# Sequential Monte Carlo Methods for Option Pricing

BY AJAY JASRA and PIERRE DEL MORAL

*Department of Mathematics, Imperial College London, SW7 2AZ, London, UK,*

`Ajay.Jasra@ic.ac.uk`

*Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux,*

*Université de Bordeaux I, 33405, France,,*

`Pierre.Del-Moral@inria.fr`

**Abstract**

In the following paper we provide a review and development of sequential Monte Carlo (SMC) methods for option pricing. SMC are a class of Monte Carlo-based algorithms, that are designed to approximate expectations w.r.t a sequence of related probability measures. These approaches have been used, successfully, for a wide class of applications in engineering, statistics, physics and operations research. SMC methods are highly suited to many option pricing problems and sensitivity/Greek calculations due to the nature of the sequential simulation. However, it is seldom the case that such ideas are explicitly used in the option pricing literature. This article provides an up-to date review of SMC methods, which are appropriate for option pricing. In addition, it is illustrated how a number of existing approaches for option pricing can be enhanced via SMC. Specifically, when pricing the arithmetic Asian option w.r.t a complex stochastic volatility model, it is shown that SMC methods provide additional strategies to improve estimation.

**Keywords:** Sequential Monte Carlo, Option pricing, Sensitivities

**Mathematics Subject Classification: 82C80, 60F99, 62F15**

**Short Title:** SMC for Option Pricing

# 1 Introduction

Monte Carlo methods have been extensively used in option pricing since the paper of [8]. Subsequently, there have been a wide variety of Monte Carlo approaches applied: quasi Monte Carlo (QMC), stratification, importance sampling (IS), control and antithetic variates, etc; see [33] for a thorough introduction.

The importance of Monte Carlo for option pricing, against other numerical approaches, is the ability to deal with high-dimensional integrals. This is either in the time parameter of the derivative (path dependent options), or in the dimension of the underlying. The rate of convergence is $O(1/\sqrt{N})$, $N$ being the number of simulated samples, supposedly, independent of the dimension. In addition, the methods are straight-forwardly extended to:

- complex stochastic volatility (SV) models (e.g. [4])

- complicated financial derivatives (e.g. [9]).

SV models are particularly useful to realistically replicate price dynamics. The latter point is relevant due to an increase in the volume traded of these instruments. Monte Carlo may also be used to calculate sensitivities/Greeks (e.g. [31]). As frequently noted in the option pricing literature, standard Monte Carlo estimates can suffer from a high level of variability, but can be improved using some of the methods mentioned in the above paragraph.

We have thus stated that Monte Carlo methods are an important tool for option pricing, but can suffer from high variability. An often used technique to deal with this problem is IS (e.g. [34]). As is well known, the idea is to change the dominating measure such that the resulting Monte Carlo estimation benefits from a lower variance. In many financial applications the simulation is sequential, that is, the underlying is sampled at discrete time-points. This often yields Radon-Nikodym derivatives that can be re-calculated sequentially in time. It is also well-known (e.g. [27, 33, 51]) that the variance of these weights increases with time; hence IS needs to be improved in order to yield accurate estimates of the option-

price.

This above problem can typically be solved using sequential Monte Carlo methods. This is a class of IS methods that are extensively used in engineering, statistics and physics. However, to our knowledge, these ideas are seldom used in the option pricing literature (see [17, 30, 44, 58, 66] for the few applications we were able to find). The purpose of this article is thus two-fold:

1. To provide an up-to date literature review of SMC methods, particularly focussed on option pricing and sensitivity analysis.

2. To illustrate that such methods can help to push the boundaries of the models for which option prices can be calculated, accurately.

In terms of 1, it seems that such methods are not well understood in the financial engineering literature, or at least the benefits of their application in option pricing is not appreciated. Therefore, we aim to review such methods and illustrate their use as well as improvement over standard approaches. Indeed, in some cases it is even possible to calculate option prices that are not subject to time discretization error [29]; this requires SMC methods. In relation to 2, an SMC algorithm is introduced to compute the value of arithmetic Asians, when the underlying is modelled by a stochastic volatility model. The volatility follows a non-Gaussian Ornstein-Uhlenbeck process [4]. This problem cannot be easily solved using ordinary IS and deterministic methods; see Section 4 for further details. Note, that it should not be seen that SMC methods are competitors to existing methods in option pricing, but simply that they enrich the methodology that can be used: in many cases SMC can be combined with existing ideas such as stratification (see e.g. [28] for some recent work).

It is remarked that the application of SMC methods can substantially reduce the variance of ordinary Monte Carlo and IS estimators; this at the cost of an increase the computational cost. In other words, the methods can be the most accurate in comparison to other approaches. Due to the above statements, the methods reviewed here in most cases could not be used at high frequency, but if solutions are required in minutes can be actively

used in finance (although see [50] for computational hardware that may make the methods even more applicable). Note, also the ideas differ from parametric IS, where the proposal lies in a parametric class, and is found to minimize some criterion (e.g. as in the cross entropy method see [63] for details). Since these methods are not SMC techniques, we do not review them here.

This article is structured as follows. In Section 2 we discuss an example to motivate the application of SMC methods for option pricing. In Section 3, SMC methods are detailed along with some of the latest developments. Illustrations are given on various examples to show that SMC methods can enhance existing Monte Carlo approaches in option pricing. In Section 4 an original SMC approach is designed for pricing Asian options, using the Barndorff-Nielsen & Shephard (BNS) SV model [4]. In Section 5 the article is concluded and some avenues for future research are discussed. There is an appendix which gives the proof of a result (in Section 3) and some details from the example in Section 4.

# 2   Motivating Example

## 2.1   Some Conventions

Some conventions are given. Recall that a Monte Carlo procedure simulates $N$ independent samples, $X^{(i)}$, from a density, $\pi$ and estimates $\pi-$integrable functions $h$ via

$$\frac{1}{N}\sum_{i=1}^{N}h(X^{(i)}).$$

Importance sampling follows much the same formula, except sampling from an $q \gg \pi$ and using the estimate

$$\sum_{i=1}^{N}w^{(i)}h(X^{(i)})$$

with $w^{(i)} = d\pi/dq(X^{(i)})$. The following notation is used. For any $(i,j) \in \mathbb{Z}^+$, $i \leq j$, $x_{i:j} := (x_i, \ldots, x_j)$. A process is written $\{X_t\}_{t\in[0,T]}$. A measurable space is denoted $(E, \mathcal{E})$. Given a sequence of spaces $E_0, \ldots, E_n$ (resp. $\sigma-$algebras $\mathcal{E}_0, \ldots, \mathcal{E}_n$) the product space is

written as $E_{[0,n]}$ (resp. product $\sigma-$ algebra $\mathcal{E}_{[0,n]}$). For a probability $\pi$ and $\pi-$integrable function $h$, the notation $\pi(h) := \int h(x)\pi(x)dx$ is sometimes used. The dirac measure on $\{x\}$ is written $\delta_x(dx')$. Probability densities are often assigned a standard notation $p$. Expectations are written generically as $\mathbb{E}$ and a subscript is added, if it is required to denote dependence upon a measure/point. Also, for $p \in \mathbb{Z}^+$, $\mathbb{T}_p = \{1, \ldots, p\}$.

## 2.2 Price Process

Throughout, the price process $\{S_t\}_{t\in[0,T]}$ $S_t \in \mathbb{R}^d$ follows a general jump-diffusion of the form

$$dS_t = \mu(S_t)dt + V_t\sigma(S_t)dZ_t$$

where $\{Z_t\}_{t\in[0,T]}$ is a Lévy process, and the volatility $\{V_t\}_{t\in[0,T]}$, may be deterministic, or may follow

$$dV_t = \alpha(V_t)dt + \beta(V_t)dU_t$$

where $\{U_t\}_{t\in[0,T]}$ is a Lévy process, that may be correlated with $\{Z_t\}_{t\in[0,T]}$. This permits a wide-class of stochastic volatility models, which can accurately replicate the stylized features of returns data (e.g. [4]). All expectations are taken w.r.t an equivalent, structure preserving, Martingale measure, $\mathbb{Q}$; this will exist for our examples.

## 2.3 Barrier Options

The first example is the calculation of barrier options. These are derivatives for which the payoff may be zero, dependent upon the path of the underlying $\{S_t\}_{t\in I}$, $I \subset [0,T]$, hitting a barrier. Our examples will concentrate upon European style options

$$\mathbb{E}_{S_0}[\Phi(\{S_t\}_{t\in I})]$$

with

$$\Phi(\{S_t\}_{t\in I}) = \mathbb{I}_A(\{S_t\}_{t\in I})e^{-rT}(S_T - K)_+$$

a barrier call option, with strike $K > 0$, interest rate $r > 0$ and $A$ the barrier set. It is assumed that the initial value of the underlying lies inside (resp. outside) for knock-out (resp. knock-in) options. For example, for a discretely monitored knock-out barrier option, $I = \{t_1, \ldots, t_m : 0 < t_1 < \cdots < t_m = T\}$, $t_0 = 0$, $d = 1$, barrier set $A = \bigotimes_{i=1}^{m}[a_{t_i}, b_{t_i}]$, the arbitrage price is:

$$(1) \qquad \int e^{-rT}(s_{t_m} - K)_+ \prod_{i=1}^{m} \left\{ \mathbb{I}_{[a_{t_i}, b_{t_i}]}(s_{t_i}) p(s_{t_i}, v_{t_i} | s_{t_{i-1}} v_{t_{i-1}}) \right\} p(v_0) d(s_{t_1:t_m}, v_{t_0:t_m})$$

where $v_{t_i}$ summarizes all the random variables induced by the volatility process (e.g. including integrated volatilities) and it is assumed that the possibly unknown transition densities can be written w.r.t a dominating measure $d(s_{t_1:t_m}, v_{t_0:t_m})$. There are a large number of publications associated to the Monte Carlo calculation of barrier options, including: [3, 35]; Zhao et al. (2006) and for the calculation of the Greeks [6, 31].

### 2.3.1    A Numerical Example

On inspection of (1), it is clear that the estimation of the barrier option is not a simple task in Monte Carlo integration. For example, even if we are able to assume that

- the transition densities are known

- it is possible to simulate from the process

or if an Euler discretization is adopted, it is still the case that many paths may yield a zero Monte Carlo estimate: they may knock-out before the terminal time. A simple remedy to this problem can be found in [35]. When the volatility is deterministic, the authors sample from the one-step process, conditioned so that it does not knock-out. This corresponds to an IS procedure with weights:

$$w = \prod_{i=1}^{m} \left\{ \int_{[a_i, b_i]} p(s_{t_i} | s_{t_{i-1}}, v_{t_{i-1}}) ds_{t_i} \right\}.$$

The weights need not be known, but an unbiased estimate of them is required; see Section 3.4.1, but the point is established in [35]. Note, as stated in [35], this idea only reduces the

variance of ordinary Monte Carlo in terms of knocking out; it may be sensible to design an IS strategy that takes into account the nature of the terminal cost.

The above idea, whilst very effective, is not always useful for large $m$ - the variance of the weights will increase with time. This is illustrated in Table 1, where a Black-Scholes model is adopted, with $d = 1$, $r = 0.01$, $\sigma = 0.75$, $K = S_0 = 10.0$, $t_i - t_{i-1} = 0.5$ and a time-homogenous barrier $A = [5.0, \infty)^m$; a down-and-out option. Monte Carlo methods are not needed here, but the idea is to show how SMC methods help; a more convincing example can be found in Section 4. The variance of the weights are approximated using the effective sample size (ESS, see [51]):

$$ESS = \frac{\left( \sum_{i=1}^N w^{(i)} \right)^2}{\sum_{i=1}^N \left( w^{(i)} \right)^2}.$$

If the simulated samples are dependent, it measures the number of samples that are independent. Here, it gives us a measure of the variability of the weights: the closer to $N$ ($= 30000$ here), the more the number of 'useful' samples. In Table 1, it is seen that the variance increases with $m$; this is well-known (e.g. the Theorem in [47]). In some cases, the algorithm can stabilize; the ideas in the subsequent Sections are still relevant, but are potentially less useful there.

It could be argued that any number of Monte Carlo enhancement methods could be used to deal with weight degeneracy; however, none of these approaches are explicitly designed to deal with this. However, SMC methods can deal with this problem; see Section 3.3.1, for some associated drawbacks. As noted in Chapter 7 of [33], the well-known likelihood ratio method [11, 32] for greek calculation, suffers from exactly this problem. As a result, SMC methods will be useful for a wide variety of option pricing and hedging problems.

| $m$ | 5 | 10 | 15 | 20 | 25 |
|-----|---|----|----|----|----|
| ESS | 21826.90 | 13389.60 | 8710.91 | 5909.51 | 4139.27 |

Table 1: Effective Sample Size for an IS Estimate of a down-and-out Barrier Option. We used a Black-Scholes model with $r = 0.01$, $\sigma = 0.75$, $a = 5.0$, $b = \infty$. 30000 samples were simulated.

# 3 Sequential Monte Carlo Methods

## 3.1 Sequential Importance Sampling

We now introduce a general methodology of sequential importance sampling (SIS), in a particular context. Sequential Monte Carlo techniques can be traced back to at least the 1950's in [40] and [60]; see [51] for a historical review. These ideas have been developed within statistics (e.g. [27]), physics (e.g. [41]), engineering (e.g. [38]) and operations research (e.g. [37]). There is no claim that the reference list is exhaustive, due to the volume of publications in this field. Much of this review focuses upon the statistical literature (with numerous exceptions), since it appears such ideas are not often used in the financial engineering literature.

The basic idea is as in ordinary IS, the only real difference being that calculations are performed sequentially. This sequential formulation allows us both to theoretically understand the algorithms as well as to derive more advanced versions. The simulation begins by simulating a collection of samples in parallel and the importance weights are computed in a sequential manner. These samples, in Section 3.2, will interact. In the statistics and engineering literature, it is typical to call the samples particles and this terminology is used interchangeably throughout the paper.

### 3.1.1 Formulation

Let $\{(E_n, \mathcal{E}_n)\}_{0 \leq n \leq m}$ be a sequence of measurable spaces. It is assumed that it is of interest to simulate from and compute expectations w.r.t a sequence of related probability measures

$\{\pi_n\}_{0 \le n \le p}$ on measurable spaces $(E_{[0,n]}, \mathcal{E}_{[0,n]})$, that is, a sequence of spaces of increasing dimension. Throughout the paper, $\{\pi_n\}_{0 \le n \le p}$ will often be referred to as 'targets'. Note, in some scenarios these targets can be artificial/arbitrary and provide the potential user with an extra degree of freedom to design the algorithm; this is illustrated in Section 3.1.2.

Introduce a sequence of probability measures of the following standard form, for $n \ge 0$:

$$Q_n(x_{0:n}) = M_n(x_n | x_{0:n-1}) Q_{n-1}(x_{0:n-1}) dx_{0:n}$$

with $M_n : E_{[0,n]} \to \mathbb{R}^+$ a probability kernel which we are able to simulate from, $Q_{-1}(\cdot) := 1$ and $M_0 = \eta_0 : E_0 \to \mathbb{R}^+$ a probability density on $E_0$.

The simulation idea, in Figure 1, is then essentially associated to the simple importance sampling identity:

$$
\text{(2)} \qquad \mathbb{E}_{\pi_n}[h_n(X_{0:n})] = \frac{\int_{E_{[0,n]}} h_n(x_{0:n}) \left\{ \prod_{i=0}^n W_i(x_{0:i}) \right\} Q_n(x_{0:n}) dx_{0:n}}{\int_{E_{[0,n]}} \left\{ \prod_{i=0}^n W_i(x_{0:i}) \right\} Q_n(x_{0:n}) dx_{0:n}}
$$

$$
\text{(3)} \qquad W_i(x_{0:i}) = \frac{\pi_i(x_{0:i})}{\pi_{i-1}(x_{0:i-1}) M_i(x_i | x_{0:i-1})}
$$

$(\pi_{-1}(\cdot) := 1)$ in the above equation (2) there is division by

$$\int_{E_{[0,n]}} \left\{ \prod_{i=0}^n W_i(x_{0:i}) \right\} Q_n(x_{0:n}) dx_{0:n}$$

to ensure that the incremental weights (3) need only be known point-wise up to a normalizing constant. The following biased (for finite $N$, but provably convergent as $N \to \infty$) estimate is employed

$$
\pi_n^N(h_n) = \sum_{l=1}^N h_n(x_{0:n}^{(l)}) w_n^{(l)}
$$

$$
\text{(4)} \qquad w_n^{(l)} = \frac{\prod_{i=0}^n W_i(x_{0:i}^{(l)})}{\sum_{j=1}^N \{ \prod_{i=0}^n W_i(x_{0:i}^{(j)}) \}}
$$

where $x_{0:n}^{(l)}$ is the $l^{th}$ sample at time $n$. From herein it is assumed, unless otherwise written, that the incremental weights are the un-normalized versions. The normalizing constant is (abusively) defined as

$$Z_n := \int_{E_{[0,n]}} \left\{ \prod_{i=0}^n W_i(x_{0:i}) \right\} Q_n(x_{0:n}) dx_{0:n}.$$

- 0. Set $n = 0$; for each $i \in \mathbb{T}_N$ sample $X_0^{(i)} \sim \eta_0$ and compute $W_0(x_0^{(i)})$.

- 1. Set $n = n + 1$, if $n = m + 1$ stop, else; for each $i \in \mathbb{T}_N$ sample $X_n^{(i)}|x_{0:n-1}^{(i)} \sim M_n(\cdot|x_{0:n-1}^{(i)})$, compute $W_n(x_{0:n}^{(i)})$ and return to the start of 1.

Figure 1: A Generic SIS Algorithm.

In order to select the $M_n$, a conditionally optimal density is [26]:

$$(5) \qquad M_n(x_n|x_{0:n-1}) = \pi_n(x_n|x_{0:n-1}).$$

The proposal is optimal, in terms of minimizing the variance of the incremental weights, conditional upon $x_{0:n-1}$ (see [25] for some limitations). Algorithms which can be expected to work well, will attempt to approximate this density; see [26] and the references there-in for details.

There are a large number of extensions of the SIS method. We list some references here: [14, 21, 25, 55]. In the paper [21], the algorithms are combined with Markov chain Monte Carlo (MCMC) methods; this is discussed later in Section 3.2.1.

### 3.1.2 The Barrier Option Revisited

In the case of the barrier option problem, in Section 2.3, if it is not possible to simulate from the process, but the transition densities are known, then we can introduce a process via $Q$. The formula (1) can be approximated using SIS, with the incremental weight

$$W_i(s_{t_{i-1}:t_i}, v_{t_{i-1}:t_i}) = \frac{p(s_{t_i}, v_{t_i}|s_{t_{i-1}}, v_{t_{i-1}})}{q(s_{t_i}, v_{t_i}|s_{t_{i-1}}, v_{t_{i-1}})}$$

$M_i = q$. Any discretely sampled option pricing problem can be written in this form. Indeed, in more generality, we may try to incorporate a terminal reward into the target densities, $\{\pi_n\}$. For example, the optimal importance density is

$$Q_m^{s_0}(v_0, s_{t_1:t_m}, v_{t_1:t_m}) \propto (s_{t_m} - K)_+ \left\{ \prod_{i=1}^m \mathbb{I}_{[a_{t_i}, b_{t_i}]}(s_{t_i}) p(s_{t_i}, v_{t_i}|s_{t_{i-1}} v_{t_{i-1}}) \right\} p(v_0).$$

Then, it is sensible to look for optimal sequences of proposals that approximate this (as in [35]). In theory, the sequence of densities

$$\pi_n^{s_0,1}(v_0, s_{t_1:t_n}, v_{t_1:t_n}) \quad \propto \quad \int (s_{t_m} - K)_+ \left\{ \prod_{i=1}^{m} \mathbb{I}_{[a_{t_i}, b_{t_i}]}(s_{t_i}) p(s_{t_i}, v_{t_i} | s_{t_{i-1}} v_{t_{i-1}}) \right\} \times$$
$$p(v_0) d(s_{t_{n+1}:t_m}, v_{t_{n+1}:t_m})$$

is a 'sensible' path to the optimal IS density; however, they cannot be computed. One strategy to circumvent this, in [42], is to introduce a monotonic transformation of the potential $(S - K)_+$, say $g$, at some time, close to the terminal time in the target density $\pi_n$:

$$\pi_n^{s_0,2}(v_0, s_{t_1:t_n}, v_{t_1:t_n}) \propto g((S_{t_n} - K)_+) \left\{ \prod_{i=1}^{n} \mathbb{I}_{[a_{t_i}, b_{t_i}]}(s_{t_i}) p(s_{t_i}, v_{t_i} | s_{t_{i-1}} v_{t_{i-1}}) \right\} p(v_0).$$

Note, in many cases $\pi_n^{s_0,1}(\cdot) \leq C_n \pi_n^{s_0,2}(\cdot)$ for $C_n \in (0, +\infty)$, $C_n \geq C_{n+1}$, therefore it is sensible to introduce the potential function at some time different than 1. In this case, the estimate

$$(6) \qquad\qquad \widehat{Z}_m \sum_{i=1}^{N} \frac{(s_{t_m}^{(i)} - K)_+}{g((s_{t_m}^{(i)} - K)_+)} w_m^{(i)}$$

can be used. It is described how to approximate the normalizing constant $Z_m$ below.

## 3.2   SIR

As we saw in Section 2.3.1, the SIS method will not always work as the time parameter increases. Before continuing, there are related methods in rare event simulation, termed multi-level splitting (e.g. [49]). These techniques are related to SMC as they are approximations of multi-level Feynman-Kac formulae [18, 12]; SMC algorithms are approximations of 'standard' Feynman-Kac formulae (as in [18]). Indeed, SMC algorithms related to splitting are given in [14]. Since most option pricing problems are not of rare-event form, we do not discuss the ideas of splitting any further; see [49] and the references therein for an introduction.

The formulation is as in Section 3.1.1; to simulate from a sequence of related probability measures on state-spaces of increasing dimension. The following resampling scheme is inserted into the SIS algorithm: At time $n$ of the algorithm $N$ particles are sampled, with replacement, from the current set of particles according to some stochastic rule with the property

(7)
$$\mathbb{E}[N_n^i | x_{0:n}^{(i)}] = N w_n^{(i)}$$

where $N_n^i$ is the number of replicates of the $i^{th}$ particle at time $n$. The most basic way this can be achieved is by resampling the particles according to the normalized weights $\{w_n^{(i)}\}_{1 \leq i \leq N}$. There are a variety of superior methods for resampling the particles: residual, stratified, systematic etc; a full description can be found in [27]. The systematic method is used here. This sequential importance sampling with resampling, is termed sequential importance sampling/resampling (SIR) (see [18, 19, 27]). The simulated paths are no longer independent, but there is a well-established convergence theory; see [18]. In addition, there is a theoretical advantage to resampling the particles; the asymptotic variance in the central limit theorem [18] can be upper-bounded, uniformly in time, with resampling; this is not necessarily the case otherwise - see [15].

The algorithm is given in Figure 2. The particles can be resampled at any time step of the algorithm, but it is best to do this when the weights are very variable; e.g. when the ESS drops below a pre-specified threshold. This is theoretically valid, as established in [20]. The reason for the above is that if the particles are resampled too often, then there are too many sampled paths which have been replicated - the paths degenerate; see Section 3.3.1 for further details. In addition, if the optimal proposal (5) can be used, then it is best to sample after the resampling operation has occurred. This will increase the number of unique samples and lower the variance in estimation.

To estimate the normalizing constants, use

$$\widehat{Z_n} = \prod_{j=1}^{r_{n-1}} \frac{\widehat{Z_{k_j}}}{Z_{k_{j-1}}},$$

- 0. Set $n = 0$; for each $i \in \mathbb{T}_N$ sample $X_0^{(i)} \sim \eta_0$ and compute $W_0(x_0^{(i)})$.

- 1. Decide whether or not to resample, and if this is performed, set all weights to $w_n^{(i)} = 1/N$ and proceed to step 2. If no resampling occurs, the weights are as in equation (4).

- 2. Set $n = n + 1$, if $n = m + 1$ stop, else; for each $i \in \mathbb{T}_N$ sample $X_n^{(i)} | x_{0:n-1}^{(i)} \sim M_n(\cdot | x_{0:n-1}^{(i)})$, compute $W_n(x_{0:n}^{(i)})$ and return to the start of 1.

Figure 2: A Generic SIR Algorithm.

with

$$(8) \qquad \widehat{\frac{Z_{k_r}}{Z_{k_{r-1}}}} = \frac{1}{N} \sum_{i=1}^{N} \prod_{j=k_{r-1}+1}^{k_r} W_j\left(x_{0:j}^{(i)}\right)$$

where $k_0 = 0$, $k_j$ is the $j^{\text{th}}$ time index at which one resamples for $j > 1$. The number of resampling steps between 0 and $n - 1$ is denoted $r_{n-1}$ and we set $k_{r_n} = n$ (to ensure that the final term includes $Z_n$ in the numerator, as is required for correctness).

### 3.2.1 SMC Samplers

A useful algorithm related to SIR is termed SMC samplers [21]. This method is designed to sample from a sequence of target distributions on a common space; i.e. $\{\pi_n\}_{0 \leq n \leq p}$ are probabilities on $(E, \mathcal{E})$. However, due to a difficulty described below, the algorithm samples from a sequence of distributions of increasing dimension. The marginal of each new density is the one of interest.

Suppose we initialize an IS algorithm with $N$ particles $\{X_0^{(i)}\}_{1 \leq i \leq N}$ sampled according to some initial density, $\eta_0$, and weight w.r.t $\pi_0$. Further, suppose that the IS works well; this can be achieved by making $\pi_0$ very simple, e.g. $\eta_0 = \pi_0$. Now, under the assumption that $\pi_0$ and $\pi_1$ are not significantly different, it might be expected that we can construct a Markov kernel $K_1(x_0, x_1)$ so as to move the particles from regions of high density of $\pi_0$ to

the corresponding regions of $\pi_1$. In such a case, the importance weight would be:

$$(9) \qquad w_1(x_1^{(i)}) = \frac{\pi_1(x_1^{(i)})}{\int_E \eta_0(x_0) K_1(x_0, x_1^{(i)}) dx_0}.$$

Equation (9) presents some major difficulties:

- In many cases the integral in the denominator (9) cannot be computed.

- In some cases, $K_1$ is not known point-wise.

For the above reasons, it appears that such a weighting scheme seems destined to fail. For example, even when there is the computational power to approximate the integral in the denominator of (9):

$$\frac{1}{N} \sum_{j=1}^{N} \eta_0(x_0^{(j)}) K_1(x_0^{(j)}, x_1^{(i)})$$

the second point will prohibit its application.

The solution is to modify the problem to a more familiar setting. Introduce a sequence of auxiliary probability measures $\{\widetilde{\pi}_n\}_{0 \leq n \leq p}$ on state-spaces of increasing dimension $(E_{[0,n]}, \mathcal{E}_{[0,n]})$, such that they admit the $\{\pi_n\}_{0 \leq n \leq p}$ as marginals. The following sequence of auxiliary densities is used (see [41] and [52]):

$$(10) \qquad \widetilde{\pi}_n(x_{0:n}) = \pi_n(x_n) \prod_{j=0}^{n-1} L_j(x_{j+1}, x_j)$$

where $\{L_n\}_{0 \leq n \leq p-1}$ are a sequence of Markov kernels that act backward in time and are termed backward Markov kernels. The algorithm samples forward using kernels $\{K_n\}$. The choice of backward kernels is made as the incremental weights are

$$(11) \qquad W_n(x_{n-1:n}) = \frac{\pi_n(x_n) L_{n-1}(x_n, x_{n-1})}{\pi_{n-1}(x_{n-1}) K_n(x_{n-1}, x_n)} \quad n \geq 1$$

which allows for a fast computation and avoids a path degeneracy effect (see Section 3.3.1). It is clear that (10) admit the $\{\pi_n\}$ as marginals, and hence, if one sequentially samples from $\{\widetilde{\pi}_n\}$ we are left with a problem that is the same as for SIR. That is, one uses the algorithm in Figure 2 with the kernel $K_n$ as a proposal (instead of $M_n$) and incremental weight (11). A discussion of the optimal choice of backward kernels can be found in [21].

### 3.2.2 The Barrier Option Revisited

Two examples of SIR algorithms are now presented. The first, is a straightforward modification of the SIS method demonstrated in Section 2.3.1. In this case the target densities are

$$\pi_n^{s_0}(v_0, s_{t_1:t_m}) \propto \left\{ \prod_{i=1}^{n} \mathbb{I}_{[a_{t_i}, b_{t_i}]}(s_{t_i}) p(s_{t_i} | s_{t_{i-1}}) \right\},$$

with importance densities as in Section 2.3.1. As a result, this approach is as in [35], except with resampling. The second idea is as in [42]: include the function $(S - K)$ in the target. This is to reduce the variance in estimation, when considering the call option. The function used is

$$g(S - K) = |(S - K)|^\kappa,$$

$\kappa$ is referred to as a temperature parameter. See [42] for a justification of this choice of $g$.



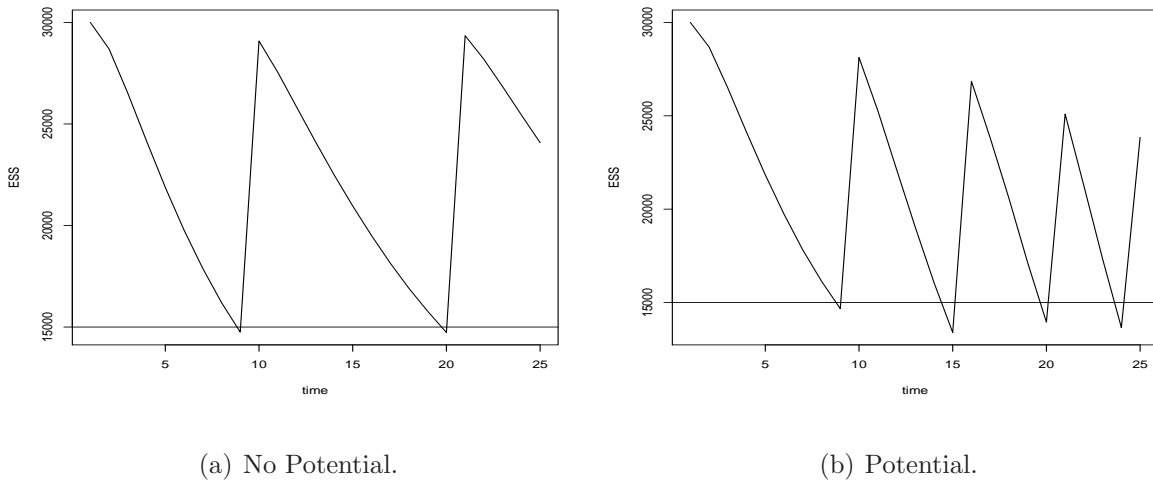(a) No Potential.

(b) Potential.

Figure 3: Controlling the Variance of the Weights. We used an SIR algorithm to apply the method of [35] to a simple Black-Scholes model (left). In the right plot we used the approach of [42]. The settings are as in Section 2.3.1 and there are $m = 25$ time steps of the algorithm and 30000 particles are simulated.

In Figure 3 (a) the performance of the first SIR method can be seen. In the Figure, we have used the Black-Scholes model with $m = 25$ and $N = 30000$ particles; the particles are

resampled when the ESS drops below 15000 particles. The algorithm exhibits very stable variance properties, resampling only 2 times. As a result, and in comparison to the results of Table 1, the algorithm is much more useful for the estimation of the down-and-out barrier option. Note that there is a marginal increase of the overall CPU time. The CPU time is less than 10 seconds on a Pentium 4 2.4 GhZ machine, coded in C++; all code is available upon request from the author.

In Figure 3 (b) the performance of the second SIR algorithm can be seen. In this case the potential is introduced at time 10, with $\kappa = 0.08$. At subsequent time steps, this value increases by 0.045. These values are set by some prior tuning, but it is possible to do this adaptively; see [43]. The issues here are:

- When to introduce the potential

- How fast the temperature should be increased.

In general, the potential could be introduced quite early in the simulation. This would allow the samples to adapt to the potential, but at the cost of increasing the variability of the weights. In practice, we have used between one-third and two-thirds of the overall algorithm time-length. The second issue is also important: to reduce the variability of (6), the temperature should be (at the final time-step) bigger than 1. However, if the temperature increases too rapidly, then algorithm has to resample too often. The stability of the weights is not as good as in Figure 3 (a). This is due to the introduction of the potential function. However, let us consider the actual quality in estimation of the down-and-out option, a significant improvement can be seen. Both approaches are run 25 times, with 30000 particles, resampling threshold of 15000, as well as computing the analytic approximation in [10]; the estimates of the options ($\pm 2$ standard deviations across the repeats) are $76.81 \pm 13.81$ and $6.03 \pm 0.43$, with analytic approximated value of 6.16. This illustrates a well-known fact about IS: the samples have to be pushed towards regions of importance, in terms of the functional of interest. The monitoring of the variance of the weights may not be enough to yield sensible estimates for option pricing.

## 3.3   SMC Methods for Sensitivities

We now review some SMC methods for estimating derivatives of expectations. This can be useful for calculating the Greeks, especially when the transition densities are not known and Euler approximations are adopted via the likelihood ratio method. There is a growing literature on Malliavin calculus approaches [6, 31], even in the case of quite sophisticated SV models [5]; these problems could be dealt with using the SMC approaches in the other Sections. Indeed the rates of convergence, in comparison to the likelihood ratio method, are faster [24], although in many cases the Malliavin weight is not known.

### 3.3.1   Path Degeneracy

In option pricing, most Greeks which are not calculated through Malliavin methods are of the form

$$(12) \qquad \Lambda(\Phi, s_0) = \nabla \left[ \int \Phi(s_{t_1:t_m}) \left\{ \prod_{i=1}^{m} p(s_{t_i}, v_{t_i} | s_{t_{i-1}}, v_{t_{i-1}}) \right\} p(v_0) d(s_{t_1:t_m}, v_{t_0:t_m}) \right].$$

Monte Carlo methods for computing this quantity can be found in Chapter 7 of [33], see also [2]. It is explicitly assumed that the transition density is known and upper bounded, and that the derivatives of this quantity can be calculated analytically.

One of the problems of estimating (12) using IS methods, is that as seen in Section 3.2.2, resampling is often required to control the variance of the weights. However, a key problem for estimating (12) is that the integral is on a path space. For example, approaches which calculate expectations (option prices) can use the same set of particles to estimate the sensitivities (Greeks) [13]. The difficulty is the following: the resampling mechanism, whilst necessary to control the variance of the weights, induces a path degeneracy effect. There may be many unique particles at the current time, but going backward in time yields paths which coalesce to the same parents. That is to say, that $x_{0:n-1}$ are not re-simulated at time $n$. Hence, due to resampling, many of the $x_{0:n-L+1}$ ($L > 0$) will be the same across the simulated samples. In other words SMC algorithms are only useful for calculating

expectations of the form:

$$\int_{E_{[n-L,n]}} h(x_{n-L:n})\pi_n(x_{n-L:n})dx_{n-L:n}$$

for some *fixed* lag $L > 0$. See [1] and [18] for more details. It should be noted that it will still be possible to compute the sensitivities of a wide-class of option prices.

### 3.3.2 Marginal Approximations

In this Section it is shown how to estimate the Greeks, when using the same set of particles to estimate the option price. The method can only be used if there is a Markovian dependence of the $(S_{t_n}, V_{t_n}) \in E_n$. For simplicity $v_0$ is dropped, but can straight-forwardly be re-introduced.

The idea is associated to the likelihood ratio method; assume the pay-off function is of the form

$$\Phi(S_{t_1}, \dots, S_{t_m}) = \prod_{i=1}^{m} \widetilde{\Phi}(S_{t_i})$$

note that there are wide class of options that can be written this way. Assume that only the process depends upon a parameter $\theta \in \Theta$. Then the sensitivity is, assuming the validity of interchanging integration and differentiation (e.g. [48])

$$\Lambda_\theta(\Phi, s_0) = \int \left\{ \prod_{i=1}^{m} \widetilde{\Phi}(S_{t_i}) \right\} \frac{\partial}{\partial \theta} \left\{ \prod_{i=1}^{m} p_\theta(s_{t_i}, v_{t_i}|s_{t_{i-1}}, v_{t_{i-1}}) \right\} d(s_{t_1:t_m}, v_{t_1:t_m})$$

where $\theta$ is the parameter of interest. To obtain the sensitivity of interest we present the following result, whose proof is housed in the appendix.

**Proposition 1.** *Let, for $n \geq 2$*

(13)
$$\widetilde{\Lambda}_n(dx_n) := \left[ \int_{E_{n-1}} \widetilde{\Lambda}_{n-1}(dx_{n-1})p_\theta(x_n|x_{n-1})\widetilde{\Phi}(s_{t_n}) \right] dx_n +$$
$$\left[ \int_{E_{n-1}} \Pi_{n-1}(dx_{n-1})\frac{\partial}{\partial \theta}\{p_\theta(x_n|x_{n-1})\}\widetilde{\Phi}(s_{t_n}) \right] dx_n$$

*where*

$$\Pi_n(dx_n) = \left[ \int_{E_{n-1}} \Pi_{n-1}(dx_{n-1})p_\theta(x_n|x_{n-1})\widetilde{\Phi}(s_{t_n}) \right] dx_n$$

$x_n = (s_{t_n}, v_{t_n})$ and

$$\widetilde{\Lambda}_1(dx_1) = \left[\frac{\partial}{\partial\theta}\{p_\theta(x_1|x_0)\}\widetilde{\Phi}(s_{t_1})\right]dx_1$$

$$\Pi_1(dx_1) = \left[p_\theta(x_1|x_0)\widetilde{\Phi}(s_{t_1})\right]dx_1.$$

*Then,* $\widetilde{\Lambda}_m(1) = \Lambda_\theta(\Phi, s_0).$

In the case of the Greeks $\Delta$ and $\Gamma$, the second expression on the RHS (13) is not needed, due to the dependence of only the first term in the product $\prod_{i=1}^{n-1} p_\theta(x_i|x_{i-1})$ on $s_0$.

The approach below is a special case of that detailed in [56] (see also [57]). The objective is to recursively approximate the measures $\{\widetilde{\Lambda}_n\}$ and $\{\Pi_n\}$. The initialization can be achieved by IS. Then, given an empirical approximation at time $n-1$ of $\Lambda_{n-1}$, say:

$$\sum_{i=1}^{N} W_{n-1}^{(i)}\delta_{x_{n-1}^{(i)}}(dx_{n-1})$$

and of $\Pi_{n-1}$, with different weights $\widetilde{W}_{n-1}$, then the following is a point-wise approximation of the signed density

$$\widetilde{\Lambda}_n^N(x_n) = \widetilde{\Phi}(s_{t_n})\left[\sum_{i=1}^{N} W_{n-1}^{(i)}p_\theta(x_n|x_{n-1}^{(i)}) + \sum_{i=1}^{N}\widetilde{W}_{n-1}^{(i)}\frac{\partial}{\partial\theta}\{p_\theta(x_n|x_{n-1}^{(i)})\}\right].$$

Assuming that new particles $\left\{X_n^{(i)}\right\}$ are sampled from some density $\Psi_n^N(x_n)$, then using IS, the following approximation is obtained

$$\widetilde{\Lambda}_n^N(dx_n) = \sum_{i=1}^{N} W_n^{(i)}\delta_{x_n^{(i)}}(dx_n)$$

where

(14)
$$W_n^{(i)} = \frac{\widetilde{\Lambda}_n^N(x_n^{(i)})}{N\Psi_n^N(x_n^{(i)})}.$$

Note

$$\widetilde{W}_n^{(i)} = \frac{\sum_{j=1}^{N}\widetilde{W}_{n-1}^{(j)}p_\theta(x_n^{(i)}|x_{n-1}^{(j)})}{N\Psi_n^N(x_n^{(i)})}.$$

The computational complexity of computing the weights is $O(N^2)$. This is required, to avoid the path degeneracy problem. It should be noted that an asymptotically biased approach, which is of $O(N)$ could, in some cases, be used from [53], with the identity

$$\frac{\partial}{\partial \theta} \int h(x)\pi_\theta(x)dx = \int h(x)\frac{\partial}{\partial \theta} \log\{\pi_\theta(x)\}\pi_\theta(x)dx.$$

See [53] for further details. We remark, also, that the work of [22, 23] can also be used, in some contexts, for greek calculation, but do not review these new techniques here.

### 3.3.3  Barrier Option Revisited

We return to the Example in Section 2.3.1. We calculate the $\Delta$ and the $\mathcal{V}$ using the recursion in (13). The model settings were as in Section 2.3.1. The algorithm is run with 10000 particles, 25 times which took 30 hours; i.e. 1 hour 20 min to run the algorithm once. The $\Delta$ was estimated as $0.22 \pm 1.14$ and the $\mathcal{V}$ was $-29.7 \pm 138.34$. Note that the error is $\pm$ 2 standard deviations for the 25 repeats. The ratio of the estimate to the $\pm$ 2 standard deviations are approximately 1/5 and perhaps too high; reducing the variability is the subject of current research. The CPU time is not substantial and shows that the given recursions are potentially useful for at computing $\Delta$ and $\Gamma$, but other path-based sensitivities may require biased estimates as in [53].

## 3.4  SMC Methods for Continuous-Time Processes

In the context of continuous-time processes, there are discrete-time SMC algorithms which can approximate the continuous-time Feynman-Kac formulae with no approximation-bias (see also [61]).

### 3.4.1  The Random Weight

An important comment, as realized by [62], on IS is the following. It is not necessary to know the weight exactly; only to find an unbiased estimate of it. More exactly, let $n = 0$

for simplicity. Suppose it is possible to find some function $\vartheta_n : E_n \times F_n \rightarrow \mathbb{R}^+$ and a probability $\rho_n$ so that:

$$w_n(x) = \mathbb{E}_{\rho_n(x,\cdot)}[\vartheta_n(x,U)].$$

then

$$
\begin{aligned}
\pi_n(h_n) &= \frac{1}{Z_n} \int_{E_n} h_n(x) w_n(x) \eta_n(x) dx \\
&= \frac{1}{Z_n} \int_{E_n} f_n(x) \left\{ \int_{F_n} \vartheta_n(x,u) \rho_n(x,u) du \right\} \eta_n(x) dx.
\end{aligned}
$$

As a result, we may perform joint IS by sampling from $\rho_n \times \eta_n$. This comment is extremely useful for option pricing; it has already adopted, by at least [35]. Moreover, it can be useful for calculating expectations w.r.t probability laws of Lévy processes.

### 3.4.2 Barrier Option Revisited

The ideas here are based upon the random weight, and the exact simulation of diffusions methodology [7, 29]. In many cases, the option price can be written as an expectation w.r.t the transition densities, but that the densities are not available up-to a constant or cannot be simulated. However, it may be that transition laws are known indirectly e.g. in terms of their characteristic function; see [36] for some solutions.

Consider (1) with deterministic volatility and the underlying is a diffusion satisfying the SDE

$$dS_t = \xi(S_t)dt + dW_t$$

with $\xi = \nabla \Xi$ and some additional assumptions in [29]. Then the transition is

$$
\begin{aligned}
p(s_{t_i}|s_{t_{i-1}}) &= \phi(s_{t_i}; s_{t_{i-1}}, (t_i - t_{i-1})) \exp\left\{\Xi(S_{t_i}) - \Xi(S_{t_{i-1}} - l(t_i - t_{i-1}))\right\} \times \\
&\quad \mathbb{E}_{\mathbb{W}_{s_{t_{i-1}}}^{s_{t_{i-1}}}} \left[ \exp\left\{ -\int_{[0,T]} \zeta(W_s) ds \right\} \right]
\end{aligned}
$$
(15)

where the expectation is taken w.r.t the law of a Brownian bridge starting at $s_{t_{i-1}}$, ending at $s_{t_i}$ and the other parameters/functions are known exactly; see [29] for details. Therefore,

we can use an SIR algorithm with an incremental weight

$$\frac{p(s_{t_i}|s_{t_{i-1}})}{q(s_{t_i}|s_{t_{i-1}})}$$

that is not known exactly. However, it is detailed in both [7] and [29] how the expectation in (15) can be estimated unbiasedly; the random weight idea can be applied. In our experience the estimation of

$$\mathbb{E}_{\mathbb{W}_{s_{t_{i-1}}}^{s_{t_{i-1}}}}\left[\exp\left\{-\int_{[0,T]}\zeta(W_s)ds\right\}\right]$$

through the methodology in [29], can lead to substantial increases in variance of the weights. In high-dimensional cases it may be difficult to use this idea, without some extra variance reduction methods. A similar idea can be found in [61] when approximating

$$\mathbb{E}_x^{\mathbb{W}}\left[g(W_T)\exp\left\{\int_{[0,T]}U(W_s)ds\right\}\right].$$

See [54] for a review of related issues.

# 4 An SMC Method for Pricing Asian Options

In the following Section we present a method to approximate the value of a fixed strike, arithmetic Asian option, when the underlying follows the system

$$dY_t = \mu dt + V_t dW_t$$
$$dV_t = -V_t dt + dU_t$$

where $Y_t$ is the log price, $\{W_t\}_{t\in[0,T]}$ is a Brownian motion and $\{U_t\}_{t\in[0,T]}$ is a pure-jumps Lévy process, such that the marginal of $V_t$ is a Gamma distribution; see [4] and the appendix for details. This latter model has been shown to fit, to an extent, the dynamics of real price data.

This problem is particularly difficult for Monte Carlo methods, even in the Black-Scholes case (see, however, [45] for some work on the exact approximation with continuously monitored average). The reason is due to the path-based nature of the payoff, which can

contribute a substantial variance in the estimate. Indeed, given the methodology explained thus far, it is not straightforward to use SMC methods for exactly the previous reason; see Section 3.3.1. Some sensible solutions (to the form of the pay-off), in the Black-Scholes scenario, can be found in [34] and [46]. Also, note that efficient techniques for partial differential equations (PDEs) can be found in [64, 65]. These methods appear to be adaptable to our case, but would lead to solving a partial integro-differential equation which is not always simple and accurate; see [16] for some ideas.

## 4.1 Strategy

The idea is to try, as in the PDE methods above, to reduce the dimension of the problem. In essence, we seek to sample from the law of the sum of the underlying at each monitoring period. The problem is that this is only known in terms of an intractable integral. However, it is shown that it is possible, using SMC methods to sample from the optimal importance density, such that it is a marginal on an extended state-space.

In our case, ignoring the discount term, the value of the option is

$$\int \left( \frac{1}{m} \sum_{i=1}^{m} e^{y_{t_i}} - K \right)_{+} \left\{ \prod_{i=1}^{m} p(y_{t_i}|y_{t_{i-1}}, v_{1:t_i}) p(v_{t_i}) \right\} d(y_{t_1:t_m}, v_{t_1:t_m})$$

where $v_{t_0}$ is known and suppressed from the notation and $v_{t_i}$ is a 2-dimensional Poisson process on $[0, \lambda \nu(t_i - t_{i-1})] \times [0, 1]$. Making the transform, $s_{t_i} = \exp\{y_{t_i}\}$, and then

$$S_{t_i} = \nu_{t_i} - \nu_{t_{i-1}} \quad i \geq 2$$
$$S_{t_1} = \nu_{t_1}$$

yields the option price

$$\int \left( \frac{\nu_{t_m}}{m} - K \right)_{+} \left\{ \prod_{i=1}^{m} \varphi_{\nu_{t_{i-1}}} \{\nu_{t_i}; \tilde{\mu}(\nu_{t_{i-2}:t_{i-1}}), \tilde{\sigma}_i(v_{1:t_i})\} p(v_{t_i}) \right\} d(\nu_{t_1:t_m}, v_{t_1:t_m})$$

where $\varphi_{\nu_{t_{i-1}}} \{\nu_{t_i}; \tilde{\mu}(\nu_{t_{i-2}:t_{i-1}}), \sigma(v_{1:t_i})\}$ is the shifted log-normal density with location parameter

$$\tilde{\mu}(\nu_{t_{i-2}:t_{i-1}}) = \log(\nu_{t_{i-1}} - \nu_{t_{i-2}}) + \mu(t_i - t_{i-1})$$

scale parameter as in [4], $\nu_{t_{-1}} = 0$, $\nu_{t_0} = S_0$ and see the Appendix for the definition of $\tilde{\sigma}$. If there was not a path dependence (see Section 3.3.1) on the integrated volatilities $v_{t_1:t_m}$, this transformation would greatly improve SMC methods. This is because one needs only estimate an integral w.r.t the marginal $\nu_{t_m}$. However, the problem is recast into a similar case as in Section 3.2.2; as a result the method detailed below, is restricted to the case where $m$ is not large ($m \le 24$).

## 4.2   Simulations

### 4.2.1   SMC Method

The SMC approach is similar to the second method in Section 3.2.2. The difference is as follows. In prior simulations it was found the temperature $\kappa$ could not be too large ($\kappa < 0.5$) to avoid a substantial weight degeneracy effect. This is when introducing the potential in the middle of the simulation. If the potential function was introduced very early, the path degeneracy effect occurred; despite the fact that the temperature is high, which can reduce the variance in estimation, the algorithm resampled too often to yield poor estimates.

To allow the temperature parameter to reach 1, an SMC sampler was adopted. That is, once time $m$ is reached the SMC sampler simulated from

$$\pi_n(\nu_{t_1:t_m}, v_{t_1:t_m}) \propto \left| \frac{\nu_{t_m}}{m} - K \right|^{\tilde{\kappa}_n} \left\{ \prod_{i=1}^{m} \varphi_{\nu_{t_{i-1}}} \{\nu_{t_i}; \tilde{\mu}(\nu_{t_{i-2}:t_{i-1}}), \tilde{\sigma}_i(v_{1:t_i})\} p(v_{t_i}) \right\}$$

for $\kappa < \tilde{\kappa}_1 < \cdots < \tilde{\kappa}_p = 1$. Here $\kappa$ is the final temperature of the SIR algorithm. In all cases, $p = 20$ and the $\kappa$'s increase at a constant amount; see [42] for automatic schemes. The initial SIR is used for computational efficiency; it is less costly to run this algorithm than SMC samplers. Note that the estimate of the option price is as in (6).

For SMC samplers, the particles are simulated according to an MCMC kernel, $K_n$, of invariant distribution $\pi_n$; details are given in the appendix. The backward kernel used in

- 0. Use an SIR algorithm (as in Figure 2) to sample from the sequence of densities:

$$\pi_n(\nu_{t_1:t_n}, v_{t_1:t_n}) \propto \left| \frac{\nu_{t_n}}{m} - K \right|^{\kappa_n} \left\{ \prod_{i=1}^{n} \varphi_{\nu_{t_{i-1}}} \{\nu_{t_i}; \tilde{\mu}(\nu_{t_{i-2}:t_{i-1}}), \tilde{\sigma}_i(v_{1:t_i})\} p(v_{t_i}) \right\}$$

  with $n \in \mathbb{T}_m$, $\{\kappa_n\}_{1 \leq n \leq m}$ given $(0 \leq \kappa_1 < \cdots < \kappa_m < 1)$. The process densities are used as the proposals.

- 1. Given the samples from $\pi_n(\nu_{t_1:t_m}, v_{t_1:t_m})$ use SMC samplers (as in Section 3.2.1) to sample from

$$\pi_n(\nu_{t_1:t_m}, v_{t_1:t_m}) \propto \left| \frac{\nu_{t_m}}{m} - K \right|^{\tilde{\kappa}_n} \left\{ \prod_{i=1}^{m} \varphi_{\nu_{t_{i-1}}} \{\nu_{t_i}; \tilde{\mu}(\nu_{t_{i-2}:t_{i-1}}), \tilde{\sigma}_i(v_{1:t_i})\} p(v_{t_i}) \right\}$$

  with $n \in \mathbb{T}_p$, $\kappa_m = \kappa < \tilde{\kappa}_1 < \cdots < \tilde{\kappa}_p = 1$. Note that this is a sequence of densities on the same space, as opposed to the sequence in step 0. which samples from one of increasing dimension. The SMC sampler uses the MCMC algorithm in the appendix and has incremental weight (16).

Figure 4: SMC Method to Approximate Arithmetic Asian Options.

the SMC sampler (see Section 3.2.1), is

$$L_{n-1}(x_n, x_{n-1}) = \frac{\pi_n(x_{n-1}) K_n(x_{n-1}, x_n)}{\pi_n(x_n)}$$

where $x_n = (\nu_{t_n}, v_{t_n})$, is such that the incremental weights at time $n$ are

$$(16) \qquad W_n(x_{n-1}) = \frac{\pi_n(x_{n-1})}{\pi_{n-1}(x_{n-1})}.$$

The choice of backward kernel is quite popular in the literature and, in our experience, often works well in practice. The procedure is summarized in Figure 4.

### 4.2.2 Importance Sampling Method

To compare the method above, the approach of [34] is modified. The main difference is to simulate the integrated volatility exactly, and to use the smallest value (over each interval)

in the numerical search for a change of drift. In this scenario, it was found that there was not always a unique root; this is in less than 1% of the cases in the results below.

### 4.2.3 Numerical Results

The algorithms are run 50 times, for similar CPU time; this was less than 150 seconds. $m = 12$, $\lambda = 1$, $t_i - t_{i-1} = 1$ ($\forall i \in \mathbb{T}_m$), $\mu = 0.07$, $\nu = 0.5$; these values, when relevant, are consistent with the parameter estimates that were reported in [39]. The SMC method was run with 10000 samples, and in the SIR part of the algorithm the potential was introduced at time 6, with $\kappa = 0.2$; this increased by 0.035 at each time step, until time 12 was reached. The process densities were used for the proposal densities. The MCMC moves, in the appendix had acceptance rates between 0.3-0.5 and this was quite a reasonable performance. The IS method was run with only 4000 samples. The higher CPU time, for IS, is due to the fact that a bisection method is run for each sample (which was allowed a maximum of 300 steps).

The variance reduction factors, across the multiple runs, for 5 different settings of $S_0$ and $K$ can be observed in Table 2. Note that whilst the option can be standardized for $S_0$, our aim is simply to show that the conclusions of this experiment are relatively invariant to the parameters used. In the Table, a substantial improvement for the same CPU times can be observed. It should be noted, however, for larger $\nu$ (in the latent process) both algorithms do not work well. The latter parameter will increase the variability of the volatility process, and it is not clear how this can be dealt with. Note the results here should not be contrasted with (slightly disappointing) those in Section 3.3.3; this technique differs substantially from the ideas there.

## 5 Summary

In this article we have provided a summary of SMC methods and shown the potential benefits for their application in option pricing and sensitivity estimation. Many of the

| Settings | Variance Reduction |
|---|---|
| $S_0 = 1,\ K = 0.1$ | $1.09 \times 10^{11}$ |
| $S_0 = 1,\ K = 0.9$ | $3.70 \times 10^{11}$ |
| $S_0 = 5,\ K = 1$ | $4.37 \times 10^{11}$ |
| $S_0 = 50,\ K = 10$ | $4.37 \times 10^{11}$ |
| $S_0 = 50,\ K = 49$ | $1.67 \times 10^{12}$ |

Table 2: Variance Reduction of an SMC Method against Importance Sampling. This is for an arithmetic Asian option, under the BNS SV model. The algorithms were both run 50 times.

examples presented in this paper have been based upon translations of algorithms that have been used in other contexts such as stochastic control (an exception is Section 4). On the basis of the work here, it is felt that the application of SMC methods focussed upon particular problems, is likely to be highly beneficial.

The reader of this article should be cautious. There is no claim that SMC methods will always work well, and indeed be uniformly superior to other Monte Carlo methods, or even deterministic approaches. In the process of working on the article, many financial models were too complex to apply standard SMC methods. What is claimed, at least, is that the methods can help to push the boundaries of models that may be used for pricing and is a useful technique for any Monte Carlo specialist in option pricing.

**Acknowledgements**

# Appendix

In this appendix the proof of Proposition 1 as well as the model and MCMC method that was used in Section 4 are given.

## Proposition 1

*Proof.* A more general proof, by induction, is given. Let $m = 2$, and $f \in \mathcal{B}_b(E_2)$ (the bounded measurable functions on $E_2$), then:

$$\widetilde{\Lambda}_2(f) = \int_{E_1} \frac{\partial}{\partial \theta}\{p_\theta(x_1|x_0)\}\widetilde{\Phi}(s_{t_1})\left[\int_{E_2} f(x_2)p_\theta(x_2|x_1)\widetilde{\Phi}(s_{t_2})dx_2\right]dx_1 +$$
$$\int_{E_1} p_\theta(x_1|x_0)\widetilde{\Phi}(s_{t_1})\left[\int_{E_2} f(x_2)\frac{\partial}{\partial \theta}\{p_\theta(x_2|x_1)\}\widetilde{\Phi}(s_{t_2})dx_2\right]dx_1$$

application of the Fubini theorem gives

$$\widetilde{\Lambda}_2(f) = \int_{E_1 \times E_2} \widetilde{\Phi}(s_{t_1})\widetilde{\Phi}(s_{t_2})f(x_2)\frac{\partial}{\partial \theta}\{p_\theta(x_1|x_0)\}p_\theta(x_2|x_1)dx_{1:2} +$$
$$\int_{E_1 \times E_2} \widetilde{\Phi}(s_{t_1})\widetilde{\Phi}(s_{t_2})f(x_2)p_\theta(x_1|x_0)\frac{\partial}{\partial \theta}\{p_\theta(x_2|x_1)\}dx_{1:2}$$
$$= \int_{E_1 \times E_2} \widetilde{\Phi}(s_{t_1})\widetilde{\Phi}(s_{t_2})f(x_2)\frac{\partial}{\partial \theta}\{p_\theta(x_2|x_1)p_\theta(x_1|x_0)\}dx_{1:2}$$
$$= \Lambda_\theta(\Phi f, s_0).$$

Now assume that the identity holds for $m = n$ and consider $n + 1$, $f \in \mathcal{B}_b(E_{n+1})$:

$$
\begin{aligned}
\Lambda_\theta(\Phi f, s_0) &= \int_{E_{[1,n+1]}} \prod_{i=1}^{n+1} \widetilde{\Phi}(s_{t_i}) f(x_{n+1}) \frac{\partial}{\partial \theta} \{ \prod_{i=1}^{n+1} p_\theta(x_i | x_{i-1}) \} dx_{1:n+1} \\
&= \int_{E_{[1,n+1]}} \prod_{i=1}^{n+1} \widetilde{\Phi}(s_{t_i}) f(x_{n+1}) \frac{\partial}{\partial \theta} \{ p_\theta(x_{n+1} | x_n) \} \prod_{i=1}^{n} p_\theta(x_i | x_{i-1}) \} dx_{1:n+1} + \\
&\quad \int_{E_{[1,n+1]}} \prod_{i=1}^{n+1} \widetilde{\Phi}(s_{t_i}) f(x_{n+1}) p_\theta(x_{n+1} | x_n) \frac{\partial}{\partial \theta} \left\{ \prod_{i=1}^{n} p_\theta(x_i | x_{i-1}) \right\} dx_{1:n+1} \\
&= \int_{E_{[n,n+1]}} \widetilde{\Phi}(s_{t_{n+1}}) f(x_{n+1}) \frac{\partial}{\partial \theta} \{ p_\theta(x_{n+1} | x_n) \} \left[ \int_{E_{[1,n-1]}} \left\{ \prod_{i=1}^{n} p_\theta(x_i | x_{i-1}) \times \right. \right. \\
&\quad \left. \left. \widetilde{\Phi}(s_{t_i}) \right\} dx_{1:n-1} \right] dx_{n:n+1} + \int_{E_{[n,n+1]}} \widetilde{\Phi}(s_{t_{n+1}}) f(x_{n+1}) p_\theta(x_{n+1} | x_n) \\
&\quad \left[ \int_{E_{[1,n-1]}} \left\{ \prod_{i=1}^{n} \widetilde{\Phi}(s_{t_i}) \right\} \frac{\partial}{\partial \theta} \left\{ \prod_{i=1}^{n} p_\theta(x_i | x_{i-1}) \right\} dx_{1:n-1} \right] dx_{n:n+1} \\
&= \int_{E_{[n,n+1]}} \widetilde{\Phi}(s_{t_{n+1}}) f(x_{n+1}) \frac{\partial}{\partial \theta} \{ p_\theta(x_{n+1} | x_n) \} \Pi_n(dx_n) dx_{n+1} + \\
&\quad \int_{E_{[n,n+1]}} \widetilde{\Phi}(s_{t_{n+1}}) f(x_{n+1}) p_\theta(x_{n+1} | x_n) \widetilde{\Lambda}_n(dx_n) dx_{n+1} \\
&= \widetilde{\Lambda}_{n+1}(f)
\end{aligned}
$$

setting $f \equiv 1$ completes the proof. $\qquad \square$

## Model

Consider a finite collection of time-points $0 = t_0 < t_1 < \cdots < t_m = T$, then the joint density of the log-price and integrated volatilities are:

$$
\left\{ \prod_{i=1}^{m} \phi\{ y_{t_i}; y_{t_{i-1}} + \mu(t_i - t_{i-1}), \tilde{\sigma}_i(v_{t_1:t_i}) \} p(v_{t_i}) \right\}
$$

where $\phi(\cdot; \mu, \sigma)$ is the normal density of mean $\mu$ and variance $\sigma$ and, with $i \in \mathbb{T}_m$, $v_{t_i} = (a^i_{1:n_i}, r^i_{1:n_i}, n_i) \in [0, \lambda\nu(t_i - t_{i-1})]^{n_i} \times [0, 1]^{n_i} \times \mathbb{N}$ is a 2-dimensional Poisson process on $[0, \lambda\nu(t_i - t_{i-1})] \times [0, 1]$ of unit rate. The functions $\tilde{\sigma}_i(\cdot)$ are as follows. Let $\bar{\sigma}_0 = v_0$ and

define

$$\gamma_{i,1} \ := \ e^{-\lambda(t_i - t_{i-1})} \sum_{j=1}^{n_i} \log\left(\frac{\lambda\nu(t_i - t_{i-1})}{a_j^i}\right) e^{\lambda(t_i - t_{i-1})r_j^i}$$

$$\gamma_{i,2} \ := \ \sum_{j=1}^{n_i} \log\left(\frac{\lambda\nu(t_i - t_{i-1})}{a_j^i}\right)$$

then

$$\bar{\sigma}_i(v_{t_1:t_i}) \ = \ e^{-\lambda(t_i - t_{i-1})}\bar{\sigma}_{i-1}(v_{t_1:t_i}) + \gamma_{i,1}$$

$$\tilde{\sigma}_i(v_{t_1:t_i}) \ = \ \gamma_{i,2} - \bar{\sigma}_i(v_{t_1:t_i}) + \bar{\sigma}_{i-1}(v_{t_1:t_i}).$$

## MCMC

The moves are based upon Metropolis-Hastings kernels; see [59] for an introduction. There are two main quantities to be updated; the $\nu_{t_1:t_m}$ and the parameters

$$(a_{1:n_1}^1, r_{1:n_1}^1, n_1, \ldots, a_{1:n_m}^m, r_{1:n_m}^m, n_m)$$

that make up the volatility process. From here-in simplify the notation to $\nu_{1:m}$.

The $\nu_{1:m}$ are updated, by picking an $i \in \mathbb{T}_m$, uniformly at random, and sampling from its process density. The move is accepted or rejected, according to a standard Metropolis-Hastings acceptance probability:

$$1 \wedge \frac{\varphi_{\nu_i'}\{\nu_{i+1}; \tilde{\mu}(\nu_{i-1}, \nu_i'), \tilde{\sigma}_{i+1}(v_{1:i+1})\}}{\varphi_{\nu_i}\{\nu_{i+1}; \tilde{\mu}(\nu_{i-1:i}), \tilde{\sigma}_{i+1}(v_{1:i+1})\}} \times \frac{\varphi_{\nu_{i+1}}\{\nu_{i+2}; \tilde{\mu}(\nu_i', \nu_{i+1}), \tilde{\sigma}_{i+2}(v_{1:i+2})\}}{\varphi_{\nu_{i+1}}\{\nu_{i+2}; \tilde{\mu}(\nu_{i:i+1}), \tilde{\sigma}_{i+2}(v_{1:i+2})\}}$$

with $1 \le i \le m-2$; a similar formula can be calculated if $i \in \{m-1, m\}$.

The $(a_{1:n_1}^1, r_{1:n_1}^1, n_1, \ldots, a_{1:n_m}^m, r_{1:n_m}^m, n_m)$ are updated in a similar way to [39]. Again, pick an $i \in \mathbb{T}_m$ uniformly at random. Then, the number of points, $n_i$ is either increased by 1 (birth), or decreased, if possible, by 1 (death); the choice is made at random. If a birth occurs, the new $a_i, r_i$ are sampled according to the process density. The Metropolis-Hastings acceptance probability, for a birth is

$$1 \wedge \frac{\prod_{j=i}^m \varphi_{\nu_{j-1}}\{\nu_j; \tilde{\mu}(\nu_{j-2:j-1}), \tilde{\sigma}_j(v_{1:i}, , v_{i:j}')\}}{\prod_{j=i}^m \varphi_{\nu_{j-1}}\{\nu_j; \tilde{\mu}(\nu_{j-2:j-1}), \tilde{\sigma}_j(v_{1:j})\}} \times \frac{\lambda\nu(t_i - t_{i-1})d(n_i + 1)}{b(n_i)}$$

where $b(\cdot), d(\cdot)$ are the probabilities of proposing birth and death moves. The death move, when $n_i = n_i + 1$, has a ratio that is the inverse of that above.

# References

[1] ANDRIEU, C., DOUCET, A. & TADIĆ, V. (2009). On-line Parameter Estimation in General State-Space Models using Pseudo-Likelihood. Technical Report, University of Bristol.

[2] ASMUSSEN, S. & GLYNN, P. (2007).*Stochastic Simulation: Algorithms and Analysis.* Springer: New York.

[3] BALDI, P., CARAMELLINO, L. & IOVINO, I. (1999).Pricing general barrier options: A numerical approach using sharp large deviations. *Math. Finance*, **9**, 239–321.

[4] BARNDORFF-NIELSEN, O. E. & SHEPHARD, N. (2001). Non-Gaussian OU-based models and some of their uses in financial economics (with discussion). *J. R. Statist. Soc. B*, **63**, 167–241.

[5] BENTH, F. E., GROTH, M. & WALLIN, O. (2007). Derivative free Greeks for the Barndorff-Nielsen and Shephard stochastic volatility model. Technical Report, University of Oslo.

[6] BERNIS, G., GOBET, E. & KOHATSU-HIGA, A. (2003). Monte Carlo evaluation of Greeks for multi-dimensional barrier and look-back options. *Math. Finance*, **13**, 99–113.

[7] BESKOS, A., PAPASPILIOPOULOS, O., ROBERTS, G. O. & FEARNHEAD, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *J. R. Statist. Soc. B*, **68**, 333–382.

[8] BOYLE, P. (1977). Options: A Monte Carlo approach. *J. Fin. Econ.*, **4**, 323–338.

32

[9] BROADIE, M. & JAIN, A. (2008). Pricing and hedging volatility derivatives. *J. Deriv.*, **15**, 7–24.

[10] BROADIE, M., GLASSERMAN, P. & KOU, S. G. (1997). A continuity correction for discrete barrier options. *Math. Finance*, **7**, 325–348.

[11] BROADIE, M., & GLASSERMAN, P. (1996). Estimating security price derivatives using simulation. *Managem. Sci.*, **42**, 269–285.

[12] CÉROU, F., DEL MORAL, P. & GUYADER, A., (2008). A non asymptotic variance theorem for unnormalized Feynman-Kac particle models. Technical Report, INRIA.

[13] CÉROU, F., LEGLAND, F. & NEWTON, N. J., (2001). *Stochastic particle methods for linear tangent equations.* in Optimal Control and PDEs - Innovations and Applications, Eds. Menaldi, J., Rofman, E. & Sulem, A., 231-240, IOS Press: Amsterdam.

[14] CHEN, Y., XIE, Y. & LIU, J. S. (2005). Stopping time resampling for sequential Monte Carlo methods. *J. R. Statist. Soc. Ser. B*, **67**, 199–219.

[15] CHOPIN, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann Statist.*, **32**, 2385–2411.

[16] CONT, R. & VOLTCHKOVA, E. (2006). Finite difference methods for option pricing in jump- diffusion and exponential Lévy models. *SIAM J. Numer. Anal.*, **43**, 1596–1626.

[17] DE FREITAS, N., NIRANJAN, M., GEE, A. H. & DOUCET, A. (2000).Sequential Monte Carlo methods to train neural networks. *Neural Computation*, **12**, 955–993.

[18] DEL MORAL, P. (2004). *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications.* Springer: New York.

[19] DEL MORAL, P., DOUCET, A. & JASRA, A. (2009). *Sequential Monte Carlo Simulation: Methods and Theory.* Book Project, In progress.

[20] DEL MORAL, P., DOUCET, A. & JASRA, A. (2008). On adaptive resampling procedures for sequential Monte Carlo methods. Technical Report, INRIA.

[21] DEL MORAL, P., DOUCET, A. & JASRA, A. (2006). Sequential Monte Carlo samplers. *J. R. Statist. Soc. B*, **68**, 411–436.

[22] DEL MORAL, P., DOUCET, A. & SINGH, S. S. (2010a). A backward particle interpretation of Feynman-Kac formulae. Technical Report, INRIA.

[23] DEL MORAL, P., DOUCET, A. & SINGH, S. S. (2010b). Forward smoothing using sequential Monte Carlo. Technical Report, University of Cambridge.

[24] DETEMPLE, J. & RINDISBACHER, M. (2007). Monte Carlo methods for derivatives of options with discontinuous payoffs. *Comp. Statist. Data Anal.*, **51**, 3393–3417.

[25] DOUCET, A., BRIERS, M. & SÉNÉCAL, S (2006). Efficient block sampling strategies for sequential Monte Carlo methods. *J. Comp. Graph. Statist.*, **15**, 693–711.

[26] DOUCET, A., GODILL, S. & ANDRIEU, C (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statist. Comp.*, **10**, 197–208.

[27] DOUCET, A., DE FREITAS, J. F. G. & GORDON, N. J. (2001). *Sequential Monte Carlo Methods in Practice*. Springer: New York.

[28] ETORE, P., FORT, G., JOURDAIN, B. & MOULINES, É. (2009). On adaptive stratification. Technical Report, CMAP, École Polytechnique.

[29] FEARNHEAD, P. , PAPASPILIOPOULOS, O. & ROBERTS, G. O. (2008). Particle filters for partially observed diffusions. *J. R. Statist. Soc. Ser B*, **70**, 755-777.

[30] FLORESCU, M., & VIENS, F. G. (2008). Stochastic volatility: option pricing using a multinomial recombining tree. *Appl. Math. Finance*, **15**, 151–181.

[31] FOURNIÉ, E., LASRY, J. M., LEBUCHOUX, J. & LIONS, P. L. (2003). Applications of Malliavin calculus to Monte Carlo methods in finance II. *Finance Stochst.*, **5**, 201–236.

[32] GELMAN, A., & MENG, X. L. (1998). Simulating normalizing constants; from importance sampling to bridge sampling to path sampling. *Statist. Sci.*, **13**, 163–185.

[33] GLASSERMAN, P. (2003). *Monte Carlo Methods in Financial Engineering.* Springer: New York.

[34] GLASSERMAN, P., HIEDELBERGER, P. & SHAHABUDDIN, P. (1999). Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Math. Finance*, **9**, 117–152.

[35] GLASSERMAN, P., & STAUM, J. (2001). Conditioning on one-step survival for barrier options. *Op. Res.*, **49**, 923–937.

[36] GLASSERMAN, P., & LIU, Z. (2008). Sensitivity estimates from characteristic functions. *Finan. Stoch.*, **12**, 507-540.

[37] GLYNN, P. W. & IGLEHART, D. L. (1989). Importance sampling for stochastic simulations. *Manage. Sci.*, **35**, 1367–1392.

[38] GORDON, N. J., SALMOND, D.J. & SMITH, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings F*, **140**, 107–113.

[39] GRIFFIN, J., & STEEL, M. (2006). Inference with non-Gaussian Ornstein-Uhlenbeck processes for stochastic volatility. *J. Econom.*, **134**, 605–644.

[40] HAMMERSLEY, J. M., & MORTON, K. W. (1954). Poor man's Monte Carlo. *J. R. Statist. Soc. Ser. B*, **16**, 23–38.

[41] JARZYNSKI, C. (1998). Nonequilibrium equality for free energy differences. *Phys. Rev. Lett* **78**, 2690–2693.

[42] JASRA, A. & DOUCET, A, (2009). Sequential Monte Carlo methods for diffusion processes. *Proc. R. Soc. A*, **465**, 3709-3727.

[43] JASRA, A., STEPHENS, D. A., DOUCET, A. & TSAGARIS, T. (2009). Inference for Lévy driven stochastic volatility models via adaptive sequential Monte Carlo. Technical Report. Imperial College London.

[44] JOHANNES, M., POLSON, N., & STROUD, J. (2006). Identifying jumps and stochastic volatility: Filtering stochastic differential equations with jumps. Technical Report, Columbia University.

[45] JOURDAIN, B. & SBAI, M. (2007). Exact retrospective Monte Carlo computation of arithmetic average Asian options. *Monte Carlo Meth. Appl.*, **13**, 135–171.

[46] KEMNA, A. G. Z. & VORST, A. C. F. (1990). A pricing method for options based on average asset values. *J. Bank. Finan.*, **14**, 113-129.

[47] KONG, A., LIU, J. S. & WONG, W. H. (1994). Sequential imputations and Bayesian missing data problems. *J. Amer. Statist. Assoc.*, **89**, 278–288.

[48] L'ECUYER, P. (1995). On the interchange of derivative and expectation for likelihood ratio derivative estimators. *Manage. Sci.*, **41**, 738748.

[49] L'ECUYER, P. & DEMERS, V. & TUFFIN, B. (2007). Rare-events, splitting, and quasi-Monte Carlo. *ACM Trans. Comp. Mod. Simul.*, **17**(2), article 9.

[50] LEE, A., YAU, C., GILES, M. B., DOUCET, A. & HOLMES, C.C. (2009). On the utility of graphics cards to perform massively parallel simulation of advanced Monte Carlo methods, Technical Report, University of Oxford.

[51] LIU, J. S. (2001). *Monte Carlo Strategies in Scientific Computing.* Springer: New York.

[52] NEAL, R. M. (2001). Annealed importance sampling. *Statist. Comp.*, **11**, 125–139.

[53] OLSSON, J., CAPPÉ, O., DOUC, R. & MOULINES, É. (2008). Sequential Monte Carlo smoothing filters with application to parameter estimation in non-linear state-space models. *Bernoulli*, **14**, 155–179.

[54] Papaspiliopoulos, O. (2010). A methodological framework for Monte Carlo probabilistic inference for diffusion processes. In *Inference and Learning in Dynamic Models* (eds Barber, D., Chippa, S. & Cemgil, A. T.), Cambridge: CUP.

[55] Pitt, M. K. & Shepherd, N. (1999). Filtering via simulation: Auxiliary particle filters. *J. Amer. Statist. Assoc.*, **94**, 590–599.

[56] Poyiadjis, G. (2006). *Particle Methods for Parameter Estimation in General State-Space Models*. PhD Thesis, University of Cambridge.

[57] Poyiadjis, G., Doucet, A. & Singh, S.S. (2009). Sequential Monte Carlo computation of the score and observed information matrix in state-space models with application to parameter estimation. Technical Report, University of Cambridge.

[58] Rambharat, B. R. & Brockwell, A. E. (2010). Sequential Monte Carlo pricing of American style options under stochastic volatility models. *Ann. Appl. Statist.* (to appear).

[59] Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer: New York.

[60] Rosenbluth, M. N. & Rosenbluth, A. W. (1955). Monte Carlo calculation of the average extension of Molecular chains. *J. Chem. Phys.*, **23**, 356–359.

[61] Rousset, M. (2006). *Continuous Time Population Monte Carlo and Computational Physics*. Ph.D. thesis, Universitié Paul Sabatier Toulouse.

[62] Rousset, M., & Doucet, A. (2006). Discussion of Beskos et al. *J. R. Statist. Soc. B*, **68**, 374–375.

[63] Rubinstein, R. & Krose, D.P. (2004). *Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*. Springer: New York.

[64] VECER, J. (2002). Unified Asian pricing. *Risk*, **6**, 113–116.

[65] VECER, J. & XU, M. (2006). Pricing Asian options in a semimartingale model. *Quant. Finan.*, **4**, 170–175.

[66] ZHAO, G., ZHOU, Y. & VAKILI, P. (2006). A new efficient simulation strategy for pricing path dependent options. *Proc. 2006 Winter Simul. Conf.*, Eds L.F. Perron, F. P. Wieland, J. Liu, B. F. Lawson, D. M. Nicol & R. M. Fujimoto.