Snell envelope with path dependent multiplicative optimality criteria

Pierre Del Moral, Peng Hu, Nadia Oudjane,
Snell envelope with path dependent multiplicative optimality criteria

Pierre Del Moral\textsuperscript{\textregistered}, Peng Hu\textsuperscript{†}, Nadia Oudjane\textsuperscript{‡}

Theme : Stochastic Methods and Models
Applied Mathematics, Computation and Simulation
Équipe-Projet ALEA

Rapport de recherche n° 7360 — August 2010 — 15 pages

Abstract: We analyze the Snell envelope with path dependent multiplicative optimality criteria. Especially for this case, we propose a variation of the Snell envelope backward recursion which allows to extend some classical approximation schemes to the multiplicatively path dependent case. In this framework, we propose an importance sampling particle approximation scheme based on a specific change of measure, designed to concentrate the computational effort in regions pointed out by the criteria. This new algorithm is theoretically studied. We provide non asymptotic convergence estimates and prove that the resulting estimator is high biased.

Key-words: Snell envelope, american option, particle model, rare events

\textsuperscript{\textregistered} 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
\textsuperscript{†} 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
\textsuperscript{‡} EDF R & D Clamart (nadia.oudjane@edf.fr)
Enveloppe de Snell avec des critères d’optimalité multiplicativement dépendants de chemin

Résumé : Nous analysons l’enveloppe de Snell avec des critères d’optimalité multiplicativement dépendants de chemin. Surtout pour ce cas, nous proposons une variation de backward récurrence de l’enveloppe de Snell qui permet d’étendre certains schémas d’approximation classique à ce cas spécial. Dans ce cadre, nous proposons un schéma d’approximation particule d’échantillonnage importance basé sur un changement de mesure spécifique, destiné à concentrer l’effort de calcul dans les régions soulignées par les critères. Ce nouvel algorithme est théoriquement étudié. Nous fournissons des estimations non convergence asymptotique et de prouver que l’estimateur resultant est surestimé.

Mots-clés : enveloppe de Snell, option américain, modèle particule, événements rares
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. In this paper, we are interested in some complicated optimality criteria, especially the multiplicatively path dependent case. In other words, given a random process \((X_k)_{0 \leq k \leq n}\) and some gain functions \((f_k)_{0 \leq k \leq n}\) and \((G_k)_{0 \leq k \leq n}\), we want to maximize the expected gain \(\mathbb{E}(f_t(X_t) \prod_{k=1}^{n-1} G_k(X_{k}))\) by choosing \(\tau\) on a set of random stopping times \(\mathcal{T}\). For example, in finance, the multiplicative optimality criteria \((G_k)_{0 \leq k \leq n}\) can be interpreted as a discount factor related to a stochastic interest rate (taking then an exponential form), or as an obstacle for exotic options such as barriers in knock out options (taking then the form of indicator functions).

In the discrete time setting, these problems associated with 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 \(\tau\) 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

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

Then the computation of the Snell envelope \((U_k)_{0 \leq k \leq n}\) amounts to solve the following backward functional equation\(^1\)

\[
u_k = f_k \vee M_{k+1}(u_{k+1}) \tag{1.1}
\]

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 amounts to replace in recursion (1.1) the pair of functions and Markov transitions \((f_k, M_k)_{0 \leq k \leq n}\) by some approximation model \((\tilde{f}_k, \tilde{M}_k)_{0 \leq k \leq n}\) on some possibly reduced measurable subsets \(\tilde{E}_k \subset E_k\). In paper \(^2\), the authors provided a general robustness lemma to estimate the error related to the resulting approximation \(\hat{u}_k\) of the Snell envelope \(u_k\), for several types of approximation models \((\tilde{f}_k, \tilde{M}_k)_{0 \leq k \leq n}\).

\(^1\) In present paper, if not specified, when one talks about the potential rare event \(G_k\) or optimality criteria \(G_k\), it means \(G_k(X_k)\)

\(^2\) Consult the last paragraph of this section for a statement of the notation used in this article.
Lemma 1.1 For any $0 \leq k < n$, on the state space $\widehat{E}_k$, we have that
\[ |u_k - \hat{u}_k| \leq \sum_{l=k}^{n} M_{k,l} |f_l - \hat{f}_l| + \sum_{l=k}^{n-1} M_{k,l} (M_{l+1} - \hat{M}_{l+1}) u_{l+1} |. \]

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 of a multiplicative optimality criteria.

Let us come back now to the multiplicatively path dependent case that we mentioned in the beginning of the article. Instead of $\mathbb{E}(f_\tau(X_\tau))$ we want to maximize $\mathbb{E}(f_\tau(X_\tau) \prod_{p=0}^{n-1} G_p(X_p))$ on the stopping times set $\mathcal{T}$. In this situation, a natural way is to consider the path $(X_0 \ldots X_k)_{0 \leq k \leq n}$ as a new Markov chain $(X_k)_{0 \leq k \leq n}$ on path spaces and associate with transitions given for any $\chi_{k-1} = (x_0, \ldots, x_{k-1}) \in (E_0 \times \cdots \times E_{k-1})$ and $\chi_k = (x_0', \ldots, x_k') \in (E_0 \times \cdots \times E_k)$ by the following formula
\[ \mathcal{M}_k(\chi_{k-1}, d\chi_k') = \delta_{\chi_{k-1},(d\chi_{k-1})} \mathcal{M}_k(x_{k-1}', dx_k') . \]

Then, let us denote by $u_k(x_0 \ldots x_k)$, the Snell envelope defined with a path version of recursion (1.1):
\[ u_k(x_0, \ldots, x_k) = [f_k(x_k) \prod_{p=0}^{k-1} G_p(x_p)] \vee \mathcal{M}_{k+1}(u_{k+1})(x_0, \ldots, x_k) , \quad (1.2) \]

for $0 \leq k < n$ with terminal value $u_n(x_0, \ldots, x_n) = f_n(x_n) \prod_{p=0}^{n-1} G_p(x_p)$. At this stage, two difficulties may arise. First, the above recursion (1.2) seem to require the approximation of a $k + 1$ dimensional function at each time step from $k = n - 1$ up to $k = 0$. Second, when the optimality criteria $G_p$ is localized in a specific region of $E_p$, for each $p$, then the product $\prod_{p=0}^{k-1} G_p(x_p)$ can be interpreted as a rare event. Hence, at first glance, the computation of Snell envelopes in the multiplicatively path dependent case seem to combine two additional numerical difficulties w.r.t. to the standard case, related to the computation of conditional expectations in a both high dimensional and rare event situation.

These issues are considered in Section 2, of the present paper. The dimensionality problem is easily bypassed by considering an intermediate standard Snell envelope $(v_k)_{0 \leq k \leq n}$, without path dependent criteria, which is directly related to the multiplicatively path dependent Snell envelope, by the relation $u_k(x_0, x_1 \ldots x_k) = \prod_{p=0}^{k-1} G_p(x_p) v_k(x_k)$, for all $0 \leq k \leq n$. Hence, computing the original Snell envelope $u_k$ can be done by using one of the many approximation schemes developed for the standard (non path dependent) case. Then, 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 $(G_k)_{0 \leq k \leq n-1}$.

In Section 3, we propose a Monte Carlo algorithm to compute the multiplicatively path dependent Snell envelope, on the base of this intermediate standard Snell envelope under a new equivalent measure defined in the previous section. This new approximation scheme is based on the stochastic mesh method introduced by M. Broadie and P. Glasserman in their seminal paper [3] (see also [7], for some recent refinements). The principal idea of original Broadie-Glasserman
model is to make a change of probability, under the assumption that the Markov transitions \( M_k(x, \cdot) \) are absolutely continuous w.r.t. some other measure \( \eta_k \) on \( E_k \), with positive Radon Nikodym derivatives \( \bar{R}_k(x, y) = \frac{dM_k(x, y)}{d\eta_k(y)} \). But in most cases, we do not know the density function of some good choice of \( \eta_k \).

So in [6], the authors provide a variation of Broadie-Glasserman model that replaces not only \( \eta_k \) but also the Radon-Nikodym derivatives \( \bar{R}_k \) with the approximation model \((\bar{\eta}_k, \bar{R}_k)\). The model introduced in the present article is an extension to multiplicatively path dependent functions.

In Section 4, the proposed Monte Carlo algorithm is theoretically analysed using an interacting particle system interpretation. We provide non asymptotic convergence estimates and prove that the resulting estimator is high biased.

For the convenience of the reader, we end this introduction with some notation used in the present article. 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 multidimensional 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 \( M_0(x_{n-1}, dx_n) \) 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 \( \sigma \)-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 Snell envelope with multiplicatively path dependent functions and change of measure

Suppose \((X_k)_{0 \leq k \leq n}\) is a Markov chain on continuous state spaces \((E_k, \mathcal{E}_k)_{0 \leq k \leq n}\) with an initial distribution \(\nu_0\) on \(E_0\), a collection of Markov transitions \(M_k(x_{k-1}, dx_k)\) from \(E_{k-1}\) to \(E_k\) and a given final time horizon \(n\). We also assume that the chain \(X_k\) is defined on a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\). In this situation, the historical process \(X_k := (X_0, \ldots, X_k)\) can be seen as a Markov chain with transitions given for any \(\chi_{k-1} = (x_0, \ldots, x_{k-1}) \in E_0 \times \cdots \times E_{k-1}\) and \(\chi'_k = (x'_0, \ldots, x'_k) \in E_0 \times \cdots \times E_k\) by the following formula

\[
\mathcal{M}_k(\chi_{k-1}, d\chi'_k) = \delta_{\chi_{k-1}}(d\chi'_k) M_k(x_{k-1}, dx_k).
\]

We denote by \((P_k)_{0 \leq k \leq n}\) a sequence of probabilities of path \((X_k)_{0 \leq k \leq n}\). For a given collection of real valued functions \((f_k)_{0 \leq k \leq n}\) and \((G_k)_{0 \leq k \leq n}\) 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 product spaces \(E_0 \times \cdots \times E_k\) by

\[
F_k(x_0, \ldots, x_k) := f_k(x_k) \prod_{0 \leq p \leq k-1} G_p(x_p), \quad \text{for all } 0 \leq k \leq n.
\]

To maximize the expected gain \(\mathbb{E}(F_{\tau}(X_{\tau}))\) w.r.t. \(\tau\) in a set of random stopping times \(\mathcal{I}\), one is interested in computing the Snell envelope \((u_k)_{0 \leq k \leq n}\) associated to the gain functions \((f_k)_{0 \leq k \leq n}\) and solution to the following recursion

\[
\begin{cases}
    u_0(x_0, \ldots, x_n) = f_n(x_n), \\
    u_k(x_0, \ldots, x_k) = F_k(x_0, \ldots, x_k) \vee \mathcal{M}_{k+1}(u_{k+1})(x_0, \ldots, x_k), \forall 0 \leq k \leq n - 1.
\end{cases}
\]

(2.1)

Now, let us consider the standard (non path dependent) Snell envelope \((v_k)_{0 \leq k \leq n}\) associated to the gain functions \((f_k)_{0 \leq k \leq n}\) and satisfying the following recursion

\[
\begin{cases}
    v_n(x_n) = f_n(x_n), \\
    v_k(x_k) = f_k(x_k) \vee [G_k(x_k)M_{k+1}(v_{k+1})(x_k)], \text{ for all } 0 \leq k \leq n - 1.
\end{cases}
\]

(2.2)

For all \(0 \leq k \leq n\), let us denote by \(v_k\) the real valued functions defined on \(E_0 \times \cdots \times E_k\), such that \(v_k(x_0, \ldots, x_k) := v_k(x_k) \prod_{p=0}^{k-1} G_p(x_p)\). By construction, one can easily check that for all \(0 \leq k \leq n\), \(u_k \equiv v_k\) and in particular \(u_0(x_0) = v_0(x_0)\). Indeed, one can verify that \((v_k)_{0 \leq k \leq n}\) follows the same recursion \((2.2)\) as \((u_k)_{0 \leq k \leq n}\). First, we note that they share the same terminal condition,

\[
v_n(x_0, \ldots, x_n) = v_n(x_n) \prod_{p=0}^{n-1} G_p(x_p) = f_n(x_k) \prod_{p=0}^{n-1} G_p(x_p) = F_n(x_0, \ldots, x_n) = u_n(x_0, \ldots, x_n).
\]

RR n° 7360
Let us now introduce the integral operator $Q$ with distribution \( \eta \). We denote the recursive relation between \( G \) of the criteria \( M \) to compute the latter. The recursion (2.2) implies that it is not relevant to interpret comes by setting \( \eta \). Furthermore, it is also important to observe that, for any measurable function \( f \), when approximating the conditional expectation \( M_k+1(v_{k+1})/v_k \) by a Monte Carlo method, it seems relevant to concentrate the simulations in the regions of $E_k$ where \( G_{k+1} \) reaches high values. Hence, to avoid the potential rare events \( G \), we propose to consider the following change of measure one the measurable product space \( (E_0 \times \cdots \times E_n, \mathcal{E}_0 \times \cdots \times \mathcal{E}_n) \),

$$dQ_n = \frac{1}{Z_n} \left[ \prod_{k=0}^{n-1} G_k \right] dP_n , \quad \text{with} \quad Z_n = \mathbb{E} \left( \prod_{k=0}^{n-1} G_k X_k \right) = \prod_{k=0}^{n-1} \eta_k (G_k) ,$$

where \( \eta_k \) is the probability measure defined on \( E_k \) such that, for any measurable function \( f \) on \( E_k \)

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

The measures \( (\eta_k)_{0 \leq k \leq n} \) defined above can be seen as the laws of \( (X_k)_{0 \leq k \leq n} \) under probability $Q$. Loosely speaking, the process \( (X_k)_{0 \leq k \leq n} \) with distribution \( (\eta_k)_{0 \leq k \leq n} \) is designed under the constrain \( (\prod_{p=0}^{k} G_p)_{0 \leq k \leq n} \). An intuitive interpretation comes by setting \( G_k(x_k) = 1_{A_k}(x_k) \) with \( A_k \subset E_k \), then the process with distribution \( \eta_k \) is just the ones surviving in the subsets \( A_k \). It follows that the measures \( \eta_k \) seem to be a relevant choice for the change of probability in our path dependent situation.

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

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

We denote the recursive relation between \( \eta_k \) and \( \eta_{k-1} \) by introducing the operators \( \Phi_k \) such that, for all \( 1 \leq k \leq n \)

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

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

$$Q_k(f)(x_{k-1}) := \int G_{k-1}(x_{k-1}) M_k(x_{k-1}, dx_k) f(x_k) .$$
In further developments of this article, we suppose that \( M_k(x_{k-1}, \cdot) \) are equivalent to some measures \( \lambda_k \), 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 \( \lambda_k \) such that:

\[
M_k(x_{k-1}, dx_k) = H_k(x_{k-1}, x_k) \lambda_k(dx_k) .
\] (2.7)

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

**Lemma 2.1** For any measure \( \eta \) on \( E_k \), recursion 2.7 defining \( v_k \) can be rewritten:

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

for any \( x_k \in E_k \), where

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

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

**Proof:**

Under Assumption 2.7, we have immediately the following formula

\[
M_{k+1}(x_k, dx_{k+1}) = H_{k+1}(x_k, x_{k+1}) \frac{\eta_k(G_k)}{\eta_k(G_k H_{k+1}(\cdot, x_{k+1}))} \eta_{k+1}(dx_{k+1}) .
\] (2.8)

Now, note that the above equation is still valid for any measure \( \eta \),

\[
M_{k+1}(x_k, dx_{k+1}) = H_{k+1}(x_k, x_{k+1}) \frac{\eta(G_k)}{\eta(G_k H_{k+1}(\cdot, x_{k+1}))} \Phi_{k+1}(\eta)(dx_{k+1}) .
\] (2.9)

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

\[
\frac{dM_{k+1}(x_k, \cdot)}{d\Phi_{k+1}(\eta)}(x_{k+1}) = H_{k+1}(x_k, x_{k+1}) \frac{\eta(G_k)}{\eta(G_k H_{k+1}(\cdot, x_{k+1}))} .
\] (2.10)

We end the proof by applying above arguments to recursion 2.2.

### 3 A particle approximation scheme

From the above discussion, we conclude that the distributions \( (\eta_k)_{0 \leq k \leq n} \) are a very good choice for the change of probability for the stochastic mesh model. In this section, we first propose a particle model to sample the random variables according to these distributions, then we describe the resulting particle scheme proposed to approximate the Snell envelope \( (v_k)_{0 \leq k \leq n} \).

By definition 2.7 of \( \Phi_{k+1} \), we have the following formula

\[
\Phi_k(\eta_{k-1}) = \eta_{k-1} K_{k, \eta_{k-1}} = \eta_{k-1} S_{k-1, \eta_{k-1}}, M_k = \Psi_{G_k}(\eta_{k-1}) M_k .
\] (3.1)
More generally, the operations \(M\) and \(\eta\) are defined as follows:

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

where the real \(\epsilon\) is such that \(\epsilon G\) takes its values \([0, 1]\).

More generally, the operations \(\Psi\) and \(S\) can be expressed as \(\Psi_G(\eta)(f) = \frac{\eta_G(f)}{\eta_G} = \eta S_G(f)\) with \(S_G(f) = \epsilon G f + (1 - \epsilon G) \Psi_G(\eta)(f)\).

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, \(\xi_0^{(N)} = (\xi_0^{(i, N)})_{1 \leq i \leq N}\), 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 \(\xi_k\) and \(\xi_k^i\) instead of \(\xi_k^{(N)}\) and \(\xi_k^{(i, N)}\). By construction, \(\xi_k\) is a particle model with a selection transition and a mutation type exploration i.e. the evolution from \(\xi_k\) to \(\xi_{k+1}\) is composed by two steps:

\[
\xi_k \in E_k^N \xrightarrow{\text{Selection}} \xi_k := (\xi_k^i)_{1 \leq i \leq N} \in E_k^N \xrightarrow{\text{Mutation}} \xi_{k+1} \in E_{k+1}^N. \tag{3.2}
\]

Then we define \(\eta_k^N\) and \(\tilde{\eta}_k^N\) as the occupation measures after the mutation and the selection steps. More precisely,

\[
\begin{align*}
\eta_k^N &= \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_k^i} \quad \text{and} \quad \tilde{\eta}_k^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_k^i}.
\end{align*}
\]

During the selection transition \(S_{k,\eta_{k-1}}\), 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 leave \(\xi_k^i\) stay on particle \(\xi_k^i\), and with probability \(1 - \epsilon G_k(\xi_k^i)\) we decide to do the following selection: \(\xi_k^j\) randomly takes the value in \(\xi_k^i\) for \(0 \leq j \leq N\) with distribution \(\frac{G_k(\xi_k^i)}{\sum_{i=1}^{N_i} G_k(\xi_k^i)}\). Note that when \(\epsilon G_k \equiv 1\), the selection is skipped (i.e. \(\tilde{\xi}_k = \xi_k\)) so that the model corresponds exactly to the Broadie-Glasserman type model analysed by P. Del Moral and P. Hu et al. \[6\]. Hence, the factor \(\epsilon\) can be interpreted as a level of selection against the rare events.

During the mutation transition \(\xi_k \sim \xi_{k+1}\), every selected individual \(\tilde{\xi}_k^i\) evolves randomly to a new individual \(\xi_{k+1}^i = x\) randomly chosen with the distribution \(M_{k+1}(\xi_k^i, dx)\), for \(1 \leq i \leq N\).

It is important to observe that by construction, \(\eta_{k+1}^N\) is the empirical measure associated with \(N\) conditionally independent and identically distributed random individual \(\xi_{k+1}^i\) with common distribution \(\Phi_{k+1}(\eta_k^N)\).
By the previous construction, we can approximate the Snell envelope \((v_k)_{0 \leq k \leq n}\). The main idea consists in taking \(\eta = \eta_N\), in Lemma 4.1 then observing that Snell envelope \((v_k)_{0 \leq k \leq n}\) is solution of the following recursion, for all \(0 \leq k < n\),

\[
v_k(x_k) = f_k(x_k) \lor \Phi_{k+1}(\eta_k^N) \left( \frac{dQ_{k+1}(x_{k+1})}{d\Phi_{k+1}(\eta_k^N)} \right) v_{k+1}.
\]

Now, if \(\Phi_{k+1}(\eta_k^N)\) is well estimated by \(\eta_{k+1}^N\), it is relevant to approximate \(v_k\) by \(\hat{v}_k\) defined by the following backward recursion

\[
\begin{cases}
\hat{v}_n &= f_n \\
\hat{v}_k(x_k) &= f_k(x_k) \lor \eta_{k+1}^N \left( \frac{dQ_{k+1}(x_{k+1})}{d\Phi_{k+1}(\eta_k^N)} \right) \hat{v}_{k+1} & \text{for all } 0 \leq k < n,
\end{cases}
\]

Note that in the above formula (3.3), 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

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

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 \(0 \leq k < n\),

\[
v_k = f_k \lor \hat{Q}_{k+1}(v_{k+1}) \quad \text{and} \quad \hat{v}_k = f_k \lor \hat{Q}_{k+1}(\hat{v}_{k+1}).
\]

4 Convergence and bias analysis

By the previous construction, we can approximate \(\Phi_{k+1}(\eta_k^N)\) by \(\eta_{k+1}^N\). 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 unbiasedness property and mean error estimates for the approximation \(\eta_{k+1}^N\) of \(\Phi_{k+1}(\eta_k^N)\).

**Lemma 4.1** 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

\[
\mathbb{E} \left( \eta_{k+1}^N(f) \left| F_k^N \right. \right) = \Phi_{k+1}(\eta_k^N)(f)
\]

and

\[
\mathbb{E} \left( \left| V_k^N(f) \right|^p \left| F_k^N \right. \right) \leq 2 a(p) \left( \Phi_{k+1}(\eta_k^N)(|f|^{p'}) \right)^{1/2}
\]

with the collection of constants

\[
a(2p)^2 = (2p)p \quad \text{and} \quad a(2p + 1)^{2p+1} = \frac{(2p + 1)^{p+1}}{\sqrt{p + 1/2}} 2^{-(p+1/2)}.
\]
We are now in position to state the main result of this paper.

A consequence of the unbiasedness property proved in Lemma 1.1 is that

\[ \sqrt{N} \mathbb{E} \left( \left| \frac{N}{|k|} - \mu_{k+1}^N \right| \right)^p \leq 2 a(p) \mathbb{E} \left( \left| f(\xi_{k+1}) \right|^p \right)^{\frac{1}{p}}. \]

We end the proof by combining the above two inequalities.

Now, by Lemma 1.1 we conclude

\[ \sqrt{N} \left| (v_k - \tilde{v}_k) \right| \leq \sum_{k<l<n} \mathbb{E} \left( (R_{k+1}^N)(v_{l+1}) \right). \quad (4.1) \]

We are now in position to state the main result of this paper.
Theorem 4.2 For any $0 \leq k \leq n$ and any integer $p \geq 1$, we have
\[
\sup_{x \in E_k} \|\tilde{v}_k - v_k\|_{L_p} \leq \sum_{k<l<n} \frac{2 a(p)}{\sqrt{N}} q_{k,l} \left[ Q_{k+l}^h(b_{k+l+1}^{p-1}v_{k+1}^{p})(x) \right]^{\frac{1}{p}}
\]
with a collection of constants $q_{k,l}$ and functions $h_k$ defined as
\[
q_{k,l} := \left( G_l \parallel h_{k+1} \prod_{m=k}^{l-1} \parallel G_m \parallel \right)^{\frac{p-1}{p}} \quad \text{and} \quad h_k(x) := \sup_{x,y \in E_{k-1}} \frac{H_k(x,x)}{H_k(y,y)}.
\]

Proof: First, decomposition (4.1) yields
\[
\sqrt{N} \|\tilde{v}_k - v_k\|_{L_p} \leq \sum_{k<l<n} \left\| \tilde{Q}_{k,l}((R_{l+1}^N)(v_{l+1}))(x) \right\|_{L_p}, \quad \text{for all} \ x \in E_k.
\]
Note that
\[
\|\tilde{Q}_{k,l}(1)\| \leq b_{k,l}, \quad \text{where} \ b_{k,l} := \|h_{k+1}\| \prod_{m=k}^{l-1} \|G_m\|.
\]
Then it follows easily that for any integrable function $f$ on $E_l$
\[
(Q_{k,l}(f))^p \leq (b_{k,l})^{p-1} \tilde{Q}_{k,l}(f^p).
\]
This yields that
\[
\left\| \tilde{Q}_{k,l}((R_{l+1}^N)(v_{l+1}))(x) \right\|_{L_p} \leq (b_{k,l})^{\frac{p-1}{p}} \mathbb{E}\left( \tilde{Q}_{k,l}((R_{l+1}^N)(v_{l+1}))(x) \right)^{\frac{1}{p}}.
\]
Applying Lemma 4.1 to the right-hand side of the above inequality, we obtain for any $x_l \in E_l$
\[
\mathbb{E}\left( ((R_{l+1}^N)(v_{l+1}))(x_l) \right)^{\frac{p}{p}} \leq 2 a(p) \left[ \int_{E_{l+1}} \Phi_{l+1}(\eta^N_l)(dL_{l+1}) \left( \frac{dQ_{l+1}(x_l, \cdot)}{d\Phi_{l+1}(\eta^N_l)}(x_{l+1})v_{l+1}(x_{l+1}) \right)^{p-1} \right]^{\frac{1}{p}}
\]
from which we find that
\[
\mathbb{E}\left( ((R_{l+1}^N)(v_{l+1}))(x_l) \right)^{\frac{p}{p}} \leq 2 a(p) \left[ \int_{E_{l+1}} Q_{l+1}(x_l, dL_{l+1}) \left( \frac{dQ_{l+1}(x_l, \cdot)}{d\Phi_{l+1}(\eta^N_l)}(x_{l+1}) \right)^{p-1} v_{l+1}(x_{l+1}) \right]^{\frac{1}{p}}
\]
By definition 1.2 of functions $h_{l+1}$ and in developing the Radon Nikodym derivative, we obtain
\[
\frac{dQ_{l+1}(x_l, \cdot)}{d\Phi_{l+1}(\eta^N_l)}(x_{l+1}) = \frac{\eta^N_l(G_l)G_l(x_l)H_{l+1}(x_l, x_{l+1})}{\eta^N_l(G_lH_{l+1})} \leq \parallel G_l \parallel h_{l+1}(x_{l+1}),
\]
RR n° 7360.
which implies
\[
\mathbb{E} \left( \left( |R_{t+1}^N| (v_{t+1}(x_t)) \right)^p \eta_k^N \right)^{\frac{1}{p}} 
\leq 2 a(p) \|G_t\|^{\frac{2-p}{p}} \left[ \int_{E_{t+1}} Q_{t+1}(x_t, dx_{t+1}) \left( h_{t+1}(x_{t+1}) \right)^{p-1} v_{t+1}(x_{t+1}) \right]^{\frac{1}{p}}
\]

Gathering the above arguments, we conclude that
\[
\| (\hat{v}_k - v_k) (x) \|_{L^2} \leq \sum_{k < l < n} \frac{2 a(p)}{\sqrt{N}} q_{k,l} \left( Q_{k,l+1}(h_{t+1}^{p'-1} v_{t+1}')(x) \right)^{\frac{1}{p'}}.
\]

**Remarks**: The constants $q_{k,l}$ could be largely reduced. In fact, $q_{k,l}$ comes from bounding $\| \prod_m \eta_k^N (G_m) \|_{L^2}$. In [4], the authors proved $\| \prod_m G_m \|_{L^2} + \text{constant}$ as a non asymptotic boundary for $\| \prod_m \eta_k^N (G_m) \|_{L^2}$. In most cases, the functions $G$ take their values in $[0, 1]$, then the majoration $\| \prod_m G_m \| \leq 1$ 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 $\hat{v}_k$ by the particle estimates $\hat{v}_k 1_{r \geq n}$. 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 [4].

To understand better the $\mathbb{L}_p$-mean error bounds in the above theorem, we deduce the following exponential concentration inequality

**Proposition 4.3** For any $0 \leq k \leq n$ any and any $\epsilon > 0$, we have
\[
\sup_{x \in E_n} \mathbb{P} \left[ |v_k(x) - \tilde{v}_k(x)| > \frac{c}{\sqrt{N}} + \epsilon \right] \leq \exp \left( -N \epsilon^2 / \epsilon^2 \right),
\]

with constant $c = \sum_{k < l < n} \frac{2 a(p)}{\sqrt{N}} q_{k,l} \left( Q_{k,l+1}(h_{t+1}^{p'-1} v_{t+1}')(x) \right)^{\frac{1}{p'}}$.

**Proof**: This result is a direct consequence from the fact that for any non negative random variable $U$ such that
\[
\exists b < \infty \text{ s.t. } \forall r \geq 1 \quad \mathbb{E} (U^r)^{\frac{1}{r}} \leq a(r) b \quad \Rightarrow \mathbb{P} (U \geq b + \epsilon) \leq \exp \left( -\epsilon^2 / (2b^2) \right).
\]

To check this claim, we develop the exponential and verify that
\[
\forall t \geq 0 \quad \mathbb{E} (e^{tU}) \leq \exp \left( \frac{(bt)^2}{2} + bt \right) \Rightarrow \mathbb{P} (U \geq b + \epsilon) \leq \exp \left( -\sup_{t \geq 0} \left( ct - \frac{(bt)^2}{2} \right) \right)
\]

Simlarly to the orginal Broadie-Glasserman model, the following proposition shows that in this model we also over-estimate the Snell envelope.
Proposition 4.4 For any $0 \leq k \leq n$ and any $x_k \in E_k$

$$E(\hat{v}_k(x_k)) \geq v_k(x_k). \quad (4.4)$$

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

$$E(\hat{v}_k(x_k)) \geq f_k(x_k) \lor E\left(\hat{Q}_{k+1}\hat{v}_{k+1}(x_k)\right)$$

$$= f_k(x_k) \lor E\left(\int_{E_{k+1}^N} \hat{Q}_{k+1}(x_k, dx_{k+1})E(\hat{v}_{k+1}(x_{k+1})|F_{k+1}^N)\right).$$

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

$$E\left(\int_{E_{k+1}^N} \hat{Q}_{k+1}(x_k, dx_{k+1})E(\hat{v}_{k+1}(x_{k+1})|F_{k+1}^N)\right) \geq E(\hat{Q}_{k+1}v_{k+1}(x_k))$$

$$= Q_{k+1}v_{k+1}(x_k).$$

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

\[\blacksquare\]
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] F. Croux, 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)

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

[6] P. Del Moral, P. Hu, N. Oudjane, B. Rémillard, On the Robustness of the Snell Envelope, preprint INRIA-00487103, 2010

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