Nationaal Lucht- en Ruimtevaartlaboratorium
National Aerospace Laboratory NLR

NLR-TP-2006-417

Sequential Monte Carlo simulation of rare event probability in stochastic hybrid systems

J. Krystul[1] and H.A.P. Blom

[1] Twente University

This report contains a paper presented at the 16th IFAC world Congress in Prague, Czech Republic, 4-8 June 2005.

Customer	National Aerospace Laboratory NLR
Contract number	----
Owner	National Aerospace Laboratory NLR
Division	Air Transport
Distribution	Unlimited
Classification of title	Unclassified
	July 2006

Approved by:

Author	Reviewer	Managing department
5-7-06	Anonymous peer reviewers	27/7/6

SEQUENTIAL MONTE CARLO SIMULATION OF RARE EVENT PROBABILITY IN STOCHASTIC HYBRID SYSTEMS

Jaroslav Krystul * **Henk A.P. Blom** **

* *University of Twente, Enschede, The Netherlands*
** *National Aerospace Laboratory NLR, Amsterdam*

Abstract: Recently (Cérou *et al.*, 2002) developed an elegant factorization of rare event probabilities appearing in diffusion processes and other strong Markov processes, and a sequential Monte Carlo simulation approach to estimate the factorized rare event probability. The paper extends this approach towards rarely switching diffusions, and demonstrates the effectiveness for a simple example.

Keywords: Monte Carlo, Markov models, Probabilistic risk assessment, Stochastic jump processes, Uncertain dynamic systems

1. INTRODUCTION

Stochastic dynamical modeling of accident risk is of high interest for the safe design of complex safety-critical operations, such as in nuclear and chemical industries, and advanced air traffic management, e.g. see (Smidts *et al.*, 1998; Labeau *et al.*, 2000; Blom *et al.*, 2003*a*) and their references. In comparison with statistical analysis of collected data (Embrechts *et al.*, 1997), stochastic dynamical modeling approach has the advantage of enabling stochastic analysis and advanced Monte Carlo (MC) simulation approaches (Doucet *et al.*, 2001).

Obtaining accurate estimates of rare event probabilities, say about 10^{-9} to 10^{-12}, is not realistic just by using straightforward MC simulation. This makes MC simulation to be a practical alternative only when it is possible to realize a high speed up. The techniques used in (Smidts *et al.*, 1998; Labeau *et al.*, 2000; Blom *et al.*, 2003*a*) for speeding up MC simulation are model specific risk decompositions. Hence there is need for a more systematic and general approach. A well known approach is importance sampling (Liu, 2003), which is based on a modification of the underlying probability distribution in such a way that the rare events occur much more frequently. Unfortunately, for rare event simulation, importance sampling alone often does not provide the required speed-up. An alternative approach to increase the relative number of visits to the rare event is to make use of the fact that there exist some well identifiable intermediate states that are visited much more often than the rare event states themselves and behave as gateway states to reach the rare event states (Townsend *et al.*, 1998). In (Cérou *et al.*, 2002) this idea has been elaborated in terms of stochastic analysis, and subsequently combined with a sequential MC based evaluation. The result is a specific Interacting Particle System (IPS) algorithm (Moral, 2004). In (Krystul and Blom, 2004) it has been shown that this IPS approach works very well for a diffusion example. The aim of this paper is to extend this IPS approach to estimate the rare event probability for

[1] This research has been performed with support of the European Commission through the HYBRIDGE project.

rarely switching diffusions (Ghosh *et al.*, 1993), (Ghosh *et al.*, 1997).

Straightforward application of the IPS approach of (Cérou *et al.*, 2002) to rarely switching diffusions has certain limitations. First, there will be few particles in a mode with small conditional probability, i.e. a "light" mode. Second, if switching rate is small it may be unlikely that there is even one switch during a simulation run. In such case, the possible switching between modes is not properly taken into account and this badly affects estimator performance. In order to improve for this, the paper develops two extensions: sampling per mode to cope with large differences in mode weights, and importance switching to cope with rare mode switching.

The paper is organized as follows. Section 2 states the problem studied. Sections 3 and 4 respectively present the factorization approach and the IPS algorithm background of (Cérou *et al.*, 2002). Sections 5 and 6 extend the IPS algorithm in order to cope with large differences in mode probabilities and rarely switching diffusions respetively. Numerical evaluation and comparison of different versions of the IPS algorithms are given in section 7. Section 8 draws conclusions.

2. THE PROBLEM CONSIDERED

Throughout this and the next sections, all stochastic processes are defined on a complete stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, \mathbb{T})$ with index set $\mathbb{T} = \mathbb{R}_+$, and $\mathbb{F}$ a right continuous filtration (an increasing family of sub-σ-algebras of $\mathcal{F}$).

Let $\{x_t, \theta_t\}$ be a switching diffusion taking its values in $\mathbb{R}^n \times \mathbb{M}$ according to

$$dx_t = a(\theta_t, x_t)dt + b(\theta_t, x_t)dW_t, \tag{1}$$

$$P_{\theta_{t+\delta}|\theta_t,x_t}(\theta|\eta, x) = \lambda_{\eta\theta}(x)\delta + o(\delta),\ \eta \neq \theta, \tag{2}$$

where $\mathbb{M}$ is a finite set of modes and $(W_t)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^n$ independent of $\{\theta_t\}$ and of initial condition (x_0, θ_0), a prescribed $(\mathbb{R}^n \times \mathbb{M})-$valued random variable. Under assumption on W_t, (x_0, θ_0), and on functions a, b and λ_{ij}, equation (1,2) admits an a.s. pathwise unique solution (Ghosh *et al.*, 1993) and which is a strong Markov process (Blom *et al.*, 2003*b*).

We set $\tau_D \overset{\triangle}{=} \inf\{t > 0 : x_t \in D\}$ for the first passage time of $\{x_t\}$ to a closed connected Borel set D. The problem addressed in the sequel is to estimate the probability $P_{hit}(0, T)$ that $\{x_t\}$ will hit the set D on the time interval $(0, T]$, $T < \infty$:

$$P_{hit}(0, T) \overset{\triangle}{=} P(\tau_D < T). \tag{3}$$

Examples of P_{hit} in air traffic are:
1) Conflict probability (Hu *et al.*, 2003; Watkins and Lygeros, 2003) in which D forms the subset in the state space where aircraft are closer to each other than some minimum separation criterion (e.g. 5 Nm in horizontal direction).
2) Collision probability (Blom *et al.*, 2003*a*) in which D is the subset in the state space where aircraft are closer to each other than their physical sizes (of order 100m in horizontal direction).

In air traffic, collision probability should be orders of magnitude smaller than conflict probability. As such the prime objective of this paper is to address the more rare collision event.

3. FACTORIZATION APPROACH

In (Cérou *et al.*, 2002; Moral, 2004) a sequence of gateway states has been used to characterize the rare event probability as a product of conditional probabilities by using Feynman-Kac model in path space. Here we explain how this product form can be obtained for a switching diffusion, the first component of which counts time. We assume that switching diffusion (1), (2) starts at $t = 0$ in a Borel set $\bar{D}_0 \subset \{0\} \times \mathbb{R}^{n-1} \times \mathbb{M}$ with a known initial probability distribution $P_{x_0,\theta_0}(\cdot)$. As in (Cérou *et al.*, 2002) we assume a sequence of nested Borel sets, $\bar{D} = \bar{D}_m \subset \cdots \subset \bar{D}_1$ which are defined as follows:

$$\bar{D}_k \overset{\triangle}{=} (0, T) \times D_k \times \mathbb{M},\ k = 1, \ldots, m \tag{4}$$

where D_k is a closed Borel set of $\mathbb{R}^{n-1}$, and $\bar{D}_1$ such that $\bar{D}_1 \cap \bar{D}_0 = \varnothing$. The first moment that $\{x_t, \theta_t\}$ hits a set $\bar{D}_k$ is defined as the stopping time:

$$\tau_k \overset{\triangle}{=} \inf\{t \geq 0 : (x_t, \theta_t) \in \bar{D}_k\},$$

$\tau_k = \infty$ if this set is empty. We want to estimate $\mathbb{P}(\tau_m < T)$, for some $T < \infty$, i.e. the probability that switching diffusion $\{x_t, \theta_t\}$ will hit the rare event set $\bar{D}$ before time T. The process $\{x_t, \theta_t\}$, before hitting $\bar{D}$, passes through a sequence of nested Borel sets (4). Following (Cérou *et al.*, 2002) we introduce the $\{0, 1\}$-valued variables $\{y_k, k = 1, \ldots, m\}$:

$$y_k(\omega) \overset{\triangle}{=} \mathbf{1}_{\{\omega: x_{\tau_k}(\omega) \in (0,T) \times D_k\}}. \tag{5}$$

Hence, for each k we have

$$y_k(\omega) = \mathbf{1}_{\{\omega:\tau_k(\omega)<T\}} = \prod_{i=1}^{k} \mathbf{1}_{\{\omega:\ \tau_i(\omega)<T\}} = \prod_{i=1}^{k} y_i(\omega). \tag{6}$$

Next we characterize $P_{hit}(0, T)$ in terms of the sequence $\{y_k\}$. By its definition,

$$P_{hit}(0, T) = \mathbb{P}(\tau_m < T) = \mathbb{E}[\mathbf{1}_{\{\tau_m<T\}}]$$

Subsequent substitution of (5) and (6) yields:

$$P_{hit}(0, T) = \mathbb{E}[y_m] = \mathbb{E}[\prod_{k=1}^{m} y_k]. \tag{7}$$

Since y_k assumes values from $\{0,1\}$,

$$\mathbb{E}[\prod_{k=1}^{m} y_k] = \prod_{k=1}^{m} \mathbb{E}[y_k | y_{k-1} = 1, \ldots, y_1 = 1]$$

Substituting this into (7) yields

$$\begin{aligned} P_{hit}(0,T) &= \prod_{k=1}^{m} \mathbb{E}[y_k | y_{k-1} = 1, \ldots, y_1 = 1] \\ &= \prod_{k=1}^{m} \mathbb{P}(\tau_k < T | \tau_{k-1} < T, \ldots, \tau_1 < T) \\ &= \prod_{k=1}^{m} \mathbb{P}(\tau_k < T | \tau_{k-1} < T) \end{aligned} \tag{8}$$

This means that (8) characterizes the probability $P_{hit}(0,T)$ of the rare event as a product of conditional probabilities of intermediate "less rare" events leading to it. Thus, if we define the conditional probabilities

$$\gamma_k \triangleq \mathbb{P}(\tau_k < T | \tau_{k-1} < T) \text{ for } k = 1, \ldots, m$$

and insert this in (8) then we get for $P_{hit}(0,T)$:

$$P_{hit}(0,T) = \prod_{k=1}^{m} \gamma_k \tag{9}$$

The estimation of the probabilities γ_k is subsequently accomplished by the IPS approach of (Cérou *et al.*, 2002).

4. IPS ALGORITHM

Let us denote $E' = \mathbb{R}^n \times \mathbb{M}$, and let $\mathcal{E}'$ be the Borel $\sigma-$algebra of E'. For any $B \in \mathcal{E}'$, $\pi_k(B)$ denotes the conditional probability of $\xi_k \triangleq (x_{\tau_k}, \theta_{\tau_k}) \in B$ given $y_{1:k} = (1,1,\ldots,1)$. Then the estimation of the probability in subsequently hitting the nested Borel sets by $\{\xi_k\}$ is characterized through the following sequence of transformations

$$\pi_{k-1}(\cdot) \xrightarrow{\text{prediction}} p_k(\cdot) \xrightarrow{\text{conditioning}} \pi_k(\cdot),$$

where $p_k(B)$ is the condition probability of $\xi_k \in B$ given $y_{1:k-1} = (1,1,\ldots,1)$. Because $\{\xi_t\}$ is a Markov sequence the prediction satisfies:

$$p_k(B) = \int_{E'} P_{\xi_k|\xi_{k-1}}(B|\xi)\pi_{k-1}(d\xi) \text{ for all } B \in \mathcal{E}', \tag{10}$$

and the conditioning satisfies:

$$\pi_k(B) = \frac{\int_B \mathbf{1}_{\{\xi \in \bar{D}_k\}} p_k(d\xi)}{\int_{E'} \mathbf{1}_{\{\xi' \in \bar{D}_k\}} p_k(d\xi')} \text{ for all } B \in \mathcal{E}'. \tag{11}$$

Then

$$\begin{aligned} \gamma_k &= P(\tau_k < T | \tau_{k-1} < T) \\ &= \mathbb{E}[y_k | y_{1:k-1} = (1,1,\ldots,1)] \\ &= \int_{E'} \mathbf{1}_{\{\xi \in \bar{D}_k\}} p_k(d\xi). \end{aligned}$$

With this each of the m terms γ_k in (9) is characterized as a solution of a sequence of "filtering" kind of equations (10,11). However, an important difference with "filtering" equations is that (10,11) are ordinary integral equations, i.e. they have no stochastic term entering them.

The sequence of transformations (10),(11) leads to the IPS algorithm of (Cérou *et al.*, 2002) to estimate $P_{hit}(0,T) = \mathbb{P}(\tau_m < T)$. In this algorithm $\gamma_k^{N_p}$, $p_k^{N_p}$ and $\pi_k^{N_p}$ denote the numerical approximations of γ_k, p_k and π_k respectively:

Step 0. Level sets

- Choose appropriate nested sequence of closed subsets of $\mathbb{R}^{n-1}$: $D = D_m \subset D_{m-1} \subset \cdots \subset D_1$, and define $\bar{D}_k = (0,T) \times D_k \times \mathbb{M}$, $k = 1, \ldots, m$.

Step 1. Initial sampling; $k = 0$.

- For $i = 1, \ldots, N_p$ generate initial state value outside $\bar{D}_1$:
  $(x_0^i, \theta_0^i) \sim P_{x_0,\theta_0}(\cdot)$ and set $\xi_0^i = (x_0^i, \theta_0^i)$
- For $i = 1, \ldots, N_p$ set the initial weights: $\omega_0^i = 1/N_p$.
- Then

$$\pi_0^{N_p} = \sum_{i=1}^{N_p} \omega_0^i \delta_{\{\xi_0^i\}}.$$

Iteration k; $k = 1, \ldots, m$ over step 2 (prediction) and step 3 (resampling)

Step 2. Prediction step: $\pi_{k-1} \longrightarrow p_k$;

- For $i = 1, \ldots, N_p$ simulate a new path (see (Krystul and Bagchi, 2004)) starting at ξ_{k-1}^i until the k-th set $\bar{D}_k$ is hit, or till $t = T$.
- This yields new particles $\{\hat{\xi}_k^i, \omega_{k-1}^i\}_{i=1}^{N_p}$.
- $p_k^{N_p}$ is the empirical distribution associated with the new cloud of particles:

$$p_k^{N_p} = \sum_{i=1}^{N_p} \omega_{k-1}^i \delta_{\{\hat{\xi}_k^i\}}.$$

- The particles which do not reach the set $\bar{D}_k$ are killed, i.e. we set $\hat{\omega}_k^i = 0$ if $\hat{\xi}_k^i \notin \bar{D}_k$ and $\hat{\omega}_k^i = \omega_{k-1}^i$ if $\hat{\xi}_k^i \in \bar{D}_k$.
- The new set of particles is $\{\hat{\xi}_k^i, \hat{\omega}_k^i\}_{i=1}^{N_p}$.
- Approximation of γ_k:

$$\gamma_k \approx \gamma_k^{N_p} = \sum_{i=1}^{N_p} \hat{\omega}_k^i.$$

  If all particles are killed, i.e. $\gamma_k^{N_p} = 0$, then the algorithm stops without $P_{hit}(0,T)$ estimate.

Step 3. Resampling step: $p_k \longrightarrow \pi_k$

- For $i = 1, \ldots, N_p$ set $\tilde{\xi}_k^i = \hat{\xi}_k^i$ and

$$\tilde{\omega}_k^i = \frac{\hat{\omega}_k^i}{\sum_{j=1}^{N_p} \hat{\omega}_k^j}, \ i = 1, \ldots, N_p,$$

- Resample with replacement N_p particles ξ_k^i according to the empirical measure
$$\pi_k^{N_p} = \sum_{i=1}^{N_p} \tilde{\omega}_k^i \delta_{\{\tilde{\xi}_k^i\}}.$$
- The new set of particles is $\{\xi_k^i, \omega_k^i\}_{i=1}^{N_p}$, with $\omega_k^i = 1/N_p$.
- If $k < m$ then repeat steps 2, 3 for $k := k+1$.
- Otherwise, stop with $P_{hit}(0,T) \approx \prod_{k=1}^m \gamma_k^{N_p}$.

In (Cérou *et al.*, 2002) it is proven that the particle estimates are unbiased, i.e.

$$\mathbb{E}[\prod_{k=1}^m \gamma_k^{N_p}] = \mathbb{P}(\tau_m < T) = P_{hit}(0,T)$$

and also that:

$$(\mathbb{E}(\prod_{k=1}^m \gamma_k^{N_p} - \prod_{k=1}^m \gamma_k)^p)^{\frac{1}{p}} \leq \frac{a_p b_m}{\sqrt{N_p}},$$

for some finite constant a_p which depends on the parameter p, and for some finite constant b_m which depends on the parameter m.

5. SAMPLING PER MODE

The initial sampling step 1 and the resampling step 3 of the IPS algorithm require very many particles for hybrid state processes the conditional mode probabilities of which may be small. Therefore, we propose "sampling per mode" for the initial sampling step 1 and for the resampling step 3. This idea of sampling per mode was successfully introduced in (Blom and Bloem, 2003) for target tracking problems.

If the initial probabilities of some particular modes are very small then it is highly unlikely to draw particles in these modes. To avoid this, at the initial sampling step we start with a fixed number of particles in each mode whatever small the initial probability is, i.e.:
Step 1H. Initial sampling; $k = 0$.

- Choose for each mode $e_i \in \mathbb{M} = \{e_1, \ldots, e_N\}$, $i = 1, \ldots, N$ an integer N_p^i, so that $N_p = \sum_{i=1}^N N_p^i$.
- For each $i = 1, \ldots, N$ sample N_p^i initial state values outside $\bar{D}_1$: $x_0^j \sim P_{x_0|\theta_0}(\cdot|e_i)$, $\theta_0^j = e_i$, set $\xi_0^j = (x_0^j, \theta_0^j)$ and assign initial weight: $\omega_0^j = \frac{P_{\theta_0}(e_i)}{N_p^i}$, for $j \in \{\sum_{k=0}^{i-1} N_p^k + 1, \ldots, \sum_{k=0}^{i-1} N_p^k + N_p^i\}$, with convention $N_p^0 = 0$.
- Then
$$\pi_0^{N_p} = \sum_{i=1}^{N_p} \omega_0^i \delta_{\{\xi_0^i\}}.$$

Similarly, for the resampling step we also keep a fixed number of particles per mode, i.e.:
Step 3H. Resampling step: $p_k \longrightarrow \pi_k$

- For each $\theta \in \mathbb{M}$ evaluate weights:
$$\tilde{\omega}_k^{\theta,i} = \hat{\omega}_k^i \cdot \mathbf{1}_{\{\theta_{\tau_k}^i = \theta\}}, \; i = 1, \ldots, N_p.$$
- For $i = 1, \ldots, N_p$ set $\tilde{\xi}_k^i = \hat{\xi}_k^i$, yielding the empirical unnormalized distribution per mode:
$$\pi_k^{\theta,N_p} \triangleq \sum_{i=1}^{N_p} \tilde{\omega}_k^{\theta,i} \delta_{\{\tilde{\xi}_k^i\}}, \qquad (12)$$
and the total weight per mode equals $\sum_{j=1}^{N_p} \tilde{\omega}_k^{\theta,j}$.
- For all $\theta \in \mathbb{M}$, resample with replacement N_p^θ values ξ_k^i according to the empirical measure (12), and assign per θ weights to particles:
$$\omega_k^\theta = \frac{\sum_{j=1}^{N_p} \tilde{\omega}_k^{\theta,j}}{N_p^\theta}.$$
- If $k < m$ then repeat steps 2 and 3H for $k := k+1$.
- Otherwise stop, with $P_{hit}(0,T) \approx \prod_{k=1}^m \gamma_k^{N_p}$.

6. IMPORTANCE SWITCHING

The possibility of small mode probabilities is covered well by resampling per mode. However for rarely switching diffusion the required number of particles increases when the switching rates decrease. During a prediction step the random paths $(\xi_{k-1:k}^i)_{i=1}^{N_p} = (x_{\tau_{k-1}:\tau_k}^i, \theta_{\tau_{k-1}:\tau_k}^i)_{i=1}^{N_p}$ are being generated to approximate the distribution p_k $(k = 1, \ldots, m)$. If the probability of some transitions (switches) is very small then, most probably, there will be few switches observed during the generation of these random paths. In order to avoid the need to increase the number of particles when the switching rates are decreasing we introduce a sequential importance switching technique.

Coefficients $\lambda_{ij}(\cdot)$ in equation (2) are responsible for switchings. In order to make the rare switches less rare we replace $\lambda_{ij}(\cdot)$ with $\hat{\lambda}_{ij}(\cdot)$, and denote the changed process by $\{\hat{X}_t, \hat{\theta}_t\}$. Then, for $n > k$, $n, k = 1, 2, \ldots$ (Krystul and Blom, 2004, pp.33-35) show for Euler approximations $\{X_t^h, \theta_t^h\}$ and $\{\hat{X}_t^h, \hat{\theta}_t^h\}$ with time step h:

$$\begin{aligned} &P_{X_{t_n}^h, \theta_{t_n}^h | X_{t_k}^h, \theta_{t_k}^h}(A, B | x_k, \theta_k) \\ &= \sum_{\theta_n \in B} \int_A \cdots \sum_{\theta_{k+1} \in \mathbb{M}} \int_{\mathbb{R}^n} \prod_{i=k+1}^n L_{t_i|t_{i-1}}(\theta_i | x_{i-1}, \theta_{i-1}) \\ &\qquad \times P_{\hat{X}_{t_i}^h, \hat{\theta}_{t_i}^h | \hat{X}_{t_{i-1}}^h, \hat{\theta}_{t_{i-1}}^h}(dx_i, \theta_i | x_{i-1}, \theta_{i-1}). \end{aligned}$$

with likelihood ratio:

$$L_{t|s}(\theta | x', \theta') \triangleq \frac{P_{\theta_t^h | X_s^h, \theta_s^h}(\theta | x', \theta')}{P_{\hat{\theta}_t^h | \hat{X}_s^h, \hat{\theta}_s^h}(\theta | x', \theta')}$$

This means, an unbiased estimates of distribution p_k $(k = 1, \ldots, m)$ in the algorithm described in the previous section can be obtained by generating random trajectories of the process $\{\hat{X}_t^h, \hat{\theta}_t^h\}$ (i.e. sampling according to $P_{\hat{x}_t,\hat{\theta}_t|\hat{x}_s,\hat{\theta}_s}(\cdot|x',\theta')$) and adjusting the weight of each particle recursively:

$$\omega_{t_j}^i = \omega_{t_{j-1}}^i \cdot L_{t_j|t_{j-1}}(\theta_j^i|x_{j-1}^i, \theta_{j-1}^i). \qquad (13)$$

We use the importance switching method described above to improve step 2 by forcing rare switchings in discrete component for the case that θ_t is a continuous time Markov chain, i.e. $L_{t|s}(\theta|x,\theta')$ is x invariant.

Step 2H. Prediction step: $\pi_{k-1} \longrightarrow p_k$;

- For $i = 1,\ldots,N_p$ simulate a new path (see (Krystul and Bagchi, 2004)) starting at $(x^i_{\tau_{k-1}}, \theta^i_{\tau_{k-1}})$ to the k-th set $\bar{D}_k$ is reached or till $t = T$;
- This yields a new set of particles $\{\hat{\xi}_k^i, \hat{\omega}_k^i\}_{i=1}^{N_p}$.
- For the i-th particle evaluate the likelihood ratio $L_{\tau_k^i|\tau_{k-1}^i}(\theta_k^i|\theta_{k-1}^i)$.
- The particles, which do not reach the set $\bar{D}_k$ are killed, i.e. we set $\hat{\omega}_k^i = 0$ if $x_k^i \notin D_k$ and otherwise $\hat{\omega}_k^i = \omega_{k-1}^i \cdot L_{\tau_k^i|\tau_{k-1}^i}(\theta_k^i|\theta_{k-1}^i)$
- The new set of particles is $\{\hat{\xi}_k^i, \hat{\omega}_k^i\}_{i=1}^{N_p}$.
- Approximation of γ_k:

$$\gamma_k \approx \gamma_k^{N_p} = \sum_{i=1}^{N_p} \hat{\omega}_k^i.$$

- If in one of the modes all particles are killed, then the algorithm stops without $P_{hit}(0,T)$ estimate.

7. NUMERICAL EXAMPLE

This section illustrates the performance of the Monte Carlo (MC) approach, the IPS algorithm of (Cérou *et al.*, 2002) and the effect of the alternative steps 1H, 2H and 3H for a switching diffusion. Table 1 presents the list of tested algorithms. There IPS stands for the algorithm of (Cérou *et al.*, 2002) in case of a switching diffusion (section 4); HIPS 1 is IPS with improved initial sampling step 1H and with resampling per mode step 3H (section 5); and HIPS 2 is HIPS 1 plus importance switching (section 6). For the example consider a Markovian switching diffusion $(x_t, \theta_t) \in \mathbb{R} \times \{e_1, e_2, e_3\}$, the evolution of which is governed by the following SDE

$$dx_t = \Big(\mu(\theta_t) + \frac{\sigma(\theta_t)^2}{2}\Big)x_t dt + \sigma(\theta_t)x_t dW_t, \quad (14)$$

$$P_{\theta_{t+\delta}|\theta_t,x_t}(\theta|\eta, x) = \lambda_{\eta\theta}(x)\delta + o(\delta), \ \eta \neq \theta. \quad (15)$$

Table 1. Tested Algorithms

Algorithm	Particle system	Initial sampling per mode	Resampling per mode	Importance switching
MC	–	Yes (1H)	–	–
IPS	Yes	–	–	–
HIPS 1	Yes	Yes (1H)	Yes (3H)	–
HIPS 2	Yes	Yes (1H)	Yes (3H)	Yes (2H)

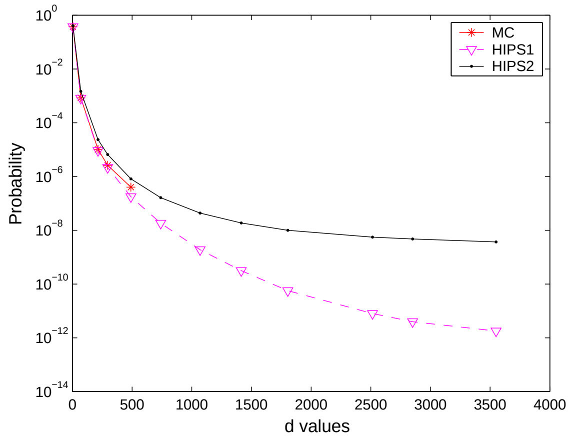

Fig. 1. Probability to hit level d. MC stops at $d_4 = 490$. IPS stops at the beginning. HIPS1 underestimates the probability because it runs on too few mode switched particles.

Initial conditions:

$$x_0 = 1; \qquad P_{\theta_0}(e_1) = 1 - 10^{-7} - 10^{-9};$$
$$P_{\theta_0}(e_2) = 10^{-7}; \qquad P_{\theta_0}(e_3) = 10^{-9};$$

Parameters:

$$\mu(e_1) = 1,\ \mu(e_2) = 4,\ \mu(e_3) = 3,$$
$$\sigma(e_1) = 1,\ \sigma(e_2) = 0.9,\ \sigma(e_3) = 1.7,$$

and with the following rates (independent of x_t):

$$\lambda_{12} = 1\cdot 10^{-4},\ \lambda_{13} = 1\cdot 10^{-6},\ \lambda_{21} = 5\cdot 10^{-5},$$
$$\lambda_{23} = 1\cdot 10^{-5},\ \lambda_{31} = 5\cdot 10^{-5},\ \lambda_{32} = 1\cdot 10^{-4}.$$

We want to estimate the probability that continuous valued component $\{x_t\}$ will hit level d before time T, i.e. $P(\tau_d(x) \le T)$ where $\tau_d(x) \triangleq \inf\{t > 0 : x_t \in [d, +\infty);\ x_0 = x\}$. For the IPS, HIPS1 and HIPS2 the decreasing nested sets $D = D_m \subset \cdots \subset D_1$ are as follows: $D_j = [d_j, +\infty)$, where the values d_j are chosen experimentally so that approximately $40\% - 50\%$ of particles that start at level D_{j-1} reach level D_j, $j = 1,\ldots,m$. For algorithm HIPS2 the importance rates are $\hat{\lambda}_{ij} = \frac{1}{30}$ for $i \neq j$. The likelihood ratio $L_{t|s}(\theta|\theta')$ can be numerically evaluated as quotient of the transition probabilities of $\{\theta_t\}$ and $\{\hat{\theta}_t\}$. This done by evaluating the matrix exponentials $e^{Q\cdot(t-s)}$ and $e^{\hat{Q}\cdot(t-s)}$, where Q and $\hat{Q}$ are the transition rate matrices of the continuous time Markov chains $\{\theta_t\}$ and $\{\hat{\theta}_t\}$ respectively.

Figure 1 presents the estimated values of rare event probability, obtained by running algorithms listed in table 1. We run 1000 simulations with 1000 particles (500 for Mode 1, 3000 for Mode 2 and 200 for Mode 3) for IPS, HIPS1 and HIPS2 algorithms, and 10^7 simulations for MC approach. The results in figure 1 show that MC stops at $d_4 = 490$ (in Mode 1) and that IPS stops immediately.

Algorithm HIPS 1 allows to avoid loss of light particles in "light" modes, e.g. in mode 3, and helps to maintain fixed mumbler of particles in each mode.

However, it is not able to cope with rare switches and as a result of which it underestimates the rare event probability.

Algorithm HIPS 2 copes well with both the problem of rare switches and the problem of "light" modes. It forces interaction between the modes by making rare switches more frequent and properly adjusting the weights of particles. The results show that the increase of the frequencies of switches has a considerable effect on probabilities of mode 2 and 3 and thus on total probability (see figure 1). This is what one should expect because the heavy particles can leave the mode 1, although very rarely, and have a great influence on modes 2 and 3. The interaction between modes 2 and 3 is not really noticeable but they have an influence on mode 1.

8. CONCLUDING REMARKS

The paper extended the sequential Monte Carlo approach of (Cérou *et al.*, 2002) to estimating rare events of rarely switching diffusions. First we have formulated the approach of (Cérou *et al.*, 2002) to explicitly include the switching diffusion situation. Then we have developed two extensions: sampling per mode to cope with large differences in mode weights, and importance switching to cope with rare mode switching. Next we evaluated the algorithms for a simple example. The algorithms tested on switching diffusion are summarized in Table 1. The obtained results show that all the proposed extensions are in fact needed for estimating rare events for a rarely switching diffusion. The best performing algorithm (HIPS2) is able to cope with differences in weights (sampling per mode), rare switches of discrete component (importance switching) and rare visits of continuous component to the rare target set (decomposition of rare event probability).

REFERENCES

Blom, H.A.P. and E.A. Bloem (2003). Joint IMM-PDA particle filter. In: *Proceedings 6th Int. Conf on Information Fusion*. Cairns, Queensland, Australia. pp. 785–792.

Blom, H.A.P., G.J. Bakker, M.H.C. Everdij and M.N.J. van der Park (2003*a*). Collision risk modeling of air traffic. In: *Proc. European Control Conf.*. Cambridge, UK.

Blom, H.A.P., G.J. Bakker, M.H.C. Everdij and M.N.J. van der Park (2003*b*). Stochastic analysis background of accident risk assessment for air traffic management. In: *IFAC Conference on Analysis and Design of Hybrid Systems*. Saint-Malo Brittany, France.

Cérou, F.P., P. Del Moral, F. LeGland and P. Lezaud (2002). Genetic genealogical models in rare event analysis. Research Report. http://www.nlr.nl/public/hosted-sites/hybridge/.

Doucet, A., N. de Freitas and N.Gordon (2001). *Sequential Monte Carlo Methods in Practice*. Springer-Verlag.

Embrechts, P., C. Kluppelberg and T. Mikosh (1997). *Modeling extrimal events for insurance and finance*. Springer.

Ghosh, M.K., A. Arapostathis and S.I. Marcus (1993). Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM J. Control Optimization* **31**, 1183–1204.

Ghosh, M.K., A. Arapostathis and S.I. Marcus (1997). Ergodic control of switching diffusions. *SIAM J. Control Optimization* **35**, 1952–1988.

Hu, J., M. Prandini and S. Sastry (2003). Probabilistic safety analysis in three-dimensional aircraft flight. In: *Proc. 42nd IEEE CDC*. Maui, USA.

Krystul, J. and A. Bagchi (2004). Approximation of first passage times of switching diffusion. In: *Proc. Sixteenth International Symposium on Mathematical Theory of Networks and Systems*. Leuven, Belgium.

Krystul, J. and H.A.P. Blom (2004). Monte carlo simulation of rare events in hybrid system. hybridge report d8.3. http://www.nlr.nl/public/hosted-sites/hybridge/.

Labeau, P.E., C. Smidts and S. Swaminathan (2000). Dynamic reliability: towards an integrated platform for probabilistic risk assessment. *Reliability Engineering and System Safety* **68**, 219–254.

Liu, J.S. (2003). *Monte Carlo Strategies in Scientific Computing*. Springer.

Moral, P. Del (2004). *Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications*. Springer.

Smidts, C., J. Devooght and P.E. Labeau (1998). Dynamic reliability: Future directions. In: *International Workshop Series on Advanced Topics in Reliability and Risk Analysis*. Greenbelt. MD, USA.

Townsend, J.D., Z. Haraszti, J.A. Freebersyser and M. Devetsikiotis (1998). Simulation of rare events in communications networks. *IEEE Communications Magazine* **36**(8), 36–41.

Watkins, O. and J. Lygeros (2003). Stochastic reachability for discrete time systems: an application to aircraft collision avoidance. In: *Proc. 42nd IEEE CDC*. Maui, USA. pp. 5314–5319.