Snell envelope with small probability criteria

Pierre Del Moral[*], Peng Hu[†], Nadia Oudjane[‡]

May 22, 2012

Abstract

We present a new algorithm to compute the Snell envelope in the specific case where the criteria to optimize is associated with a small probability or a rare event. This new approach combines the Stochastic Mesh approach of Broadie and Glasserman with a particle approximation scheme based on a specific change of measure designed to concentrate the computational effort in regions pointed out by the criteria. The theoretical analysis of this new algorithm provides non asymptotic convergence estimates. Finally, the numerical tests confirm the practical interest of this approach.

Key-words : Snell envelope, American option, Bermudan option, Stochastic Mesh, particle methods, rare events.

[*]Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux , Université de Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France, Pierre.Del-Moral@inria.fr

[†]Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux , Université de Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France, Peng.Hu@inria.fr

[‡]EDF R&D Clamart and FiME (Laboratoire de Finance des Marchés de l'Energie (Dauphine, CREST, EDF R& D) www.fime-lab.org) (nadia.oudjane@edf.fr)

Contents

1 Introduction

The Snell envelope is related to the calculation of the optimal stopping time of a random process based on a given optimality criteria. Several approximation schemes have been proposed recently to numerically compute the Snell envelope. In this paper, we are interested in some specific optimality criteria related to the realization of a small probability or even rare events. In other words, given a random process $(X_k)_{0\leq k\leq n}$ and some *payoff* functions $(f_k)_{0\leq k\leq n}$, we want to maximize an expected gain $\mathbb{E}(f_\tau(X_\tau))$ by choosing τ on a set of random stopping times $\mathcal{T}$. When the payoff functions f_k are localized in a small region of the space, standard Monte Carlo simulations usually fail, because of the difficulty in ensuring enough simulation samples to realize the (relative-)rare events. For example, in finance, when $f(x) = (K - x)^+$, the so-called put option value is difficult to compute when K is much smaller than the initial asset price x_0. In even more complicated cases, we can consider the maximization of $\mathbb{E}(f_\tau(X_\tau)\prod_{k=0}^{\tau-1} B_k(X_k))$ for a given class of functions $(B_k)_{0\leq k\leq n}$ modeling an obstacle. For instance in the case of *barrier options*, $(B_k)_{0\leq k\leq n}$ take the form of indicator functions.

In this paper, we propose a Monte Carlo algorithm to compute the Snell envelope, combining the Stochastic Mesh method introduced by M. Broadie and P. Glasserman [3] and a judicious interacting particle scheme which allows to concentrate the computational effort in the regions of interest w.r.t. the criteria. The principal idea of Broadie-Glasserman model is to operate a change of measure to replace conditional expectations by simple expectations. Besides, the change of measures can also be used with a variance reduction purpose to accelerate Monte Carlo methods. However, in general, the choice of an efficient (in term of variance) change of measure, with an explicit Radon-Nikodym derivative, leading to an *easy-to-simulate* distribution is difficult. Precisely, the authors in [9] proposed an

adaptive scheme based on an original interacting particle algorithm to approximate rare event expectations, allowing us to bypass the tricky steps of guessing a correct change of measure. In the present paper, we extend this adaptive scheme for the recursive computation of the conditional expectations appearing in the context of optimal stopping problems. The main idea of the present paper is then to mix the interacting particle algorithm in [9] with the Stochastic Mesh algorithm of Broadie and Glassserman [3].

This article is organized as follows. In Section 2, notations and generalities on the Snell envelope are presented. Moreover, some specific examples are outlined to motivate the scope of the paper. In Section 3, we introduce a change of measure which allows to concentrate the computational effort in the regions of interest w.r.t. the criteria. In Section 4, we propose an interacting particle scheme to approximate the resulting (changed) measure. Section 5, is devoted to the theoretical analysis of this new Stochastic Mesh algorithm based on an interacting particle scheme. We provide non asymptotic convergence estimates and prove that the resulting estimator is positively biased. Finally, some numerical simulations are performed, in Section 7, showing the practical interest of the proposed algorithm.

2 Preliminary

For the convenience of the reader, we begin by introducing some notations and basic results that will be used throughout the paper.

2.1 Notations

We denote respectively by $\mathcal{P}(E)$, and $\mathcal{B}(E)$, the set of all probability measures on some measurable space $(E,\mathcal{E})$, and the Banach space of all bounded and measurable functions f equipped with the uniform norm $\|f\|$. We let $\mu(f) = \int \mu(dx)\ f(x)$, be the Lebesgue integral of a function $f \in \mathcal{B}(E)$, w.r.t. a measure $\mu \in \mathcal{P}(E)$.
We recall that a bounded integral kernel $M(x,dy)$ from a measurable space $(E,\mathcal{E})$ into an auxiliary measurable space $(E',\mathcal{E}')$ is an operator $f \mapsto M(f)$ from $\mathcal{B}(E')$ into $\mathcal{B}(E)$ such that the functions

$$x \mapsto M(f)(x) := \int_{E'} M(x,dy) f(y)$$

are $\mathcal{E}$-measurable and bounded, for any $f \in \mathcal{B}(E')$. In the above displayed formulae, dy stands for an infinitesimal neighborhood of a point y in E'. Sometimes, for indicator functions $f = 1_A$, with $A \in \mathcal{E}$, we also use the notation $M(x,A) := M(1_A)(x)$. The kernel M also generates a dual operator $\mu \mapsto \mu M$ from $\mathcal{M}(E)$ into $\mathcal{M}(E')$ defined by $(\mu M)(f) := \mu(M(f))$. A Markov kernel is a positive and bounded integral operator M with $M(1) = 1$. Given a pair of bounded integral operators (M_1, M_2), we let $(M_1 M_2)$ be the composition operator defined by $(M_1 M_2)(f) = M_1(M_2(f))$. Given a sequence of bounded integral operators M_n from some state space E_{n-1} into another E_n, we set $M_{k,l} := M_{k+1} M_{k+2} \cdots M_l$, for any $k \leq l$, with the convention $M_{k,k} = Id$, the identity operator. In the context of finite state spaces, these integral operations coincide with the traditional matrix operations on multi-dimensional state spaces.
We also assume that the reference Markov chain X_n with initial distribution $\eta_0 \in \mathcal{P}(E_0)$, and elementary transitions $M_n(x_{n-1}, dx_n)$ from E_{n-1} into E_n is defined on some filtered

probability space $(\Omega, \mathcal{F}, \mathbb{P}_{\eta_0})$, and we use the notation $\mathbb{E}_{\mathbb{P}_{\eta_0}}$ to denote the expectations w.r.t. $\mathbb{P}_{\eta_0}$. In this notation, for all $n \geq 1$ and for any $f_n \in \mathcal{B}(E_n)$, we have that

$$\mathbb{E}_{\mathbb{P}_{\eta_0}}\{f_n(X_n)|\mathcal{F}_{n-1}\} = M_n f_n(X_{n-1}) := \int_{E_n} M_n(X_{n-1}, dx_n)\ f_n(x_n)$$

with the σ-field $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ generated by the sequence of random variables X_p, from the origin $p = 0$ up to the time $p = n$. We also use the conventions $\prod_{\emptyset} = 1$, and $\sum_{\emptyset} = 0$.

2.2 Robustness Lemma

In the discrete time setting, the Snell envelope are defined in terms of a given Markov process $(X_k)_{k\geq 0}$ taking values in some sequence of measurable state spaces $(E_n, \mathcal{E}_k)_{k\geq 0}$ adapted to the natural filtration $\mathcal{F} = (\mathcal{F}_k)_{k\geq 0}$. We let $\eta_0 = \text{Law}(X_0)$ be the initial distribution on E_0, and we denote by $M_k(x_{k-1}, dx_k)$ the elementary Markov transition of the chain from E_{k-1} into E_k. For a given time horizon n and any $k \in \{0, \ldots, n\}$, we let $\mathcal{T}_k$ be the set of all stopping times τ taking values in $\{k, \ldots, n\}$. For a given sequence of non negative measurable functions f_k on E_k, we define a target process $Z_k = f_k(X_k)$. Then $(U_k)_{0\leq k\leq n}$ the Snell envelope of process $(Z_k)_{0\leq k\leq n}$ is defined by a recursive formula:

$$U_k = Z_k \vee \mathbb{E}(U_{k+1}|\mathcal{F}_k)$$

with terminal condition $U_n = Z_n$. The main property of the Snell envelope defined as above is

$$U_k = \sup_{\tau\in\mathcal{T}_k} \mathbb{E}(Z_\tau|\mathcal{F}_k) = \mathbb{E}(Z_{\tau_k^*}|\mathcal{F}_k) \ \text{ with } \ \tau_k^* = \min\{k \leq j \leq n \ :\ U_j = Z_j\} \in \mathcal{T}_k\ .$$

Then the computation of the Snell envelope $(U_k)_{0\leq k\leq n}$ amounts to solving the following backward functional equation.

$$u_k = f_k \vee M_{k+1}(u_{k+1}) \tag{2.1}$$

for any $0 \leq k < n$ with the terminal condition $u_n = f_n$.

But at this level of generality, we can hardly have a closed solution of the function u_k. In this context, lots of numerical approximation schemes have been proposed. Most of them amount to replacing in recursion (2.1) the pair of functions and Markov transitions $(f_k, M_k)_{0\leq k\leq n}$ by some approximation model $(\widehat{f}_k, \widehat{M}_k)_{0\leq k\leq n}$ on some possibly reduced measurable subsets $\widehat{E}_k \subset E_k$. In paper [10], the authors provided the following robustness lemma to estimate the error related to the resulting approximation $\widehat{u}_k$ of the Snell envelope u_k, for several types of approximation models $(\widehat{f}_k, \widehat{M}_k)_{0\leq k\leq n}$.

Lemma 2.1 *For any $0 \leq k < n$, on the state space $\widehat{E}_k$, we have that*

$$|u_k - \widehat{u}_k| \leq \sum_{l=k}^{n} \widehat{M}_{k,l}|f_l - \widehat{f}_l| + \sum_{l=k}^{n-1} \widehat{M}_{k,l}|(M_{l+1} - \widehat{M}_{l+1})u_{l+1}|\ ,$$

where $\widehat{M}_{k,l}$ is the composition operator defined as $\widehat{M}_{k,l} = \widehat{M}_{k+1}\widehat{M}_{k+2}\ldots\widehat{M}_l$, for any $k \leq l$.

This lemma provides a natural way to compare and combine different approximation models. In the present paper, this Lemma will be applied in the specific framework for the small probability criteria.

2.3 Motivations

The choice of nonhomogeneous state spaces $\mathbf{E}_n$ is not innocent. In several application areas the underlying Markov model is a path-space Markov chain:

$$\mathbf{X}_n = (X_0, \ldots, X_n) \in \mathbf{E}_n = (E_0 \times \ldots \times E_n) \ . \tag{2.2}$$

The elementary prime variables X_n represent an elementary Markov chain with Markov transitions $M_k(x_{k-1}, dx_k)$ from E_{k-1} into E_k. In this situation, the historical process $\mathbf{X}_n$ can be seen as a Markov chain with transitions given for any $\mathbf{x}_{k-1} = (x_0, \ldots, x_{k-1}) \in \mathbf{E}_{k-1}$ and $\mathbf{y}_k = (y_0, \ldots, y_k) \in \mathbf{E}_k$ by the following formula

$$\mathbf{M}_k(\mathbf{x}_{k-1}, d\mathbf{y}_k) = \delta_{\mathbf{x}_{k-1}}(d\mathbf{y}_{k-1}) \ M_k(y_{k-1}, dy_k) \ .$$

As we will see in this sequel, this path space framework is, for instance, well suited when dealing with path dependent options as Asian options or Barrier options. Besides, this path space framework is also well suited for the analysis of the Snell envelope under different probability measures.

The multiplicatively path dependent case Now come back to the multiplicatively path dependent Snell envelope that we mentioned in the introduction and formalize the the path space model. For a given collection of real valued functions $(f_k)_{0\leq k\leq n}$ and $(B_k)_{0\leq k\leq n}$, defined on $(E_k)_{0\leq k\leq n}$, we define a class of real valued functions $(F_k)_{0\leq k\leq n}$ defined on the path spaces $(\mathbf{E}_k)_{0\leq k\leq n}$ by

$$F_k(\mathbf{x}_k) := f_k(x_k) \prod_{0\leq p\leq k-1} B_p(x_p) \ , \quad \text{for all } 0 \leq k \leq n \ ,$$

for all $\mathbf{x}_k = (x_0, \cdots, x_k) \in \mathbf{E}_k$. Instead of $\mathbb{E}(f_\tau(X_\tau))$ we want to maximize the expected gain $\mathbb{E}(F_\tau(\mathbf{X}_\tau))$ w.r.t. τ in a set of random stopping times $\mathcal{T}$. In other words, one is interested in computing the Snell envelope $(\mathbf{u}_k)_{0\leq k\leq n}$ associated to the gain functions $(F_k)_{0<k\leq n}$; it satisfies the recursion:

$$\left\{ \begin{array}{l} \mathbf{u}_n(\mathbf{x}_n) = F_n(\mathbf{x}_n) \\ \mathbf{u}_k(\mathbf{x}_k) = F_k(\mathbf{x}_k) \vee \mathbf{M}_{k+1}(\mathbf{u}_{k+1})(\mathbf{x}_k), \forall \ 0 \leq k \leq n-1 \ . \end{array} \right. \tag{2.3}$$

At this stage, two difficulties may arise. First, the above recursion seems to require the approximation of high dimensional conditional expectations, defined on the path spaces E_k, at each time step from $k = n-1$ up to $k = 0$. Second, when the optimality criteria B_p is localized in a specific region of E_p, for each p, then the product $\prod_{p=0}^{k-1} B_p(x_p)$ can be interpreted as a rare event. Hence, at first glance, the computation of Snell envelopes in the multiplicatively path dependent case seems to combine two additional numerical difficulties w.r.t. the standard case, related to the computation of conditional expectations in both high dimensional and rare event situations. The dimensionality problem is easily bypassed by considering an intermediate standard Snell envelope without path dependent criteria, which is directly related to the multiplicatively path dependent Snell envelope. Indeed, consider the standard (non path dependent) Snell envelope $(v_k)_{0\leq k\leq n}$ satisfying the following recursion:

$$\left\{ \begin{array}{rcl} v_n(x_n) & = & f_n(x_n) \\ v_k(x_k) & = & f_k(x_k) \vee \left[B_k(x_k) M_{k+1}(v_{k+1})(x_k) \right] \ , \ \forall \ 0 \leq k \leq n-1 \ . \end{array} \right. \tag{2.4}$$

For all $0 \leq k \leq n$, let us denote by $\mathbf{v}_k$ the real valued functions defined on $\mathbf{E}_k$, such that $\mathbf{v}_k(\mathbf{x}_k) := v_k(x_k) \prod_{p=0}^{k-1} G_p(x_p)$. By construction, one can easily check that for all $0 \leq k \leq n$, $\mathbf{u}_k \equiv \mathbf{v}_k$ and in particular $\mathbf{u}_0(\mathbf{x}_0) = v_0(x_0)$. Indeed, one can verify that $(\mathbf{v}_k)_{0 \leq k \leq n}$ follow the same recursion (2.3) as $(\mathbf{u}_k)_{0 \leq k \leq n}$ and have the same terminal condition. Now that we have underlined the link between $\mathbf{u}_k$ and v_k, the computation of the original Snell envelope $\mathbf{u}_k$ can be done by using one of the many approximation schemes developed for the standard (non path dependent) case.

Besides, to deal with the rare event problem, we propose a change of measure which allows to concentrate the computational effort in the regions of interest w.r.t. the criteria $(B_k)_{0 \leq k \leq n-1}$.

Rare event associated with Payoff function Another Snell envelope problem associated with a small probability event comes from the payoff function when $f(X_n)$ is difficult to simulate. An example arises from the Bermudan put options when the strike K is much smaller than the initial price of the underlying asset. In this case, the standard Monte Carlo approach is not able to concentrate the computational effort in regions where the payoff function $x \mapsto f(x) = (K - x)^+$ does not vanish to zero. In full generality, for a payoff function f concentrated in a relative small region of the space, the choice of an efficient change of measure for computing the recursive conditional expectations is difficult. This problem becomes even more tricky when the number of the underlying assets is greater than three. In the following section, we propose a simple adaptive scheme that allows to approximate an efficient change of measure without requiring any a priori information.

3 Snell envelope and change of measure

Now, recall the reduced Snell envelope for the multiplicatively path dependent case:

$$\left\{ \begin{array}{rcl} v_n(x_n) & = & f_n(x_n) \\ v_k(x_k) & = & f_k(x_k) \vee \left[B_k(x_k) M_{k+1}(v_{k+1})(x_k) \right] \ , \ \forall \ 0 \leq k \leq n-1 \ . \end{array} \right.$$

The above recursion implies that it is not relevant to compute precisely the conditional expectation $M_{k+1}(v_{k+1})(x_k)$ when the value of the criteria $B_k(x_k)$ is zero or very small, or when the gain function f_k is zero or very small. Hence from a variance reduction point of view, when approximating the conditional expectation $M_{k+1}(v_{k+1})(x_k)$ by a Monte Carlo method, it seems relevant to concentrate the simulations in the regions of E_{k+1} where B_{k+1} and/or f_{k+1} reach high values. Hence, to avoid the potential rare events B, we consider a change of measure on the measurable product space $(E_0 \times \cdots \times E_n, \mathcal{E}_0 \times \cdots \times \mathcal{E}_n)$, with the following form

$$d\mathbb{Q}_n = \frac{1}{Z_n} \left[\prod_{k=0}^{n-1} G_k(X_k) \right] d\mathbb{P}_n \ , \quad \text{with} \quad Z_n = \mathbb{E} \left(\prod_{k=0}^{n-1} G_k(X_k) \right) = \prod_{k=0}^{n-1} \eta_k(G_k) \ , \qquad (3.1)$$

where $(G_k)_{0 \leq k < n}$ is a sequence of non-negative functions defined on $(E_k)_{0 \leq k < n}$ (typically $G_k := B_k$, and G_k is written instead of B_k in further development of this article) and η_k is

the probability measure defined on E_k such that, for any measurable function f on E_k

$$\eta_k(f) := \frac{\mathbb{E}\Big(f(X_k)\prod_{p=0}^{k-1} G_p(X_p)\Big)}{\mathbb{E}\Big(\prod_{p=0}^{k-1} G_p(X_p)\Big)} \ .$$

The measures $(\eta_k)_{0\leq k\leq n}$ defined above can be seen as the laws of random states $(\bar{X}_k)_{0\leq k\leq n}$ under the probability measures $(\mathbb{Q}_k)_{0\leq k\leq n}$. More interestingly, in Section 4 we will see that the sequence of random states $(\bar{X}_k)_{0\leq k\leq n}$ forms a nonlinear Markov chain with transitions $\bar{X}_k \rightsquigarrow \bar{X}_{k+1}$ that depends on the current distribution η_k, at time k. The behavior of this chain is dictated by the potential functions $(G_k)_{0\leq k\leq n}$ and the Markov transitions $(M_k)_{\leq k\leq n}$ of the reference process $(X_k)_{0\leq k\leq n}$. Regions with high G_k-values are visited more likely.

To illustrate this remark, we examine the situation where $G_k(x_k) = B_k(x_k) := 1_{A_k}(x_k)$ with $A_k \subset E_k$. In this situation, $law(X_k|X_p \in A_p,\ p < k) = law(\bar{X}_k) = \eta_k$ is the conditional distribution of X_k given the fact that $X_p \in A_p$, for any $p < k$. In this special case, the process $(\bar{X}_k)_{0\leq k\leq n}$ is restricted to regions related to the choice of the sequence $(A_k)_{0\leq k\leq n}$. This change of measure is know as the optimal twisted measure for sampling a Markov chain restricted to the subset regions A_k. More general change of measure are addressed in section 6. These models are direct extension of 3.1 to potential functions that depend on the transition of the reference Markov chain.

When the rare event problem comes from the payoff, we can construct a collection of G_k to force the particle step by step to achieve the payoff. But in this case, there is no more explicit obstacle B_k to help us to construct such potential functions. A choice of G_k is provided in section 7.2. For further reading, readers are referred to [9]. The authors have proposed several choices to minimize the variance.

At this stage, it is important to emphasize that the analysis of the both case where the choice of G_k is explicit or not, are mathematically equivalent. The only difference comes from the fact that the recursion 2.4 has additional term B_k compared to 2.1. And the mathematical analysis of the later is easier and can be induced directly from the former (by deleting all the B_k appeared in the Snell envelope recursion in the analysis). So only the analysis of the multiplicatively path dependent case are provided in this paper.

Furthermore, it is also important to observe that, for any measurable function f on E_k

$$\eta_k(f) = \frac{\eta_{k-1}(G_{k-1}M_k(f))}{\eta_{k-1}(G_{k-1})} \ . \tag{3.2}$$

We denote the recursive relation between η_k and η_{k-1} by introducing the operators Φ_k such that, for all $1 \leq k \leq n$

$$\eta_k = \Phi_k(\eta_{k-1}) \ . \tag{3.3}$$

Let us now introduce the integral operator Q_k such that, for all $1 \leq k \leq n$

$$Q_k(f)(x_{k-1}) := \int G_{k-1}(x_{k-1})M_k(x_{k-1}, dx_k)f(x_k) \ . \tag{3.4}$$

In further developments of this article, we suppose that $M_k(x_{k-1}, \cdot)$ are equivalent to some measures λ_k, for any $0 \leq k \leq n$ and $x_{k-1} \in E_{k-1}$, i.e. there exists a collection of positive functions H_k and measures λ_k such that:

$$M_k(x_{k-1}, dx_k) = H_k(x_{k-1}, x_k)\lambda_k(dx_k) \ . \tag{3.5}$$

Now, we are in a position to state the following Lemma.

Lemma 3.1 *For any measure η on E_k, recursion (2.4) defining v_k can be rewritten:*

$$v_k(x_k) = f_k(x_k) \vee Q_{k+1}(v_{k+1})(x_k) = f_k(x_k) \vee \Phi_{k+1}(\eta) \left(\frac{dQ_{k+1}(x_k, \cdot)}{d\Phi_{k+1}(\eta)} v_{k+1} \right) ,$$

for any $x_k \in E_k$, where

$$\frac{dQ_{k+1}(x_k, \cdot)}{d\Phi_{k+1}(\eta)}(x_{k+1}) = \frac{G_k(x_k) H_{k+1}(x_k, x_{k+1}) \eta(G_k)}{\eta(G_k H_{k+1}(\cdot, x_{k+1}))} ,$$

for any $(x_k, x_{k+1}) \in E_k \times E_{k+1}$.

Proof:
Under Assumption (3.5), we have immediately the following formula

$$M_{k+1}(x_k, dx_{k+1}) = H_{k+1}(x_k, x_{k+1}) \frac{\eta_k(G_k)}{\eta_k(G_k H_{k+1}(\cdot, x_{k+1}))} \eta_{k+1}(dx_{k+1}) . \tag{3.6}$$

Now, note that the above equation is still valid for any measure η,

$$M_{k+1}(x_k, dx_{k+1}) = H_{k+1}(x_k, x_{k+1}) \frac{\eta(G_k)}{\eta(G_k H_{k+1}(\cdot, x_{k+1}))} \Phi_{k+1}(\eta)(dx_{k+1}) . \tag{3.7}$$

Hence, the Radon Nikodym derivative of $M_{k+1}(x_k, dx_{k+1})$ w.r.t. $\Phi_{k+1}(\eta)$ is such that

$$\frac{dM_{k+1}(x_k, \cdot)}{d\Phi_{k+1}(\eta)}(x_{k+1}) = H_{k+1}(x_k, x_{k+1}) \frac{\eta(G_k)}{\eta(G_k H_{k+1}(\cdot, x_{k+1}))} . \tag{3.8}$$

We end the proof by applying the arguments above to recursion (2.4).

4 A particle approximation scheme

In this section, we first propose a particle model to sample the random variables according to these distributions. This sample scheme is then combined with the Stochastic Mesh scheme to finally provide an original particle algorithm to approximate the Snell envelope $(v_k)_{0 \leq k \leq n}$.
By definition (3.3) of Φ_{k+1}, we have the following formula

$$\Phi_k(\eta_{k-1}) = \eta_{k-1} K_{k,\eta_{k-1}} = \eta_{k-1} S_{k-1,\eta_{k-1}} M_k = \Psi_{G_{k-1}}(\eta_{k-1}) M_k . \tag{4.1}$$

Where $K_{k,\eta_{k-1}}$, $S_{k-1,\eta_{k-1}}$ and $\Psi_{G_{k-1}}$ are defined as follows:

$$\begin{cases} K_{k,\eta_{k-1}}(x_{k-1}, dx_k) &= (S_{k-1,\eta_{k-1}} M_k)(x_{k-1}, dx_k) \\ &= \int S_{k-1,\eta_{k-1}}(x_{k-1}, dx'_{k-1}) M_k(x'_{k-1}, dx_k) , \\ S_{k-1,\eta_{k-1}}(x, dx') &= \epsilon G_{k-1}(x) \delta_x(dx') + (1 - \epsilon G_{k-1}(x)) \Psi_{G_{k-1}}(\eta_{k-1})(dx') \\ \Psi_{G_{k-1}}(\eta_{k-1})(dx) &= \frac{G_{k-1}(x)}{\eta_{k-1}(G_{k-1})} \eta_{k-1}(dx) , \end{cases}$$

where the real ϵ is such that ϵG takes its values in $[0,1]$.
More generally, the operations Ψ and S can be expressed as $\Psi_G(\eta)(f) = \frac{\eta(Gf)}{\eta(G)} = \eta S_\eta(f)$ with $S_\eta(f) = \epsilon Gf + (1-\epsilon G)\Psi_G(\eta)(f)$. We recall from [8] that $\eta_k = law(\bar{X}_k)$, where $\bar{X}_{k-1} \rightsquigarrow \bar{X}_k$ is a Markov chain with transitions $K_{k,\eta_{k-1}}$ defined above.

The particle approximation provided in the present paper is defined in terms of a Markov chain $\xi_k^{(N)} = (\xi_k^{(i,N)})_{1\leq i\leq N}$ on the product state spaces E_k^N, where the given integer N is the number of particles sampled in every instant. The initial particle system, $\xi_0^{(N)} = \left(\xi_0^{(i,N)}\right)_{1\leq i\leq N}$, is a collection of N i.i.d. random copies of X_0. We let $\mathcal{F}_k^N$ be the sigma-field generated by the particle approximation model from the origin, up to time k. To simplify the presentation, when there is no confusion we suppress the population size parameter N, and we write ξ_k and ξ_k^i instead of $\xi_k^{(N)}$ and $\xi_k^{(i,N)}$. By construction, ξ_k is a particle model with a selection transition and a mutation type exploration i.e. the evolution from ξ_k to ξ_{k+1} is composed by two steps:

$$\xi_k \in E_k^N \xrightarrow[S_{k,\eta_k^N}]{\text{Selection}} \widehat{\xi}_k := \left(\widehat{\xi}_k^i\right)_{1\leq i\leq N} \in E_k^N \xrightarrow[M_{k+1}]{\text{Mutation}} \xi_{k+1} \in E_{k+1}^N \ . \tag{4.2}$$

Then we define η_k^N and $\widehat{\eta}_k^N$ as the occupation measures after the mutation and the selection steps. More precisely,

$$\eta_k^N := \frac{1}{N}\sum_{1\leq i\leq N} \delta_{\xi_k^i} \quad \text{and} \quad \widehat{\eta}_k^N := \frac{1}{N}\sum_{1\leq i\leq N} \delta_{\widehat{\xi}_k^i} \ .$$

During the selection transition S_{k,η_k^N}, for $0\leq i\leq N$ with a probability $\epsilon G_k(\xi_k^i)$ we decide to skip the selection step i.e. we let $\widehat{\xi}_k^i$ stay on particle ξ_k^i, and with probability $1-\epsilon G_k(\xi_k^i)$ we decide to do the following selection: $\widehat{\xi}_k^i$ randomly takes the value in ξ_k^j for $0\leq j\leq N$ with distribution $\frac{G_k(\xi_k^j)}{\sum_{l=1}^N G_k(\xi_k^l)}$. Note that when $\epsilon G_k \equiv 1$, the selection is skipped (i.e. $\widehat{\xi}_k = \xi_k$) so that the model corresponds exactly to the Broadie-Glasserman type model analysed by P. Del Moral et al. [10]. Hence, the factor ϵ can be interpreted as a level of selection against the rare events.
During the mutation transition $\widehat{\xi}_k \rightsquigarrow \xi_{k+1}$, every selected individual $\widehat{\xi}_k^i$ evolves randomly to a new individual $\xi_{k+1}^i = x$ randomly chosen with the distribution $M_{k+1}(\widehat{\xi}_k^i, dx)$, for $1\leq i\leq N$. It is important to observe that by construction, η_{k+1}^N is the empirical measure associated with N conditionally independent and identically distributed random individual ξ_{k+1}^i with common distribution $\Phi_{k+1}(\eta_k^N)$.

Now, we are in a position to describe precisely the new approximation scheme proposed to estimate the Snell envelope $(v_k)_{0\leq k\leq n}$. The main idea consists in taking $\eta = \eta_k^N$, in Lemma 3.1, then observing that Snell envelope $(v_k)_{0\leq k\leq n}$ is solution of the following recursion, for all $0\leq k<n$,

$$v_k(x_k) = f_k(x_k) \vee \Phi_{k+1}(\eta_k^N)\left(\frac{dQ_{k+1}(x_k,\cdot)}{d\Phi_{k+1}(\eta_k^N)} v_{k+1}\right) \ .$$

Now, if $\Phi_{k+1}(\eta_k^N)$ is well estimated by η_{k+1}^N, it is relevant to approximate v_k by $\widehat{v}_k$ defined

by the following backward recursion

$$\left\{ \begin{array}{rcl} \widehat{v}_n & = & f_n \\ \widehat{v}_k(x_k) & = & f_k(x_k) \vee \eta_{k+1}^N \left(\dfrac{dQ_{k+1}(x_k,\cdot)}{d\Phi_{k+1}(\eta_k^N)} \widehat{v}_{k+1} \right) \quad \text{for all } 0 \leq k < n \ , \end{array} \right. \tag{4.3}$$

Note that in the above formula (4.3), the function v_k is defined not only on E_k^N but on the whole state space E_k.
To simplify notations, we set

$$\widehat{Q}_{k+1}(x_k, dx_{k+1}) = \eta_{k+1}^N(dx_{k+1}) \frac{dQ_{k+1}(x_k,\cdot)}{d\Phi_{k+1}(\eta_k^N)}(x_{k+1}) \ .$$

Finally, with this notation, the real Snell envelope $(v_k)_{0\leq k\leq n}$ and the approximation $(\hat{v}_k)_{0\leq k\leq n}$ are such that, for all $0 \leq k < n$,

$$\begin{array}{rcl} v_k & = & f_k \vee Q_{k+1}(v_{k+1}) \\ \widehat{v}_k & = & f_k \vee \widehat{Q}_{k+1}(\widehat{v}_{k+1}) \ . \end{array}$$

In the change of measure interpretation presented in section 3, the particle algorithm developed above can be seen as a stochastic acceptance-rejection technique with recycling transitions. This type of particle sampling model has been used in other contexts, including financial risk analysis in [4, 6]. For an overview of these novel particle algorithms in financial mathematics, we refer the interested reader to the book [5].

5 Convergence and bias analysis

By the previous construction, we can approximate $\Phi_{k+1}(\eta_k^N)$ by η_{k+1}^N. In this section, we will first analyze the error associated with that approximation and then derive an error bound for the resulting Snell envelope approximation scheme. To simplify notations, in further development, we consider the random fields V_k^N defined as

$$V_k^N := \sqrt{N} \ \left(\eta_k^N - \Phi_k(\eta_{k-1}^N) \right) \ .$$

The following lemma shows the conditional zero-bias property and mean error estimates for the approximation η_{k+1}^N of $\Phi_{k+1}(\eta_k^N)$.

Lemma 5.1 *For any integer $p \geq 1$, we denote by p' the smallest even integer greater than p. In this notation, for any $0 \leq k \leq n$ and any integrable function f on E_{k+1}, we have*

$$\mathbb{E}\left(\eta_{k+1}^N(f) | \mathcal{F}_k^N \right) = \Phi_{k+1}(\eta_k^N)(f)$$

and

$$\mathbb{E}\left(\left| V_k^N(f) \right|^p | \mathcal{F}_k^N \right)^{\frac{1}{p}} \leq 2 \ a(p) \ \left[\Phi_{k+1}(\eta_k^N)(|f|^{p'}) \right]^{\frac{1}{p'}}$$

with the collection of constants

$$a(2p)^{2p} = (2p)_p \ 2^{-p} \quad \textit{and} \quad a(2p+1)^{2p+1} = \frac{(2p+1)_{p+1}}{\sqrt{p+1/2}} \ 2^{-(p+1/2)} \ .$$

Proof : The conditional zero-bias property is easily proved as follows

$$\begin{aligned}\mathbb{E}\left(\eta_{k+1}^N(f)|\eta_k^N\right) &= \frac{1}{N}\sum_{i=1}^N \mathbb{E}(f(\xi_{k+1}^i)|\eta_k^N) \\ &= \frac{1}{N}\sum_{i=1}^N K_{k+1,\eta_k^N}(f)(\xi_k^i) \\ &= (\eta_k^N K_{k+1,\eta_k^N})(f) = \Phi_{k+1}(\eta_k^N)(f) \ .\end{aligned}$$

Then the above equality implies

$$\mathbb{E}\left(\left|\left[\eta_{k+1}^N - \Phi_{k+1}(\eta_k^N)\right](f)\right|^p |\mathcal{F}_k^N\right)^{\frac{1}{p}} \leq \mathbb{E}\left(\left|\left[\eta_{k+1}^N - \mu_{k+1}^N\right](f)\right|^p |\mathcal{F}_k^N\right)^{\frac{1}{p}} \ ,$$

where $\mu_{k+1}^N := \frac{1}{N}\sum_{i=1}^N \delta_{Y_{k+1}^i}$ stands for an independent copy of η_{k+1}^N given η_k^N. Using Khintchine's type inequalities yields that

$$\begin{aligned}\sqrt{N}\ \mathbb{E}\left(\left|[\eta_{k+1}^N - \mu_{k+1}^N](f)\right|^p |\mathcal{F}_k^N\right)^{\frac{1}{p}} &\leq 2\ a(p)\ \mathbb{E}\left(\left|f\left(\xi_{k+1}^1\right)\right|^{p'} \mid \mathcal{F}_k^N\right)^{\frac{1}{p'}} \\ &= 2\ a(p)\ \left[\Phi_{k+1}(\eta_k^N)(|f|^{p'})\right]^{\frac{1}{p'}} \ .\end{aligned}$$

We refer the reader to lemma 7.3.3 on page 223 of [8], for the proof of this kind of Khintchine's inequalities. We end the proof by combining the above two inequalities.

■

A consequence of the zero-bias property proved in Lemma 5.1 is that

$$\mathbb{E}(\widehat{Q}_{k+1}(f)(x_k)|\eta_k^N) = Q_{k+1}(f)(x_k) \ .$$

To estimate the error between v_k and the approximation $\hat{v}_k$, it is useful to introduce the following random integral operator R_k^N such that for any measurable function on E_{k+1},

$$R_{k+1}^N(f)(x_k) = \sqrt{N}\left(\widehat{Q}_{k+1}(f)(x_k) - Q_{k+1}(f)(x_k)\right) \ .$$

Note that

$$R_{k+1}^N(f)(x_k) := \int V_{k+1}^N(dx_{k+1})\ \frac{dQ_{k+1}(x_k,.)}{d\Phi_{k+1}(\eta_k^N)}(x_{k+1})\ f(x_{k+1}) \ ,$$

then, applying again Lemma 5.1 implies the following Khintchine's type inequality

$$\begin{aligned}&\mathbb{E}(\left|R_{k+1}^N(v_{k+1})(x_k)\right|^p |\eta_k^N)^{\frac{1}{p}} \\ &\leq 2\ a(p)\left[\int_{E_{k+1}} \Phi_{k+1}(\eta_k^N)(dx_{k+1})\left(\frac{dQ_{k+1}(x_k,\cdot)}{d\Phi_{k+1}(\eta_k^N)}(x_{k+1})v_{k+1}(x_{k+1})\right)^{p'}\right]^{\frac{1}{p'}}\end{aligned}$$

Let $\widehat{Q}_{k,l} = \widehat{Q}_{k+1}\widehat{Q}_{k+2}\ldots\widehat{Q}_l$ for any $0 \leq k < l \leq n$, then it follows easily, by recursion, that

$$\mathbb{E}(\widehat{Q}_{k,l}(f)(x_k)|\eta_k^N) = Q_{k,l}(f)(x_k) \ .$$

Now, by Lemma 2.1, we conclude

$$\sqrt{N}\,|(v_k - \widehat{v}_k)| \le \sum_{k<l<n} \widehat{Q}_{k,l}|(R^N_{l+1})(v_{l+1})| \ . \tag{5.1}$$

We are now in position to state the main result of this paper.

Theorem 5.2 *For any $0 \le k \le n$ and any integer $p \ge 1$, we have*

$$\sup_{x\in E_k} \|(\widehat{v}_k - v_k)(x)\|_{L_p} \le \sum_{k<l<n} \frac{2\ a(p)}{\sqrt{N}} q_{k,l}\ \left[Q_{k,l+1}(h^{p'-1}_{l+1} v^{p'}_{l+1})(x)\right]^{\frac{1}{p'}} \ ,$$

with a collection of constants $q_{k,l}$ and functions h_k defined as

$$q_{k,l} := \left[\|h_{k+1}\| \prod_{m=k}^{l} \|G_m\|\right]^{\frac{p'-1}{p'}} \quad \textit{and} \quad h_k(x_k) := \sup_{x,y\in E_{k-1}} \frac{H_k(x,x_k)}{H_k(y,x_k)} \ . \tag{5.2}$$

Proof : First, decomposition (5.1) yields

$$\sqrt{N}\,\|(\widehat{v}_k - v_k)(x)\|_{L_p} \le \sum_{k<l<n} \left\|\widehat{Q}_{k,l}|(R^N_{l+1})(v_{l+1})|(x)\right\|_{L_p} \ , \quad \text{for all } x \in E_k \ .$$

Note that

$$\|\widehat{Q}_{k,l}(1)\| \le b_{k,l} \ , \quad \text{where} \quad b_{k,l} := \|h_{k+1}\| \prod_{m=k}^{l-1} \|G_m\| \ .$$

Then it follows easily that for any integrable function f on E_l

$$(\widehat{Q}_{k,l}(f))^p \le (b_{k,l})^{p-1} \widehat{Q}_{k,l}(f^p) \ .$$

This yields that

$$\left\|\widehat{Q}_{k,l}\left|(R^N_{l+1}))(v_{l+1})\right|(x)\right\|_{L_p} \le (b_{k,l})^{\frac{p-1}{p}} \mathbb{E}\left(\widehat{Q}_{k,l}\left(\left|(R^N_{l+1}))(v_{l+1})\right|^p(x)\right)^{\frac{1}{p}} \ .$$

Applying Lemma 5.1 to the right-hand side of the above inequality, we obtain for any $x_l \in E_l$

$$\begin{aligned}
&\mathbb{E}\left(\left|(R^N_{l+1}))(v_{l+1})(x_l)\right|^p |\eta^N_l\right)^{\frac{1}{p}} \\
&\quad \le 2\ a(p) \left[\int_{E_{l+1}} \Phi_{l+1}(\eta^N_l)(dx_{l+1}) \left(\frac{dQ_{l+1}(x_l,\cdot)}{d\Phi_{l+1}(\eta^N_l)}(x_{l+1}) v_{l+1}(x_{l+1})\right)^{p'}\right]^{\frac{1}{p'}}
\end{aligned}$$

from which we find that

$$\begin{aligned}
&\mathbb{E}\left(\left|(R^N_{l+1}))(v_{l+1})(x_l)\right|^p |\eta^N_l\right)^{\frac{1}{p}} \\
&\quad \le 2\ a(p) \left[\int_{E_{l+1}} Q_{l+1}(x_l, dx_{l+1}) \left(\frac{dQ_{l+1}(x_l,\cdot)}{d\Phi_{l+1}(\eta^N_l)}(x_{l+1})\right)^{p'-1} v_{l+1}(x_{l+1})^{p'}\right]^{\frac{1}{p'}}
\end{aligned}$$

By definition (5.2) of functions h_{l+1} and in developing the Radon Nikodym derivative, we obtain

$$\frac{dQ_{l+1}(x_l,\cdot)}{d\Phi_{l+1}(\eta_l^N)}(x_{l+1}) = \frac{\eta_l^N(G_l)G_l(x_l)H_{l+1}(x_l,x_{l+1})}{\eta_l^N(G_lH_{l+1})(\cdot,x_{l+1})} \leq \|G_l\| h_{l+1}(x_{l+1}) \ ,$$

which implies

$$\begin{aligned}
&\mathbb{E}\left(\left|(R_{l+1}^N))(v_{l+1})(x_l)\right|^p |\eta_l^N\right)^{\frac{1}{p}} \\
&\quad \leq 2\ a(p)\|G_l\|^{\frac{p'-1}{p'}} \left[\int_{E_{l+1}} Q_{l+1}(x_l,dx_{l+1})\,(h_{l+1}(x_{l+1}))^{p'-1}\, v_{l+1}(x_{l+1})^{p'}\right]^{\frac{1}{p'}}
\end{aligned}$$

Gathering the above arguments, we conclude that

$$\|(\widehat{v}_k - v_k)(x)\|_{L_p} \leq \sum_{k<l<n} \frac{2\ a(p)}{\sqrt{N}} q_{k,l}\ \left(Q_{k,l+1}(h_{l+1}^{p'-1}v_{l+1}^{p'})(x)\right)^{\frac{1}{p'}} \ .$$

■

Remarks : The constants $q_{k,l}$ could be largely reduced. In fact, $q_{k,l}$ comes from bounding $\|\prod_m \eta_m^N(G_m)\|_{L_p}$. In [7], the authors proved $\|\prod_m G_m\|_{L_2} + \frac{constant}{N}$ as a non asymptotic boundary for $\|\prod_m \eta_m^N(G_m)\|_{L_2}$. In most cases, the functions G take their values in $[0,1]$, then the boundary $\|\prod_m G_m\| \leq 1$ holds, but $\|\prod_m G_m\|_{L_2}$ is very small.

When the function G vanishes in some regions of the state space, we also mention that the particle model is only defined up to the first time $\tau^N = k$ such that $\eta_k^N(G_k) = 0$. We can prove that the event $\{\tau^N \leq n\}$ has an exponentially small probability to occur, with the number of particles N. In fact, the estimates presented in the above theorems can be extended to this singular situation by replacing $\widehat{v}_k$ by the particle estimates $\widehat{v}_k 1_{\tau^N \geq n}$. The stochastic analysis of these singular models are quite technical, for further details we refer the reader to section 7.2.2 and section 7.4 in the book [8].

It is also very natural to assume the functions $(v_k)_{0\leq k\leq n}$ are bounded by M in the sense that

$$\left(Q_{k,l+1}(v_{l+1}^p)(x)\right)^{\frac{1}{p}} < M$$

, for any integer p. Then a new weak bound

$$\frac{2\ a(p)\ (n-k)}{\sqrt{N}} M\left(1 \vee \left(\|h\|^2\|G\|^{n-k}\right)\right)$$

is provided to simplify the notations, where $\|h\| = \max_k \|h_k\|$ and $\|G\| = \max_k \|G_k\|$ To understand better the $\mathbb{L}_p$-mean error bounds in the theorem, we deduce the following exponential concentration inequality:

Proposition 5.3 *For any $0 \leq k \leq n$ and any $\epsilon > 0$, we have*

$$\sup_{x\in E_k} \mathbb{P}\left(|v_k(x) - \widehat{v}_k(x)| > \frac{c}{\sqrt{N}} + \epsilon\right) \leq exp\left(-N\epsilon^2/c^2\right) \ , \tag{5.3}$$

with constant $c = 2(n-k)M\left(1 \vee \left(\|h_k\|^2\|G\|^{n-k}\right)\right)$.

Proof : This result is a direct consequence from the fact that for any non negative random variable U such that

$$\exists b < \infty \text{ s.t. } \forall r \geq 1 \qquad \mathbb{E}\left(U^r\right)^{\frac{1}{r}} \leq a(r)\ b \ \Rightarrow \mathbb{P}\left(U \geq b + \epsilon\right) \leq \exp\left(-\epsilon^2/(2b^2)\right) .$$

To check this claim, we develop the exponential and verify that

$$\forall t \geq 0 \ \ \mathbb{E}\left(e^{tU}\right) \leq \exp\left(\frac{(bt)^2}{2} + bt\right) \Rightarrow \mathbb{P}(U \geq b + \epsilon) \leq \exp\left(-\sup_{t \geq 0}(\epsilon t - \frac{(bt)^2}{2})\right)$$

■

Similarly to Broadie-Glasserman model, the following proposition shows that in this model we also over-estimate the Snell envelope.

Proposition 5.4 *For any $0 \leq k \leq n$ and any $x_k \in E_k$*

$$\mathbb{E}\left(\widehat{v}_k(x_k)\right) \geq v_k(x_k) \ . \tag{5.4}$$

Proof:
We can easily prove this inequality with a simple backward induction. The terminal condition $\widehat{v}_n = v_n$ implies directly the inequality at instant n. Assuming the inequality at time $k+1$, then the Jensen's inequality implies

$$\begin{aligned}
\mathbb{E}\left(\widehat{v}_k(x_k)\right) &\geq f_k(x_k) \vee \mathbb{E}\left(\widehat{Q}_{k+1}\widehat{v}_{k+1}(x_k)\right) \\
&= f_k(x_k) \vee \mathbb{E}\left(\int_{E_{k+1}^N} \widehat{Q}_{k+1}(x_k, dx_{k+1})\mathbb{E}\left(\widehat{v}_{k+1}(x_{k+1})|\mathcal{F}_{k+1}^N\right)\right) \ .
\end{aligned}$$

By the induction assumption at time $k+1$, we have

$$\begin{aligned}
\mathbb{E}\left(\int_{E_{k+1}^N} \widehat{Q}_{k+1}(x_k, dx_{k+1})\mathbb{E}\left(\widehat{v}_{k+1}(x_{k+1})|\mathcal{F}_{k+1}^N\right)\right) &\geq \mathbb{E}\left(\widehat{Q}_{k+1}v_{k+1}(x_k)\right) \\
&= Q_{k+1}v_{k+1}(x_k) \ .
\end{aligned}$$

Then the inequality still holds at time k, which completes the proof.

■

6 Applications and extensions

In this section, we apply the Feynman-Kac methodology developed in section 4 to two type of importance sampling Monte Carlo techniques. We start with some important observations related to potential functions on transitions spaces.

For potential functions $G_k(X_k, X_{k+1})$ depending on the local transitions (X_k, X_{k+1}) of the reference process, the change of measure has the same form as in 3.1, replacing X_k

by the Markov chain $\mathcal{X}_k = (X_k, X_{k+1})$. In this situation, the Snell envelop $\mathbf{v}_k(x_0, \ldots, x_k)$ associated with the payoff functions given bellow:

$$F_k(x_0, \ldots, x_k) = f_k(x_k) \prod_{0 \leq p < k} G_p(x_p, x_{p+1}),$$

has the form

$$\mathbf{v}_k(x_0, \ldots, x_k) = v_k(x_k) \prod_{0 \leq p < k} G_p(x_p, x_{p+1}). \tag{6.1}$$

The sequence of functions $(u_k)_{0 \leq k \leq n}$ satisfies the backward recursion:

$$\begin{aligned} u_n &= f_n \\ u_k(x_p) &= f_p(x_p) \vee \int M_{k+1}(x_k, dx_{k+1}) G_k(x_k, x_{k+1}) u_{k+1}(x_{k+1}). \end{aligned} \tag{6.2}$$

This equation has exactly the same form as 2.4, by replacing the function $B_k(x_k)$ by the function $G_k(x_k, x_{k+1})$.

We illustrate these properties in two situations.

The first one concerns the design of more general change of reference measure. For instance, let us suppose we are given a judicious Markov transition $M'_k(x_{k-1}, x_k)$ such that $M'_k(x_{k-1}, \cdot)$ is absolutely continuous w.r.t. $M_k(x_{k-1}, \cdot)$. In this situation, we have

$$\begin{aligned} &\mathbb{E}(f_n(X_n) \prod_{0 \leq p < n} G_p(X_p)) \\ &\qquad = \mathbb{E}\left(f_n(X'_n) \prod_{0 \leq p < n} \left[G_p(X'_p) \frac{dM_{p+1}(X'_p, \cdot)}{dM'_{p+1}(X'_p, \cdot)}(X'_{p+1}) \right] \right), \end{aligned} \tag{6.3}$$

where $(X'_p)_{0 \leq p \leq n}$ is a Markov chain with initial condition $\eta'_0 = \eta_0 = law(X_0)$, and Markov transitions M'_p. We can rewrite 6.3 as follows:

$$\mathbb{E}(f_n(X_n) \prod_{0 \leq p < n} G_p(X_p)) = \mathbb{E}(f_n(X'_n) \prod_{0 \leq p < n} G'_p(X'_p, X'_{p+1})),$$

with $G'_p(x_p, x_{p+1}) = G_p(x_p) \frac{dM_{p+1}(x_p, \cdot)}{dM'_{p+1}(x_p, \cdot)}(x_{p+1})$.

The second example concerns the design of an importance sampling strategy. Suppose we are given a sequence of positive payoff functions $(f_k)_{0 \leq k \leq n}$, with $f_0 \equiv 1$. In this situation, we have

$$\mathbb{E}(f_n(X_n)) = \mathbb{E}(\prod_{0 \leq p < n} G_p(X_p, X_{p+1}))$$

, with the potential function $G_p(x_p, x_{p+1}) = \frac{f_{p+1}(x_{p+1})}{f_p(x_p)}$. In this context, the Snell envelop 6.1 and 6.2 are given by the backward recursion:

$$\begin{aligned} u_n &= 1 \\ u_p(x_p) &= 1 \vee \int M_{p+1}(x_p, dx_{p+1}) G_p(x_p, x_{p+1}) u_{p+1}(x_{p+1}). \end{aligned}$$

7 Numerical simulations

In this section, we give numerical examples to test our new algorithm, the *Stochastic Mesh with Change of Measure (SMCM)*, on Bermudan options from dimension 1 up to 5, compared with the standard *Stochastic Mesh (SM)* algorithm without change of measure.

7.1 Prices dynamics and options model

In our numerical tests we have considered a simple Black-Scholes price model. However, notice that both algorithms (*SM* and *SMCM*) can be applied in a general Markovian framework. The asset prices are modeled by a d-dimensional Markov process (S_t) such that each component (i.e. each asset) follows a geometric Brownian motion under the risk-neutral measure, that is, for assets $i = 1, \cdots, d$,

$$dS_t(i) = S_t(i)(rdt + \sigma dz_t^i) \ , \tag{7.1}$$

where z^i, for $i = 1, \cdots, d$ are independent one dimensional standard Brownian motions. Unless otherwise specified, the interest rate r is set to 10% annually and the volatility is supposed to be the same for all assets, $\sigma = 20\%$ annually. The starting prices of the assets are for all $i = 1, \cdots, d$, $S_{t_0}(i) = 1$. We consider two types of Bermudan options with maturity $T = 1$ year and 11 equally distributed exercise opportunities at dates $t_k = kT/n$ with $k = 0, 1, \cdots, n = 10$, associated with two different payoffs:

1. Geometric average put option with payoff $(K - \prod_{i=1}^d S_T(i))_+$,
2. Arithmetic average put option with payoff $(K - \frac{1}{d}\sum_{i=1}^d S_T(i))_+$,

Note that the geometric average put payoff involves the process $\prod_{i=1}^d S(i)$ which can be identified to a one-dimensional non standard exponential Brownian motion. For this specific case of geometric put payoff, we chose to vary, in our simulations, the short term interest rate and the volatility with the number of underlying assets d, such that the option value remains the same for all d:

$$r(d) = r/d \ , \quad \text{and} \quad \sigma(d) = \sigma/\sqrt{d} \ . \tag{7.2}$$

Then, we chose as a benchmark value the estimate obtained by the standard Stochastic Mesh approach with $N = 6400$ mesh points for $d = 1$ asset. These benchmark values are reported on Table 1.

Strike	$K = 0.95$	$K = 0.85$	$K = 0.75$
Option value	0.0279	0.0081	0.0015

Table 1: Benchmark values for the geometric put option obtained by using the Stochastic Mesh method with 10000 particles. $n = 11$ exercise opportunities, $T = 1$, $S_0 = 1$ and $r = 10\%/d$, $\sigma_i = 20\%/\sqrt{d}$ for the geometric payoff and $r = 10\%$, $\sigma_i = 20\%$ for the arithmetic payoff.

7.2 Choice of potential functions

We consider the Markov chain $(X_k)_{0\leq k\leq n}$, taking values on $E_k = \mathbb{R}^{+d}$, obtained by discretization of the time-continuous process S defined by (7.1) at times of exercise opportunities, $0 = t_0 < \cdots < t_n = T$, such that for all $k = 0, \cdots, n$, $X_k = S_{t_k}$.
Now, we can introduce the sequence of positive functions $(G_k)_{1\leq k\leq n}$, defining the change of measure (3.1), as follows:

$$\begin{cases} G_0(x_1) = (f_1(x_1) \vee \varepsilon)^\alpha \ , \\ \\ G_k(x_k, x_{k+1}) = \frac{(f_{k+1}(x_{k+1})\vee\varepsilon)^\alpha}{(f_k(x_k)\vee\varepsilon)^\alpha} \ , \quad \text{for all } k = 1, \cdots, n-1 \ , \end{cases} \tag{7.3}$$

where f_k are the payoff functions and $\alpha \in (0, 1]$ and $\varepsilon > 0$. In this choice of potential function G, the parameter α has to be fine-tuned to the particular class of rare events of interest. In our simulations we set $\alpha = 1/5$ and $\varepsilon = 10^{-7}$.

7.3 Numerical results

For each example, we have performed the algorithm for different numbers of mesh points $N = 100, 200, 400, 800, 1600, 3200, 6400$. 1000 runs of both algorithms (*Stochastic Mesh (SM)* and *Stochastic Mesh with Change of Measure (SMCM)*) were performed to compute the mean and confidence intervals of each estimate.

Simulations results are reported in Figure 1, 2 and 3 for the geometric and arithmetic put payoff, with strikes corresponding to standard out of the money puts to deep out of the money puts: $K = 0.95$, $K = 0.85$ and $K = 0.75$. Notice that both algorithms (the Stochastic Mesh algorithm with and without Change of Measure) have been implemented without any standard variance reduction technique (control variate, stratification, ...). In term of complexity, the Stochastic Mesh algorithm with Change of Measure is equivalent to the standard Stochastic Mesh algorithm: the complexity is in both cases quadratic with the number of mesh points $O(N^2)$ since the number of operations required to operate the change of measure is negligible.

We have reported on our graphs to types of estimates:

- the *Positively-biased estimator* provided by the backward induction on the value function;
- the *Negatively-biased estimator* provided by the associated optimal exercise policy. This estimate is obtained via a two-step procedure: first, the optimal policy is approximated in the backward induction on the value function, then the policy is evaluated using the standard forward Monte Carlo procedure. Note that the resulting estimator is known to provide a lower bound (in average) to the option price. In our simulation, we have used $N_{forward} = 10000$ Monte Carlo forward simulations.

As expected, one can observe on Table 2, that the *SMCM* algorithm allows to obtain an estimate, $\hat{v}_{SMCM}$, with the same complexity but with a smaller variance than the standard *SM* algorithm estimate, $\hat{v}_{SM}$, especially for deep out the money options.
More surprisingly, one can observe on Table 2 and Figure 1, 2 and 3 that the *SMCM* algorithm also allows to reduce significantly the estimator bias which is known to compose the

growing part of the error when the number of underlying assets increases. For instance, one can notice that the *SMCM* algorithm achieves the convergence in average of the Positively-biased estimate to the Negatively-biased estimate for a number of mesh points much smaller than for the *SM* algorithm. Hence, the *SMCM* could also be a way to deal with high dimensional optimal stopping problems since the algorithm complexity remains insensitive to the dimension whereas the convergence rate is not significantly reduced.

Payoff	K	$d=1$	$d=2$	$d=3$	$d=4$	$d=5$
Geometric	0.95	1 (1%)	1 (3%)	1 (6%)	1 (9%)	1 (10%)
Put	0.85	5 (2%)	8 (6%)	6 (11%)	4 (14%)	3 (14%)
	0.75	18 (6%)	28 (11%)	18 (17%)	16 (18%)	11 (16%)
Arithmetic	0.95	1 (1%)	3 (2%)	3 (7%)	4 (13%)	5 (18%)
Put	0.85	5 (2%)	13 (6%)	24 (19%)	56 (24%)	100 (20%)
	0.75	18 (6%)	71 (15%)	363 (14%)	866 (16%)	$-$ ($-$)

Table 2: Variance ratio ($\frac{Var(\hat{v}_{SM})}{Var(\hat{v}_{SMCM})}$) and Bias ratio ($\frac{\mathbb{E}(\hat{v}_{SM})-\mathbb{E}(\hat{v}_{SMCM})}{E(\hat{v}_{SM})}$) (within parentheses) computed over 1000 runs for $N = 3200$ mesh points. (For the arithmetic put, when $d = 5$ and $K = 0.75$, the 1000 estimates provided by the standard SM algorithm were all equal to zero, hence the associated variance ratio has not been reported).

References

[1] V. Bally, D. Talay *The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function*. Probab. Th. Related Fields, **104**, 43-60, (1996).

[2] V. Bally, D. Talay *The law of the Euler scheme for stochastic differential equations: II. Approximation of the density*. Monte Carlo Methods and Applications, **2**, 93-128, (1996).

[3] M. Broadie and P. Glasserman,. *A Stochastic Mesh Method for Pricing High-Dimensional American Options*. Journal of Computational Finance, vol. 7, 35-72, (2004).

[4] R. Carmona, J.-P. Fouque and D. Vestal. Interacting Particle Systems for the Computation of Rare Credit Portfolio Losses. *Finance and Stochastics*, vol. 13, no. 4, 2009 pp. 613-633 (2009).

[5] R. Carmona, P. Del Moral, P. Hu and N. Oudjane (eds.), Springer Proceedings in Mathematics. vol. 12, Springer-Verlag, 2012.

[6] R. Carmona, P. Del Moral, P. Hu and N. Oudjane. An Introduction to Particle Methods with Financial Applications, in *Numerical Methods in Finance* of the Springer Proceedings in Mathematics, vol. 12, pp. 3–50 Springer-Verlag, 2012.

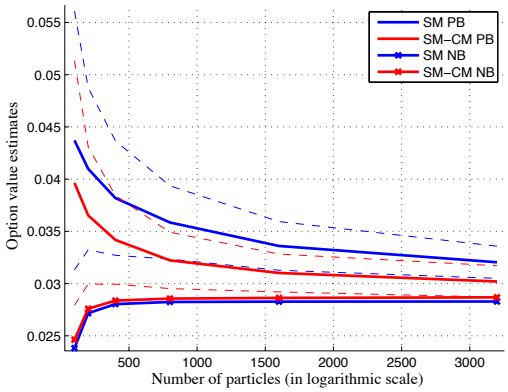
[7] F. Cérou, P. Del Moral, A. Guyader *A non asymptotic variance theorem for unnormalized Feynman-Kac particle models*. HAL-INRIA RR-6716 (2008). To appear in the journal : Annales de l'Institut Henri Poincar (2010)

[8] P. Del Moral, *Feynman-Kac formulae. Genealogical and interacting particle systems with applications*. Probability and its Applications, Springer Verlag, New York (2004).

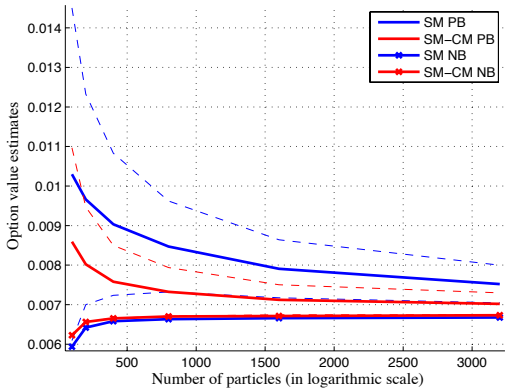
[9] P. Del Moral and J. and Garnier, *Genealogical particle analysis of rare events*. Annals of Applied Probability, vol. 15, 2496–2534, (2005).

[10] P. Del Moral, P. Hu, N. Oudjane, B. Rémillard, *On the Robustness of the Snell Envelope*, SIAM J. Finan. Math., Vol. 2, pp. 587-626, 2011.

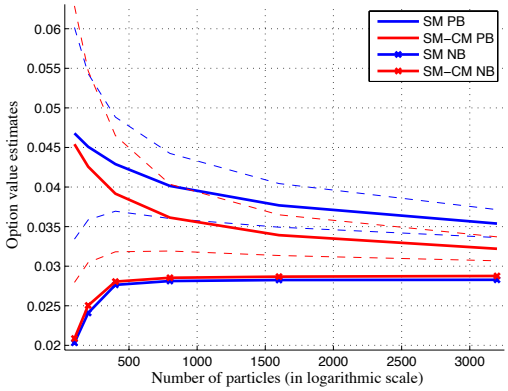
[11] G. Liu and L. J. Hong. *Revisit of stochastic mesh method for pricing American options*. Operations Research Letters, 37(6), 411-414 (2009).



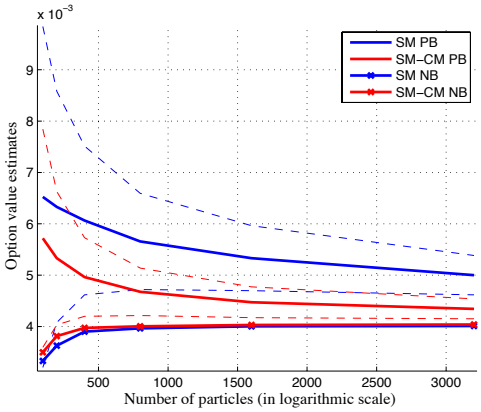
(a) Geometric Put with $d = 3$ assets



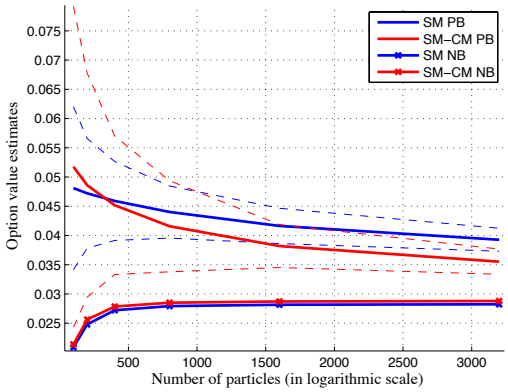
(b) Arithmetic Put with $d = 3$ assets



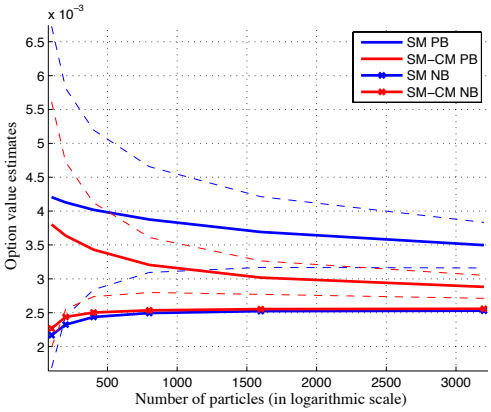
(c) Geometric Put with $d = 4$ assets



(d) Arithmetic Put with $d = 4$ assets

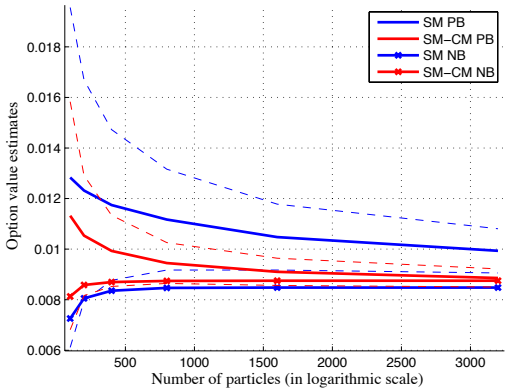


(e) Geometric Put with $d = 5$ assets

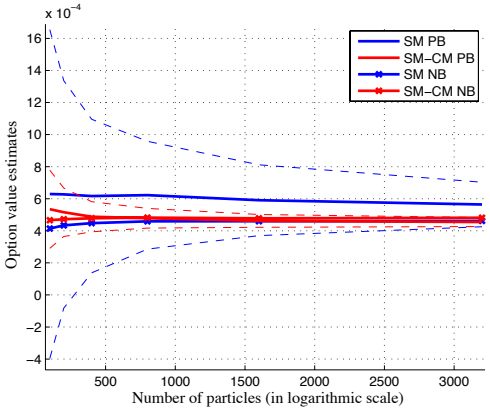


(f) Arithmetic Put with $d = 5$ assets

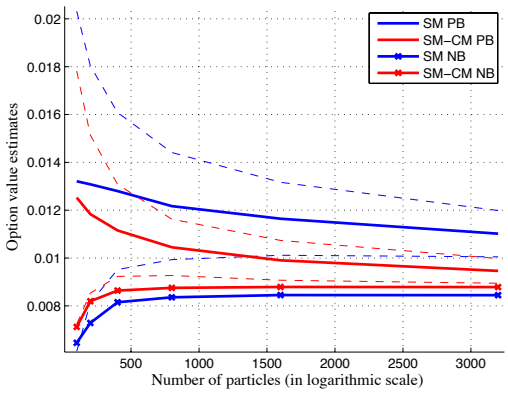
Figure 1: Positively-biased option values estimates (average estimates with 95% confidence interval computed over 1000 runs) and Negatively-biased option values estimates (average estimates over the 1000 runs each forward estimate being evaluated over 10000 forward Monte Carlo simulations), computed by the *SM* algorithm (in blue line) and the *SMCM* algorithm (in red line), as a function of the number of mesh points for geometric (on the left column) and arithmetic (on the right column) put options with **strike** $K = 0.95$.



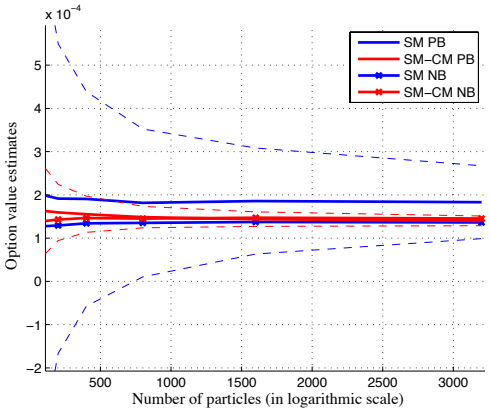
(a) Geometric Put with $d = 3$ assets



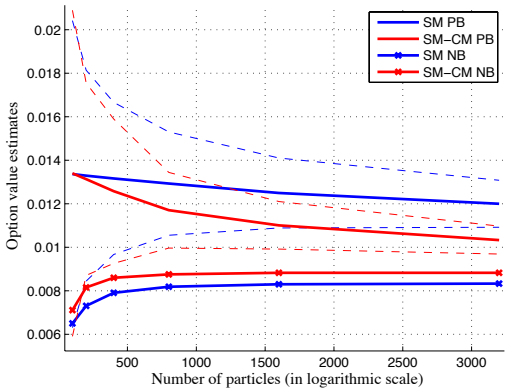
(b) Arithmetic Put with $d = 3$ assets



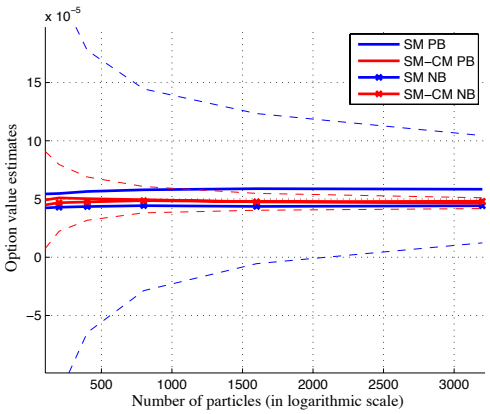
(c) Geometric Put with $d = 4$ assets



(d) Arithmetic Put with $d = 4$ assets

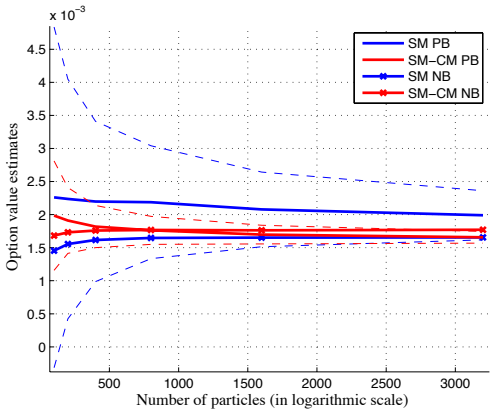


(e) Geometric Put with $d = 5$ assets

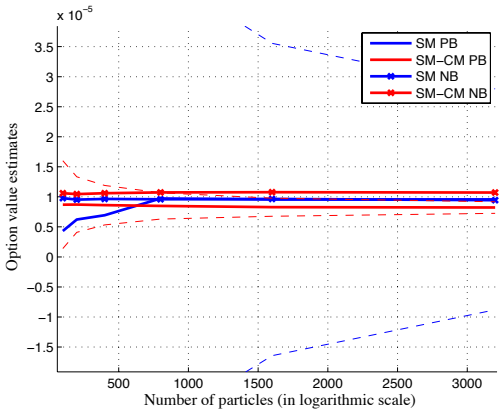


(f) Arithmetic Put with $d = 5$ assets

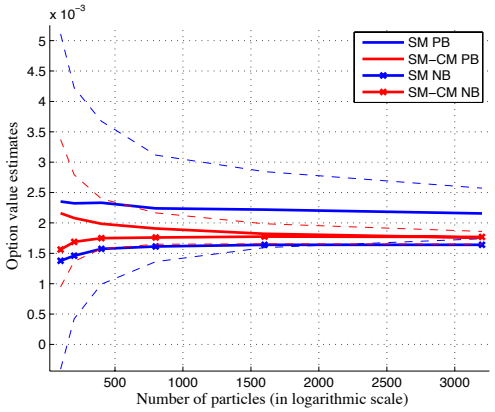
Figure 2: Positively-biased option values estimates (average estimates with 95% confidence interval computed over 1000 runs) and Negatively-biased option values estimates (average estimates over the 1000 runs each forward estimate being evaluated over 10000 forward Monte Carlo simulations), computed by the *SM* algorithm (in blue line) and the *SMCM* algorithm (in red line), as a function of the number of mesh points for geometric (on the left column) and arithmetic (on the right column) put options with **strike** $K = 0.85$.



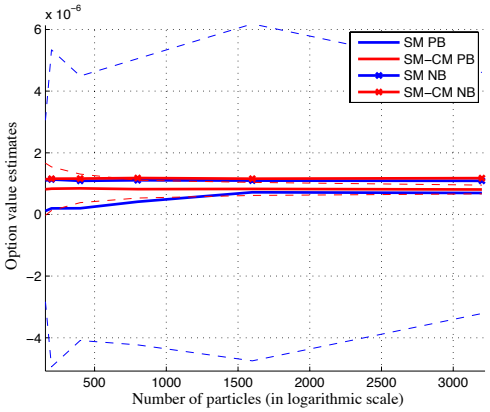
(a) Geometric Put with $d = 3$ assets



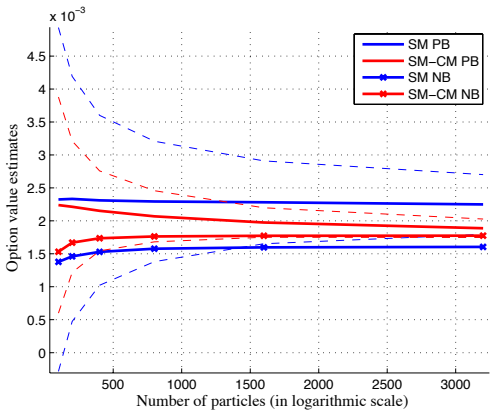
(b) Arithmetic Put with $d = 3$ assets



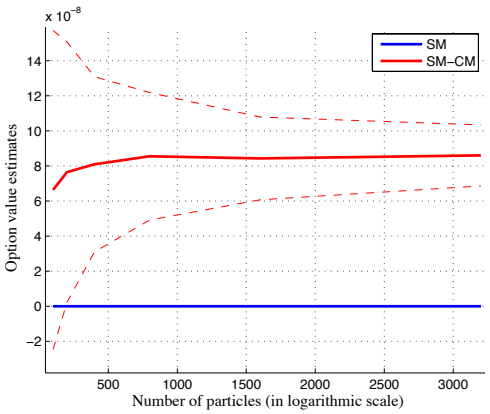
(c) Geometric Put with $d = 4$ assets



(d) Arithmetic Put with $d = 4$ assets



(e) Geometric Put with $d = 5$ assets



(f) Arithmetic Put with $d = 5$ assets

Figure 3: Positively-biased option values estimates (average estimates with 95% confidence interval computed over 1000 runs) and Negatively-biased option values estimates (average estimates over the 1000 runs each forward estimate being evaluated over 10000 forward Monte Carlo simulations), computed by the *SM* algorithm (in blue line) and the *SMCM* algorithm (in red line), as a function of the number of mesh points for geometric (on the left column) and arithmetic (on the right column) put options with **strike** $K = 0.75$. (For the clarity of the graph (f), the Negatively-biased estimate is not reported, the associated variance (for 10 000 forward Monte Carlo simulations) being relatively strong).