

## PARTICLE FILTERS FOR MULTISCALE DIFFUSIONS <sup>\*,\*\*</sup>

ANASTASIA PAPAVALIOU<sup>1</sup>

**Abstract.** We consider multiscale stochastic systems that are partially observed at discrete points of the slow time scale. We introduce a particle filter that takes advantage of the multiscale structure of the system to efficiently approximate the optimal filter.

**Résumé.** On considère des systèmes stochastiques multi-échelles qui sont partiellement observés à des points discrets de l'échelle de temps lente. On introduit un filtre à particule qui utilise la structure multi-échelle du système pour approcher efficacement le filtre optimal.

### INTRODUCTION

We are interested in the problem of estimating a function of a multiscale process that can be approximated by a diffusion which lives the slow scale, when it is partially observed. Such problems come up in many applications, such as molecular dynamics, climate modelling or estimation of stochastic volatility using agent-based models (see [8] for general discussion of multiscale models or [4] and [9] for applications to kinetic Monte-Carlo and climate modelling respectively).

In this paper, we focus on the problem of estimating the slow component of a continuous multiscale process from partial and discrete observations of it. More specifically, we have an  $\mathbb{R}^{p+q}$  process  $X^\epsilon = (X_t^\epsilon)_{t \geq 0} = (X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)})_{t \geq 0}$  that satisfies the following multiscale stochastic differential equation:

$$\begin{cases} dX_t^{(1,\epsilon)} &= a(X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)})dt + \sigma_1(X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)})dW_t^{(1)} \\ dX_t^{(2,\epsilon)} &= \frac{1}{\epsilon}b(X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)})dt + \frac{1}{\sqrt{\epsilon}}\sigma_2(X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)})dW_t^{(2)} \end{cases} \quad (1)$$

where  $X_t^{(1,\epsilon)} \in \mathbb{R}^p$ ,  $X_t^{(2,\epsilon)} \in \mathbb{R}^q$  and  $W_t^{(1)}$  and  $W_t^{(2)}$  are two independent Wiener processes in  $\mathbb{R}^p$  and  $\mathbb{R}^q$  respectively. Let  $\mu$  be the initial distribution, i.e.  $\mu = \mathcal{L}(X_0^\epsilon)$ . We denote by  $\mu_1$  and  $\mu_2$  the marginals on  $X_0^{(1,\epsilon)}$  and  $X_0^{(2,\epsilon)}$  respectively.

We observe the process  $(X_t^\epsilon)_{t \geq 0}$  through  $Y^\epsilon = (Y_{k\Delta}^\epsilon)_{k=0,\dots,T}$ , where  $\Delta \sim \mathcal{O}(1)$ , i.e. the observations live in the the same scale as  $X_t^{(1,\epsilon)}$ , which we call the *slow time scale*. In fact, let us assume for simplicity that  $\Delta = 1$ . The process  $Y^\epsilon$  is given by

$$Y_k^\epsilon = h(X_k^{(1,\epsilon)}, v_k), \quad (2)$$

where  $(v_k)_k$  are i.i.d. random variables with known distribution.

---

\* The author has been partially supported by a Marie Curie International Reintegration Grant, MIRC-CT-2005-029160.

\*\* The author would like to thank Professor I.G. Kevrekidis for suggesting this problem to her.

<sup>1</sup> Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK. Email: a.papavasiliou@warwick.ac.uk

Our goal is to compute the conditional distribution of the slow process  $X_t^{(1,\epsilon)}$  given the observations, at observation points  $k = 1, \dots, T$ , or, equivalently, compute the expectations

$$\pi_k^\epsilon(f) := \mathbb{E} \left( f(X_k^{(1,\epsilon)}) | Y_\ell^\epsilon, \ell = 1, \dots, k \right) \tag{3}$$

for all continuous bounded functions  $f$  on  $\mathbb{R}^p$ , i.e.  $f \in \mathcal{C}_b(\mathbb{R}^p)$ , and  $k = 0, \dots, T$ .

In [2], the authors discuss this problem for an arbitrary diffusion  $(X_t)_{t \geq 0}$  and they develop a particle filter that approximates the conditional distribution (also called *optimal filter*). The additional difficulty compared to discrete systems is that of simulating the process  $(X_t)_t$  between the observation points, i.e. simulate  $X_{(k+1)}$  given  $X_k$ . In [2], this is done by applying the Euler discretization scheme. The step size is given as a function of the number of particles used by the particle filter, chosen so as to optimize the convergence rate. An alternative approach has been recently suggested in [5].

In the case of multiscale diffusions, both the Euler discretization scheme and the MCMC method described in [5] become inefficient. However, if we are only interested in the slow scale marginal of the optimal filter given by (3), we can avoid these problems by replacing the multiscale diffusion by the approximation of the slow scale constructed by applying the averaging principle. When the averaged equation is not available in closed form, we construct a further approximation of its drift and variance using short simulations of the multiscale system (see [3, 7]).

In section 1, we review some of the basic results of [2] for discretely and partially observed diffusions. In section 2, we describe the algorithm and analyze the approximation error in the case were the averaged equation is available in closed form. In section 3, we do the same for the case were the averaged equation is not available in closed form and we apply the heterogeneous multiscale method to approximate it. Finally, in section 4, we discuss how to extent this approach to continuous observation processes.

### 1. DISCRETELY AND PARTIALLY OBSERVED DIFFUSIONS: A REVIEW

Suppose that  $X = (X_t)_{t \geq 0}$ , with  $X_t \in \mathbb{R}^d$ , is a diffusion of the form

$$dX_t = a(X_t)dt + \sigma(X_t)dW_t, \tag{4}$$

with initial distribution  $\mu = \mathcal{L}(X_0)$ . The diffusion process is observed through

$$Y_k = h(X_k, v_k), \tag{5}$$

where  $(v_k)_{k \geq 0}$  are i.i.d. random variables, such that the conditional probability admits a density  $g$ , i.e.  $\mathbb{P}(Y_k \in dy | X_k = x) = g(x, y)dy$ , and  $g$  is bounded and explicitly known. We want to approximate the optimal filter

$$\pi_k = \mathbb{P}(X_k | Y_\ell, \ell = 1, \dots, k). \tag{6}$$

In [2], the authors approximate (6) using a combination of the Euler method and the discrete particle filter. The exact algorithm is as follows:

- *Initialization* ( $k=0$ ): Simulate  $N$  independent random variables  $(\xi_0^j)_{j=1}^N$  from the initial distribution  $\mu$ .
- For  $k > 0$ 
  - (1) *Evolution*: Simulate

$$\hat{\xi}_k^j \sim p_1^{(N)}(\xi_{k-1}^j, \cdot),$$

where  $p_1^{(N)}(x, \cdot)$  is the forward Euler approximation with step  $\frac{1}{\sqrt{N}}$  of the transition kernel

$$p_1(x, \cdot) = \mathbb{P}(X_1 \in \cdot | X_0 = x).$$

(2) *Resampling*: Simulate  $N$  new random variables  $(\xi_k^j)_{j=1}^N$  from

$$\xi_k^j \sim \frac{1}{N} \sum_{j=1}^N \frac{w_k^j}{\sum_{i=1}^M w_k^i} \delta_{\xi_k^j},$$

where the weights  $(w_k^j)_{j=1}^N$  are the likelihood of observing  $Y_k$  if  $X_k = \hat{\xi}_k^j$ , i.e.  $w_k^j := g(\hat{\xi}_k^j, Y_k)$ .

Then, the particle filter  $\pi_k^N = \frac{1}{N} \sum_{j=1}^N \delta_{\xi_k^j}$  converges weakly to the optimal filter  $\pi_k$  defined in (6). More precisely, the following holds:

**Theorem 1.1** (Del Moral – Jacod – Protter, [2]). *For all bounded Borel functions  $f \in \mathcal{B}_b(\mathbb{R}^d)$ , all  $k = 0, \dots, T$  and all  $N \geq 0$ , the approximation error will be bounded by*

$$\mathbb{E} |\pi_k^N(f) - \pi_k(f)| \leq \frac{C_k}{\sqrt{N}} \|f\|_\infty, \quad (7)$$

under the following assumptions

- (1) The functions  $a(\cdot)$  and  $\sigma(\cdot)$  are two times differentiable with bounded derivatives of all orders up to two.
- (2) The covariance matrix is uniformly non degenerate, i.e.  $\sigma \sigma^t(\cdot) > \eta > 0$ .

The constant  $C_k$  depends on the drift and variance of the diffusion, the likelihood function  $g$  and  $k$ .

If the likelihood function  $g(x, y)$  is bounded above and below, i.e. there exists a constant  $K$  such that  $\frac{1}{K} \leq g(x, y) \leq K$  for all  $x$  and  $y$ , then the constant  $C_k$  in theorem 1.1 takes the following form:

$$C_k = (2 + 2\alpha) \frac{(8K^{2T})^{k+1} - 8K^{2T}}{8K^{2T} - 1},$$

where  $\alpha$  is such that

$$\sup_x |p_1^{(N)} f(x) - p_1 f(x)| \leq \frac{\alpha}{\sqrt{N}} \|f\|_\infty,$$

with  $p_1^{(N)}$  and  $p_1$  as above. So,  $\alpha$  is the constant that appears in the upper bound of the error of the approximation of the distribution of  $X_1$  by the forward Euler method (see [1]). Consequently, if we apply the Euler discretization method to the multiscale system (1), the constant  $\alpha$  will be of order  $\alpha \sim \mathcal{O}(\frac{1}{\epsilon})$ .

**Corollary 1.2.** *Under the assumptions of theorem 1.1, if the diffusion process and the observations are of the form (1) and (2) respectively, then the error of the slow scale marginal of the particle filter described above becomes*

$$\mathbb{E} \left| \pi_k^{\epsilon, N}(f) - \pi_k^\epsilon(f) \right| \leq \frac{C'_k}{\epsilon \sqrt{N}} \|f\|_\infty, \quad (8)$$

The above corollary shows that if the diffusion process that we want to estimate is a multiscale diffusion, the particle filter described in [2] will no longer be efficient, just as the Euler discretization method will not be efficient.

## 2. THE MULTISCALE CASE

Since the observations live in the slow scale, we can only hope to get a good approximation of the slow scale marginal of the optimal filter and, consequently, we focus on the approximation of  $\pi_k^\epsilon(f)$  given by (3). A quite natural thing to do in order to avoid simulating the whole multiscale process – which, as we have already seen, is problematic – is to try and replace the slow scale process  $(X_t^{1, \epsilon})_t$  by a diffusion  $(\bar{X}_t)_t$  in  $\mathbb{R}^p$  that does not

depend on the fast scale process  $(X_t^{2,\epsilon})_t$ . This is, indeed, possible under the following assumption:  $\exists \lambda > 0$  such that  $\forall x_1 \in \mathbb{R}^p$  and  $\forall x_2, x'_2 \in \mathbb{R}^q$ ,

$$\langle x_2 - x'_2, b(x_1, x_2) - b(x_1, x'_2) \rangle + \|\sigma_2(x_1, x_2) - \sigma_2(x_1, x'_2)\|^2 \leq -\lambda |x_2 - x'_2|^2, \quad (9)$$

where  $\langle \cdot, \cdot \rangle$ ,  $|\cdot|$  and  $\|\cdot\|$  denote the Euclidean inner product and norm in  $\mathbb{R}^q$  and the Frobenius norm in  $\mathbb{R}^q \times \mathbb{R}^q$ , respectively. In other words, we require both  $b$  and  $\sigma$  to grow sublinearly. This assumption implies that if we fix  $X_t^{1,\epsilon} \equiv x_1$  in (1),  $X_t^{2,\epsilon}$  converges to its unique invariant distribution  $\nu_{x_1}$  exponentially fast, with rate  $\frac{\lambda}{\epsilon}$ . In fact, the necessary assumption is not (9) but this exponential ergodicity property. We approximate the process  $(X_t^{1,\epsilon})_t$  by the diffusion process  $(\bar{X}_t)_t$  satisfying

$$d\bar{X}_t = \bar{a}(\bar{X}_t)dt + \bar{\sigma}(\bar{X}_t)dW_t, \quad \bar{X}_0 \sim \mu_1 = \mathcal{L}(X_0^{1,\epsilon}), \quad (10)$$

where

$$\bar{a}(x) = \int_{\mathbb{R}^q} a(x, z) \nu_x(dz) \quad (11)$$

and

$$\bar{\sigma}(x) = \left( \int_{\mathbb{R}^q} \sigma_1(x, z)^2 \nu_x(dz) \right)^{\frac{1}{2}} \quad (12)$$

From now on, let us assume that the assumptions of theorem 1.1 and (9) hold. Then, it is a well-known result, often referred to as the *averaging principle*, that  $X_t^{1,\epsilon} \rightarrow \bar{X}_t$  as  $\epsilon \rightarrow 0$ . More specifically, the following holds (see [6]):

$$\sup_{0 \leq t \leq T} \left| \mathbb{E}f(X_t^{1,\epsilon}) - \mathbb{E}f(\bar{X}_t) \right| \leq C_{f,T} \cdot \epsilon, \quad \forall f \in \mathcal{C}_b(\mathbb{R}^p). \quad (13)$$

This estimate suggests that we can approximate  $\pi_k^\epsilon(f)$  given by (3) by  $\bar{\pi}_k(f)$  defined by

$$\bar{\pi}_k(f) := \mathbb{E} \left( f(\bar{X}_k) \mid \bar{Y}_\ell = Y_\ell^\epsilon, \ell = 1, \dots, k \right), \quad (14)$$

where  $\bar{Y}_k = h(\bar{X}_k, v_k)$  and  $(v_k)_k$  are i.i.d. random variables as in (2). Indeed, it is a straight forward consequence of (13) and Proposition 2.1 of [2] that

$$\mathbb{E} |\pi_k^\epsilon(f) - \bar{\pi}_k(f)| \leq C_1 \epsilon \|f\|_\infty.$$

If we cannot compute  $\bar{\pi}_k$  explicitly, we approximate it by the particle filter  $\bar{\pi}_k^N$  described in section 1. Then, the total error will be bounded by

$$\mathbb{E} |\pi_k^\epsilon(f) - \bar{\pi}_k^N(f)| \leq C \left( \epsilon + \frac{1}{\sqrt{N}} \right) \|f\|_\infty. \quad (15)$$

Thus, if  $\epsilon$  is small, it is much more efficient to approximate  $\pi_k^\epsilon$  by  $\bar{\pi}_k^N$  rather than  $\pi_k^{\epsilon,N}$  in (8), i.e. if we are willing to accept an approximation error of order  $\delta$ , we will, in general, achieve this with a much smaller number of simulations (and computing time) if we compute  $\bar{\pi}_k^N$  rather than  $\pi_k^{\epsilon,N}$ .

### 3. APPROXIMATING THE AVERAGED EQUATION

In the previous section, we argued that it is, in general, more efficient to approximate the slow marginal of the optimal filter  $\pi_k^\epsilon$  by replacing the slow component of multiscale diffusion by another diffusion, which we call *averaged diffusion*, and then applying the particle filter algorithm, rather than applying it directly to the multiscale diffusion. However, in order to simulate the averaged diffusion  $(\bar{X}_t)_t$  that replaces the slow scale process  $(X_t^{1,\epsilon})_t$ , we need to know its drift and its standard deviation given by (11) and (12) respectively. In

most cases, these are not going to be explicitly known. Then, we replace (11) and (12) by their Monte Carlo estimates, as in [3].

First, we define a new family of diffusion processes as follows. For each  $x \in \mathbb{R}^p$ , we define the process  $Z_t(x)$  as the solution of the following stochastic differential equation:

$$dZ_t(x) = b(x, Z_t(x))dt + \sigma_2(x, Z_t(x))dV_t, \quad Z_0(x) \sim \mu_2, \quad (16)$$

where  $V_t$  is an  $\mathbb{R}^q$ -valued Wiener process. We also define a new approximation to the transition kernel  $\bar{p}_1(x, \cdot)$ , where  $\bar{p}_t(x, \cdot) := \mathbb{P}(\bar{X}_t \in \cdot | \bar{X}_0 = x)$ , so that we can simulate from it exactly, as follows:

(1) For  $k = 0$ :

Simulate  $M$  independent random variables  $(\zeta_{k,n}^i)_{i=1}^M$  from the forward Euler approximation to the distribution of  $Z_n(x)$  defined in (16), with step  $\delta t$  and initial distribution  $\mu_2$ . Simulate  $\xi_1$  from

$$\xi_1 \sim \text{Gsn} \left( \Delta t \left( \frac{1}{M} \sum_{i=1}^M a(x, \zeta_{k,n}^i) \right), \Delta t \left( \frac{1}{M} \sum_{i=1}^M \sigma_1(x, \zeta_{k,n}^i)^2 \right) \right),$$

where we denote by  $\text{Gsn}(\mu, \tau^2)$  the Gaussian distribution with mean  $\mu$  and variance  $\tau^2$ . Note that we implicitly assume that  $q = 1$ , in order to simplify notation.

(2) For  $k = 1, \dots, \lfloor \frac{1}{\Delta t} \rfloor - 1$ :

For all  $i = 1, \dots, M$ , set  $\zeta_{k,0}^i = \zeta_{k-1,n}^i$  and simulate  $\zeta_{k,n}^i$  from the forward Euler approximation to the transition kernel  $\mathbb{P}(Z_n(\xi_k) \in \cdot | Z_0(\xi_k) = \zeta_{k,0}^i)$  with step  $\delta t$ . Then, simulate  $\xi_{k+1}$  from

$$\xi_{k+1} \sim \text{Gsn} \left( \Delta t \left( \frac{1}{M} \sum_{i=1}^M a(\xi_k, \zeta_{k,n}^i) \right), \Delta t \left( \frac{1}{M} \sum_{i=1}^M \sigma_1(\xi_k, \zeta_{k,n}^i)^2 \right) \right).$$

(3) For  $k = \lfloor \frac{1}{\Delta t} \rfloor$ :

As in the previous step, set  $\zeta_{k,0}^i = \zeta_{k-1,n}^i$  and simulate  $\zeta_{k,n}^i$  from the forward Euler approximation to the transition kernel  $\mathbb{P}(Z_n(\xi_k) \in \cdot | Z_0(\xi_k) = \zeta_{k,0}^i)$  with step  $\delta t$ . Then, simulate  $\tilde{X}_1$  from

$$\tilde{X}_1 \sim \text{Gsn} \left( (1 - k\Delta t) \left( \frac{1}{M} \sum_{i=1}^M a(\xi_k, \zeta_{k,n}^i) \right), (1 - k\Delta t) \left( \frac{1}{M} \sum_{i=1}^M \sigma_1(\xi_k, \zeta_{k,n}^i)^2 \right) \right).$$

One can extend the weak convergence theorem in [3] for  $\sigma_1 \neq 0$ , to get an estimate of the approximation error of the transition kernel. More specifically,

$$|\mathbb{E}f(\tilde{X}_1) - \mathbb{E}f(\bar{X}_1)| \leq C_f \left( \Delta t + \delta t + \frac{e^{-\frac{1}{2}\lambda n}}{1 - e^{-\frac{1}{2}\lambda n}} (\Delta t + \Delta t^2) + \frac{\Delta t}{M} \right) \quad (17)$$

Let us now define a new particle filter, similar to the one in section 1, only the evolution of the particles between observation points follows the algorithm above, for  $\delta t = \Delta t = \frac{1}{\sqrt{N}}$  and  $n = M = 1$ . The choice  $M = 1$  might seem surprising at first, but actually gives optimal bounds (see [3], section 2.4). The reason is that the Monte-Carlo estimation is done by averaging both in time and independent realizations but averaging in time also improves the initialization error. So, it is in theory preferable to average one long path rather than many short ones.

Let us name this new particle filter  $\tilde{\pi}_k^N$ . Notice that for these values of  $\Delta t, \delta t, M, n$ , (17) becomes

$$|\mathbb{E}f(\tilde{X}_1) - \mathbb{E}f(\bar{X}_1)| \leq C'_f \frac{1}{\sqrt{N}}$$

Then, the total approximation error becomes

$$\mathbb{E} |\pi_k^\epsilon(f) - \tilde{\pi}_k^N(f)| \leq C \left( \epsilon + \frac{1}{\sqrt{N}} \right) \|f\|_\infty. \tag{18}$$

To study the efficiency of this particle filter, suppose that we want to achieve a total error of order  $\mathcal{O}(\epsilon)$ . Then, if we apply the particle filter algorithm of section 1 to the multiscale system, the number of simulations needed will be of order  $\mathcal{O}(\frac{1}{\epsilon^3}(p + q))$ : at each step, we simulate  $N\sqrt{N}(p + q)$  random variables – we need  $\sqrt{N}(p + q)$  simulations for the evolution of each particle and we have  $N$  particles – and we need  $N \sim \mathcal{O}(\frac{1}{\epsilon^4})$ , since the total error is given by (8).

On the other hand, if we approximate the optimal filter by  $\tilde{\pi}_k^N$ , we need  $N \sim \mathcal{O}(\frac{1}{\epsilon^2})$  to get a total error of order  $\mathcal{O}(\epsilon)$ . For this particle filter, the number of random variables we simulate at each step is  $N(\sqrt{N}p + Nq)$  –  $N$  is the number of particles and we simulate  $\sqrt{N}p + Nq$  random variables for the evolution of each particle. Notice that since  $M = 1$ , we estimate the drift and variance of (10) using the final value of only one path of the appropriate process  $Z_t(\cdot)$ . The reason why we discard the rest of the path is to allow the distribution of  $Z_t(\cdot)$  to get close to the invariant distribution of the process. Consequently, we need a total of  $\mathcal{O}(\frac{1}{\epsilon^3}p + \frac{1}{\epsilon^4}q)$  simulations, which shows that we can, indeed, achieve substantial improvement in the efficiency of the algorithm by replacing the multiscale system by the averaged diffusion, even when this not explicitly known.

**Remark 3.1.** *In order to approximate the drift and variance of the averaged process, we need to be able to simulate random variables from the invariant distributions  $\nu_x$ , for the appropriate  $x$ . We do that by simulating the process  $Z_t(x)$ , whose distribution converges exponentially fast to the invariant distribution  $\nu_x$ . Notice, however, that if  $x$  and  $x'$  are close, the distributions  $\nu_x$  and  $\nu_{x'}$  will also be close as a result of the smoothness of the drift and variance. Thus, we can improve the efficiency of the algorithm further by correlating the simulations of  $Z_t(x)$  and  $Z_t(x')$  as in [10] or by using the simulations of one process to initialize the other.*

#### 4. CONCLUSIONS

This analysis can also be applied for more general observation processes. For example, suppose that we observe  $(Y_k^\epsilon)_{k=1}^T$ , where  $Y_t^\epsilon$  is the solution of the following SDE:

$$dY_t^\epsilon = h(X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)}, Y_t^\epsilon)dt + \tau(X_t^{(1,\epsilon)}, X_t^{(2,\epsilon)}, Y_t^\epsilon)dV_t, \quad Y_0 = 0. \tag{19}$$

Then, we can replace (19) by its averaged approximation

$$d\bar{Y}_t = \bar{h}(\bar{X}_t, \bar{Y}_t)dt + \bar{\tau}(\bar{X}_t, \bar{Y}_t)dV_t, \quad Y_0 = 0, \tag{20}$$

for

$$\bar{h} = \int_{\mathbb{R}^q} h(x, z, y)\nu_x(dz) \quad \text{and} \quad \bar{\tau} = \left( \int_{\mathbb{R}^q} \tau(x, z, y)^2\nu_x(dz) \right)^{\frac{1}{2}}.$$

Notice that, by the averaging principle,  $(X_t^{\epsilon,1}, Y_t^\epsilon) \rightarrow (\bar{X}_t, \bar{Y}_t)$ , as  $\epsilon \rightarrow 0$ . Then, we can apply the particle filter described in [2] for this type of observation process and approximate  $\pi_k^\epsilon$  by

$$\bar{\pi}_k = \mathbb{P}(\bar{X}_k | \bar{Y}_\ell = Y_\ell^\epsilon, \ell = 1, \dots, k).$$

We expect that the efficiency of the algorithm will also be improved in this case.

In this paper, we introduced a particle filter for the estimation of a quantity  $(X_t^{(1,\epsilon)})$  that can be approximated by a diffusion given discrete and partial observations of it, in the case where this quantity is the slow component of a multiscale diffusion of the form (1). The main idea is that rather than evolving the particles by simulating the full multiscale system which can be very inefficient, it is better to do a short runs of of the full multiscale system and use these simulations to locally estimate the drift and variance of the diffusion that approximates

the evolution of the partially observed quantity. Depending on the multiscale system and the approximate diffusion, one can use different methods for the estimation of the diffusion parameters and the evolution of the particles that follow the diffusion, rather than the Monte-Carlo estimation and the Euler simulation discussed above.

## REFERENCES

- [1] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equation: I. Convergence rate of the distribution function, *Probab. Theory Relat. Fields* 104: 43-60, 1996.
- [2] P. Del Moral, J. Jacod, and P. Protter. The Monte-Carlo method for filtering with discrete-time observations, *Probab. Theory Relat. Fields* 120: 346-368, 2001.
- [3] Weinan E, D. Liu and E. Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.* 58(11): 1544-1585, 2005.
- [4] Weinan E, D. Liu and E. Vanden-Eijnden. Nested stochastic simulation algorithm for chemical kinetic systems with multiple time scales. *J. Comp. Phys.* (to appear).
- [5] P. Fearnhead, O. Papaspiliopoulos and G. O. Roberts. Particle filtering for diffusions avoiding time-discretisations. *Proceedings of NSSPW*, 2006.
- [6] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, 2nd edition, Springer-Verlag, 1998.
- [7] C. W. Gear, I. G. Kevrekidis and C. Theodoropoulos. "Coarse" integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods. *Comp. Chem. Engng.* 26: 941-963, 2002.
- [8] D. Givon, R. Kupferman and A. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity* 17: R55-R127, 2004.
- [9] A.J. Majda, I. Timofeyev and E. Vanden-Eijnden. A mathematical framework for stochastic climate models. *Comm. Pure App. Math.* 54: 891-974, 2001.
- [10] A. Papavasiliou and I. G. Kevrekidis. Variance Reduction for the Equation-free Simulation of Multiscale Stochastic Systems, *Multiscale Model. Simul.* (to appear).