• Discussion of "Exact and Computationally Efficient Likelihood-Based Estimation for Discretely Observed Diffusion Processes" by A. Beskos, O. Papaspiliopoulos, G.O. Roberts and P. Fearnhead.

• Mathias Rousset (University of Toulouse) and Arnaud Doucet (University of British Columbia).

• Arnaud Doucet
Department of Statistics,
333-6356 Agricultural Road,
University of British Columbia,
Vancouver, BC, V6T 1Z2,
Canada.
Email: arnaud@ubc.stat.ca

• The authors are to be congratulated for this impressive paper which solves many problems and opens many avenues of investigation. We present here a direct application of their methodology to time-discretization error-free filtering of partially observed diffusions. Let us consider the following diffusion where $X_0 \sim \pi_0$ and for $t > 0$

$$dX_t = \alpha(X_t)dt + dB_t.$$

This diffusion is partially observed at times $\{t_k\}_{k\geq 1}$ (where $t_k > t_{k-1}$) and the conditional density of the k-th observation Y_{t_k} given by $g_k(y_{t_k}|x_{t_k})$ is known analytically. We are interested in estimating sequentially the distributions

$$p(x_{t_0:t_k}|y_{t_1:t_k})dx_{t_0:t_k} = \mathbb{P}(X_{t_0:t_k} \in dx_{t_0:t_k}|Y_{t_1:t_k} = y_{t_1:t_k})$$

where $t_0 = 0$, $x_{t_0:t_k} = (x_{t_0}, x_{t_{2,\ldots,}} x_{t_k})$ and $y_{t_1:t_k} = (y_{t_1}, y_{t_{2,\ldots,}} y_{t_k})$. To achieve this we propose to use a Sequential Monte Carlo (SMC) [1]. The distributions are approximated by a large number N of weighted random samples $\left\{X^{(i)}_{t_0:t_k}\right\}$ $(i = 1, \ldots, N)$ termed particles. The particles are sampled using

$$X^{(i)}_{t_0:t_k} \sim q_k\left(x_{t_0:t_k}\middle| y_{t_1:t_k}\right)$$

with

$$q_k\left(x_{t_0:t_k}\middle| y_{t_1:t_k}\right) = \pi_0\left(x_{t_0}\right)\prod_{n=1}^{k} q_n(x_{t_n}|y_{t_n}, x_{t_{n-1}})$$

where $\left\{q_n(x_{t_n}|y_{t_n}, x_{t_{n-1}})\right\}$ are importance distributions known pointwise. In the standard SMC framework, these particles should be reweighted according to normalized weights proportional to

$$w_k\left(X^{(i)}_{t_0:t_k}\right) = \frac{p(X^{(i)}_{t_0:t_k}, y_{t_1:t_k})}{q_k\left(X^{(i)}_{t_0:t_k}\middle| y_{t_1:t_k}\right)} = w_{k-1}\left(X^{(i)}_{t_0:t_{k-1}}\right)\frac{\tilde{p}_{\Delta t_k}(X^{(i)}_{t_{k-1}}, X^{(i)}_{t_k})g(y_{t_k}|X^{(i)}_{t_k})}{q_k(X^{(i)}_{t_k})|y_{t_k}, X^{(i)}_{t_{k-1}})}$$

where

$$\begin{aligned}\tilde{p}_{\Delta t_k}(x_{t_{k-1}}, x_{t_k}) &= \mathbb{P}(X_{t_k} \in dx_{t_k} | X_{t_{k-1}} = x_{t_{k-1}})/dx_{t_{k-1}} \\ &= \mathcal{N}_{\Delta t_k}(x_{t_k} - x_{t_{k-1}}) \exp(A(x_{t_k}) - A(x_{t_{k-1}})) a(x_{t_{k-1}}, x_{t_k})\end{aligned}$$

with $A(u) = \int^u \alpha(z)\, dz$ and $a(x_{t_{k-1}}, x_{t_k}) = \mathbb{E}_{\mathbb{W}^{(x_{t_{k-1}}, x_{t_k})}}\big(\exp(-\frac{1}{2}\int_{t_{k-1}}^{t_k} (\alpha^2 + \alpha')(\omega_s) ds)\big)$. The particles are resampled whenever the variance of $\left\{ w_k \left(X_{t_0:t_k}^{(i)} \right) \right\}$ is too large. Clearly this SMC algorithm cannot be implemented as $\tilde{p}_{\Delta t_k}(x_{t_{k-1}}, x_{t_k})$ does not admit a closed-form expression. However a straightforward argument shows that it is not necessary to know $w_k \left(X_{t_0:t_k}^{(i)} \right)$ exactly. Only an unbiased positive estimate $\widehat{w}_k \left(X_{t_0:t_k}^{(i)} \right)$ of $w_k \left(X_{t_0:t_k}^{(i)} \right)$ is necessary to obtain asymptotically consistent SMC estimates under weak assumptions. Hence all the techniques developed by the authors to estimate $\tilde{p}_{\Delta t_k}(x_{t_{k-1}}, x_{t_k})$ unbiasedly can be applied straightforwardly. The need for positive estimates restricts us to diffusions similar to those of EA1 or EA2.

So as to be efficient, the SMC method requires to design "good" importance distributions and to obtain estimates of the importance weights with low variance. To design the importance distributions, a suggestion consists of approximating analytically $\tilde{p}_{\Delta t_k}(x_{t_{k-1}}, x_{t_k})$ by a Gaussian distribution using a local linearization technique [2] and combining this approximate prior with $g(y_{t_k}|x_{t_k})$ or a linearized version of it to obtain $q_k(x_{t_k}|y_{t_k}, x_{t_{k-1}})$.
It is also crucial to reduce the variance of the estimates of the weights given by

$$\mathrm{var}\left(\widehat{w}_k\left(X_{t_0:t_k}\right)\right) = \mathrm{var}\left(w_k\left(X_{t_0:t_k}\right)\right) + \mathbb{E}\left(\mathrm{var}\left(\widehat{w}_k\left(X_{t_0:t_k}\right)|X_{t_0:t_k}\right)\right).$$

To achieve this, if the Poisson estimator of section 6 is used, one could sample for each particle P Poisson random variables but use the same Brownian bridge to sample retrospectively for computational savings. However, a large Poisson parameter λ and a large P may be needed to obtain a reasonable variance.

References

[1] Doucet, A., de Freitas, N. and Gordon, N.J. (editors) (2001) *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag.

[2] Durham, G.B. and Gallant, A.R. (2002) Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes (with discussion). *J. Bus. Econ. Statist.*, **20**, 297-338.