and 20 000 particles are initially used for each particle filter in Figs. 1–2 and Figs. 3–4, respectively, where particle means an independent copy of the signal. Graphs of the average mean square error (MSE) in the position estimates over the simulated time for 200 runs are provided in Figs. 1 and 3. For each run, we simulate over a period of 50 time units with observation arriving at every 0.25 time unit. Here, MSE at time t_k denotes the Euclidean distance between the true signal position and the

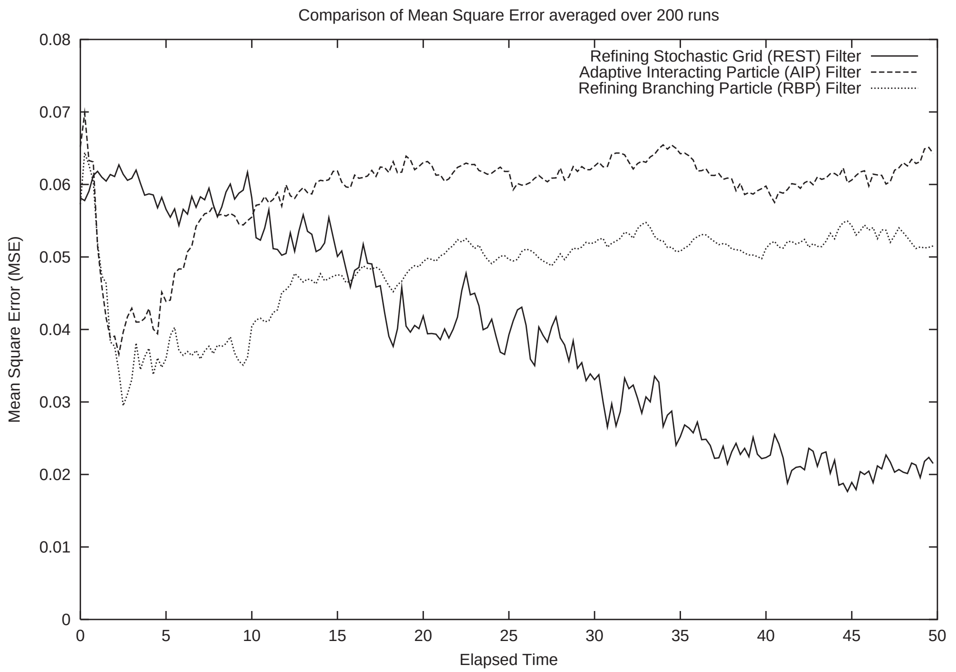

Fig. 3. REST filter initialized with 2500 particles, AIP filter initialized with 20 000 particles, RBP filter initialized with 20 000 particles.

approximated filter at the time of each observation. From Figs. 1 and 3, we see that it may take a longer time for the REST filter to localize the target due to the fact that an initial computational burst is disallowed in this filter, but after localization, the MSE for the REST filter becomes smaller and smaller with tiny fluctuation as time elapses. For the AIP and RBP filters, they can localize the target faster but may lose the target later on. Although all the three filtering algorithms, being at least adaptive, are readily able to localize the target, the REST filter is the best one. From Figs. 1 and 3, we also see that the RBP filter is more efficient than the AIP filter when the number of particles is not too large, although there is little difference between their efficiency when the number of particles is large. From Figs. 2 and 4, we find that the computation time for the REST filter is much less than that of the AIP and RBP filters.

From the simulation results, we can conclude that the new method of solving nonlinear filtering problems numerically introduced in this paper, which uses a Markov chain to push particles about fixed grid points, provides a mathematically sound solution to a general class of such problems and a practical solution to the given specific problem.

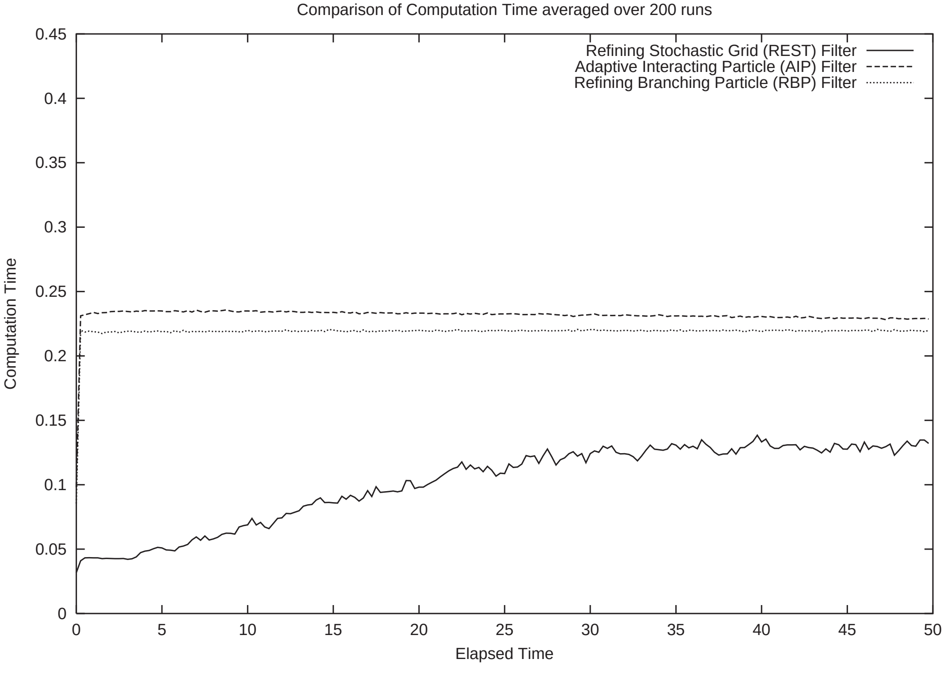

Fig. 4. REST filter initialized with 2500 particles, AIP filter initialized with 20 000 particles, RBP filter initialized with 20 000 particles.

Acknowledgements

The authors gratefully acknowledge the support and sponsorship of Lockheed Martin Naval Electronic and Surveillance Systems, Lockheed Martin Canada, Acoustic Positioning Research, the Pacific Institute of Mathematical Sciences, the Natural Science and Engineering Research Council (NSERC) through the Prediction in Interacting Systems (PINTS) centre of the Mathematics of Information Technology and Complex Systems (MITACS) network of centres of excellence, and the University of Alberta. The authors also thank Surrey Kim for providing the simulation results.

References

Arnold, L., Theodosopulu, M., 1980. Deterministic limit of the stochastic model of chemical reactions with diffusion. Adv. Appl. Probab. 12, 367–379.

Ballantyne, D.J., Chan, H.Y., Kouritzin, M.A., 2000. A novel branching particle method for tracking. Signal and Data Processing of Small Targets 2000, in: O.E. Drummond (Ed.), Proceedings of SPIE, Vol. 4048, Orlando, FL, pp. 277–287.

Ballantyne, D.J., Kouritzin, M.A., Long, H.W., Sun, W., 2002. Discrete-space particle filters for reflecting diffusions. Proceedings of the 2002 IEEE Aerospace Conference, Big Sky, MT.

Bhatt, A.G., Kallianpur, G., Karandikar, R.L., 1999. Robustness of the nonlinear filter. Stochastic Process. Appl. 81, 247–254.

Blount, D., 1991. Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion. Ann. Probab. 19, 1440–1462.

Blount, D., 1994. Density dependent limits for a nonlinear reaction–diffusion model. Ann. Probab. 22, 2040–2070.

Clark, J.M.C., 1978. The design of robust approximations to the stochastic differential equations of nonlinear filtering. In: Skwirzynski, J.K. (Ed.), Communication Systems and Random Process Theory. Sijthoff and Noordhoff, Alphen a/d Rijn, pp. 721–734.

Crisan, D., Lyons, T., 1997. Nonlinear filtering and measure-valued processes. Probab. Theory Related Fields 109, 217–244.

Crisan, D., Gaines, J., Lyons, T., 1998. Convergence of a branching particle method to the solution of the Zakai equation. SIAM J. Appl. Math. 58, 1568–1590.

Crisan, D., Del Moral, P., Lyons, T., 1999. Discrete filtering using branching and interacting particle systems. Markov Process. Related Fields 5, 293–318.

Del Moral, P., 1996. Nonlinear filtering: interacting particle solution. Markov Process. Related Fields 2, 555–579.

Del Moral, P., Miclo, L., 2000. Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to non-linear filtering. Séminaire de Probabilités, Vol. XXXIV, Lecture Notes in Mathematics, Vol. 1729. Springer, Berlin, pp. 1–145.

Del Moral, P., Kouritzin, M.A., Miclo, L., 2001. On a class of discrete generation interacting particle systems. Electron. J. Probab. 6, 1–26.

Di Masi, G.B., Runggaldier, W.J., 1981. Continuous-time approximations for the nonlinear filtering problem. Appl. Math. Optim. 7, 233–245.

Di Masi, G.B., Runggaldier, W.J., 1982. On approximation methods for nonlinear filtering. In: Mitter, S.K., Moro, A. (Eds.), Nonlinear Filtering and Stochastic Control, Lecture Notes in Mathematics, Vol. 972. Springer, Berlin, pp. 249–259.

Ethier, S.N., Kurtz, T.G., 1986. Markov Processes, Characterization and Convergence. Wiley, New York.

Freidlin, M., 1985. Functional Integration and Partial Differential Equations. Princeton University Press, Princeton, NJ.

Fukushima, M., Oshima, Y., Takeda, M., 1994. Dirichlet Forms and Symmetric Markov Processes. Walter de Gruyter, Berlin, New York.

Hu, Y., Kallianpur, G., Xiong, J., 2002. An approximation for the Zakai equation. Appl. Math. Optim. 45, 23–44.

Kotelenez, P., 1986. Law of large numbers and central limit theorem for linear chemical reaction with diffusion. Ann. Probab. 14, 173–193.

Kotelenez, P., 1988. High density limit theorems for nonlinear chemical reaction with diffusion. Probab. Theory Related Fields 78, 11–37.

Kouritzin, M., Long, H.W., 2002. Convergence of Markov chain approximations to stochastic reaction diffusion equations. Ann. Appl. Probab. 12, 1039–1070.

Kurtz, T.G., 1971. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. J. Appl. Probab. 9, 344–356.

Kushner, H.J., 1977. Probability Methods for Approximations in Stochastic Control and for Elliptic Equations. Academic Press, New York.

Kushner, H.J., 1979. A robust discrete state approximation to the optimal nonlinear filter for a diffusion. Stochastics 3, 75–83.

Kushner, H.J., Huang, H., 1985. Limits for partial differential equations with wide band stochastic coefficients and an application to filtering theory. Stochastics 14, 115–148.

Pardoux, E., 1982. Equations du filtrage non lineaire de la prediction et du lissage. Stochastics 6, 193–231.

Zakai, M., 1969. On the optimal filtering of diffusion processes. Z. Wahrsch. Verw. Gebiete. 11, 230–243.