Feynman-Kac representation of fully nonlinear PDEs and applications

Huyen Pham

Feynman-Kac representation of fully nonlinear PDEs and applications

Huyên PHAM
Laboratoire de Probabilités et Modèles Aléatoires
CNRS, UMR 7599
Université Paris Diderot, Paris, France
pham at math.univ-paris-diderot.fr
and John Von Neumann Institute, Ho-Chi-Minh City

September 2, 2014

Abstract

The classical Feynman-Kac formula states the connection between linear parabolic partial differential equations (PDEs), like the heat equation, and expectation of stochastic processes driven by Brownian motion. It gives then a method for solving linear PDEs by Monte Carlo simulations of random processes. The extension to (fully)nonlinear PDEs led in the recent years to important developments in stochastic analysis and the emergence of the theory of backward stochastic differential equations (BSDEs), which can be viewed as nonlinear Feynman-Kac formulas. We review in this paper the main ideas and results in this area, and present implications of these probabilistic representations for the numerical resolution of nonlinear PDEs, together with some applications to stochastic control problems and model uncertainty in finance.

MSC Classification (2000): 60H30, 65C99, 93E20

Keywords: Nonlinear PDE, Hamilton-Jacobi-Bellman equation, backward stochastic differential equation, randomization of controls, discrete-time approximation.

1 Introduction

Let us consider the parabolic heat equation:

$$\left\{ \begin{array}{rcll} \dfrac{\partial v}{\partial t} + \dfrac{1}{2}\Delta_x v & = & 0, & \text{on } [0,T) \times \mathbb{R}^d, \\ v(T,.) & = & h, & \text{on } \mathbb{R}^d. \end{array} \right. \tag{1.1}$$

It is well-known that the solution to (1.1) is given by:

$$v(t,x) \;=\; \int h(y) K(T-t,x,y) dy,$$

where $K(t,x,y) = \frac{1}{(4\pi t)^{\frac{d}{2}}} e^{-|x-y|^2/4t}$ is the heat kernel on $\mathbb{R}^d$. By introducing the d-dimensional Brownian W on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and from the Gaussian distribution of W_t, we observe that the solution v can be represented also as:

$$v(t,x) \;=\; \mathbb{E}\big[h(x + W_{T-t})\big], \quad (t,x) \in [0,T] \times \mathbb{R}^d. \tag{1.2}$$

The probabilistic representation (1.2) gives a Monte-Carlo method for computing an approximation of v by the empirical mean:

$$v(t,x) \;\simeq\; \bar{v}^N(t,x) \;:=\; \frac{1}{N} \sum_{i=1}^{N} h(x + W^i_{T-t}),$$

where $(W^i)_{1 \le i \le N}$ is an N-sample drawn from an (exact) simulation of W. The convergence of $\bar{v}^N$ to v is ensured by the law of large numbers, when N goes to infinity, while the rate of convergence, obtained from the central limit theorem, is equal to $1/\sqrt{N}$, and independent of the dimension d of the heat equation. More generally, let us consider the linear parabolic partial differential equation (PDE):

$$\left\{ \begin{array}{rcll} \dfrac{\partial v}{\partial t} + \mathcal{L} v + f & = & 0, & \text{on } [0,T) \times \mathbb{R}^d, \\ v(T,.) & = & h, & \text{on } \mathbb{R}^d, \end{array} \right. \tag{1.3}$$

where $\mathcal{L}$ is the second order Dynkin operator:

$$\mathcal{L} v \;=\; b(x).D_x v + \frac{1}{2} \mathrm{tr}(\sigma\sigma^\intercal(x) D_x^2 v). \tag{1.4}$$

Under suitable conditions on the functions b, σ, f and h defined on $\mathbb{R}^d$, there exists a unique solution v to (1.3), which may be represented by the Feynman-Kac formula:

$$v(t,x) \;=\; \mathbb{E}\Big[\int_t^T f(X_s^{t,x}) ds + h(X_T^{t,x}) \Big], \quad (t,x) \in [0,T] \times \mathbb{R}^d, \tag{1.5}$$

where $X^{t,x}$ is the solution to the (forward) diffusion process,

$$dX_s \;=\; b(X_s) ds + \sigma(X_s) dW_s, \quad s \ge t,$$

starting from x at time t. Notice that the Feynman-Kac formula (1.5) can be easily derived from Itô's formula when v is smooth. Indeed, in this case, by defining the pair of processes (Y, Z):

$$Y_t := v(t, X_t), \qquad Z_t := \sigma^\intercal(X_t) D_x v(t, X_t), \quad 0 \leq t \leq T,$$

and applying Itô's formula to $v(s, X_s)$ between t and T, with v satisfying the PDE (1.3), we get:

$$Y_t = h(X_T) + \int_t^T f(X_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T. \tag{1.6}$$

This equation can be viewed as a backward stochastic equation in the pair of adapted processes (Y, Z) w.r.t. the filtration $\mathbb{F}^W$ generated by the Brownian motion W, determined from a terminal condition $h(X_T)$, and originally appeared in [4]. By taking conditional expectation in (1.6), we retrieve the Feynman-Kac formula (1.5). This probabilistic representation leads to a numerical method for solving the linear PDE, relying on Monte-Carlo simulations of the forward diffusion process X, whose convergence rate does not depend on the dimension of the problem, hence not suffering in principle of the curse of dimensionality encountered in deterministic numerical methods. On the other hand, it is also useful for computing explicitly the solution v in some particular models for X, e.g. geometric Brownian motion in the Black Scholes model for option pricing in finance.

In this paper, we address the problem of nonlinear PDEs, and shall review the recent developments about their probabilistic representation, i.e. nonlinear Feynman-Kac formulae. We shall first consider in Section 2 the case of semi-linear PDEs, i.e. when nonlinearity appears only on the first order derivative, and show how it is related to the theory of backward stochastic differential equations (BSDEs) introduced in [20], and leads to probabilistic scheme for solving semi-linear PDEs. We next consider in Section 3 the challenging problem of fully nonlinear PDEs, i.e. when nonlinearity enters also on the second order derivative. Such framework arises in many applications, for example in stochastic control in finance (portfolio optimization, risk management, model uncertainty). We shall present the randomization approach for dealing with such nonlinear context, and show how fully nonlinear PDES are represented in terms of randomized BSDEs with nonpositive jumps. This provides an original probabilistic scheme for solving fully nonlinear PDEs.

2 Backward SDEs and semi-linear PDEs

2.1 A short overview of BSDEs

Let us introduce some standard notations in the theory of backward stochastic differential equations (BSDEs). On a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a d-dimensional Brownian motion W over a finite time interval $[0, T]$, and its natural filtration $\mathbb{F} = \mathbb{F}^W$, we denote by:

- $\mathcal{P}_\mathbb{F}$: σ-algebra of $\mathbb{F}$-predictable subsets of $[0, T] \times \Omega$

- $\mathbb{S}^2_{\mathbb{F}}$: set of real-valued càd-làg $\mathbb{F}$-adapted processes Y such that

$$\mathbb{E}\big[\sup_{0\le t\le T}|Y_t|^2\big] \;<\; \infty,$$

- $\mathbb{L}^2_{\mathbb{F}}(W)$: set of $\mathbb{R}^d$-valued $\mathcal{P}_{\mathbb{F}}$-measurable processes Z such that

$$\mathbb{E}\left[\int_0^T |Z_t|^2 dt\right] \;<\; \infty.$$

We are given as data:

- a terminal condition ξ, which is an $\mathcal{F}_T$-measurable real-valued random variable
- a generator $f = (f_t(y,z))_{0\le t\le T}$, which is an $\mathcal{P}_{\mathbb{F}}\otimes\mathcal{B}(\mathbb{R}\times\mathbb{R}^d)$-measurable real-valued map, where $\mathcal{B}(\mathbb{R}\times\mathbb{R}^d)$ denotes the Borel σ-field of $\mathbb{R}\times\mathbb{R}^d$.

A (one dimensional) BSDE in differential form is written as

$$dY_t \;=\; -f_t(Y_t,Z_t)dt + Z_t dW_t, \quad 0\le t\le T, \;\; Y_T=\xi, \tag{2.1}$$

and a solution to (2.1) is a pair $(Y,Z)\in\mathbb{S}^2_{\mathbb{F}}\times\mathbb{L}^2_{\mathbb{F}}(W)$ satisfying

$$Y_t \;=\; \xi + \int_t^T f_s(Y_s,Z_s)ds - \int_t^T Z_s dW_s, \quad 0\le t\le T. \tag{2.2}$$

Existence and uniqueness of a solution to the BSDE (2.1) is proved in the seminal paper [20] under the following Lipschitz and square integrability assumptions:

(H1)

(i) f is uniformly Lipschitz in (y,z), i.e. there exists a positive constant C_f s.t. for all (y,z,y',z'):

$$|f_t(y,z)-f_t(y',z')| \;\le\; C_f\Big(|y-y'|+|z-z'|\Big), \;\; dt\otimes d\mathbb{P}\; a.e.$$

(ii) ξ and $\{f_t(0,0), t\in[0,T]\}$ are square integrable:

$$\mathbb{E}\Big[|\xi|^2 + \int_0^T |f_t(0,0)|^2 dt\Big] \;<\; \infty.$$

Notice that when the generator f does not depend on y and z, the solution to the BSDE (2.1), which is then simply a backward stochastic equation as in [4], is directly obtained from the martingale representation theorem applied to the Brownian martingale $M_t := \mathbb{E}\Big[\xi+\int_0^T f_s ds\big|\mathcal{F}_t\Big]$, $0\le t\le T$, which gives the existence of an integrand $Z\in\mathbb{L}^2_{\mathbb{F}}(W)$ s.t.

$$M_t \;=\; M_0 + \int_0^t Z_s dW_s, \quad 0\le t\le T.$$

Indeed, by defining

$$Y_t := M_t - \int_0^t f_s ds \;=\; \mathbb{E}\Big[\xi+\int_t^T f_s ds\big|\mathcal{F}_t\Big], \quad 0\le t\le T,$$

we see that (Y, Z) satisfies (2.2). In the general case where f depends on (y, z), the existence and uniqueness is proved by a fixed point argument under the Lipschitz assumption in **(H1)**(i).

Let us now consider the particular case of interest when the generator is linear, i.e. in the form:

$$f_t(y, z) \;=\; \delta_t y + \alpha_t.z + \gamma_t,$$

for some bounded $\mathbb{F}$-adapted processes (δ_t) valued in $\mathbb{R}$, (α_t) valued in $\mathbb{R}^d$, and $(\gamma_t) \in \mathbb{H}^2_{\mathbb{F}}$ the set of rela-valued $\mathbb{F}$-adapted processes s.t. $\mathbb{E}[\int_0^T |\gamma_t|^2 dt] < \infty$. By discounting and Girsanov's change of measure, the solution (in Y) to the linear BSDE (2.2) is given by the linear expectation:

$$Y_t \;=\; \mathbb{E}^{\mathbb{P}^\alpha}\big[e^{\int_t^T \delta_s ds}\xi + \int_t^T e^{\int_t^u \delta_u du}\gamma_s ds \big| \mathcal{F}_t\big],$$

where $\mathbb{P}^\alpha$ is the probability measure equivalent to $\mathbb{P}$ under which

$$W^\alpha \;:=\; W - \int \alpha dt, \quad \text{is a } \mathbb{P}^\alpha - \text{Brownian motion.} \tag{2.3}$$

Such context of linear BSDE arises typically in option pricing in finance, where $\mathbb{P}^\alpha$ is the martingale measure, Y is the fair price for the option payoff ξ, and Z the hedging portfolio. A trivial remark in that, in this linear case, if $\xi \geq 0$ and $\gamma \geq 0$, then the solution Y to the linear BSDE is also nonnegative. This is the key observation for showing comparison theorem for BSDEs: given two pairs (ξ, f) and (ξ', f') of terminal data/generators satisfying **(H1)**, and (Y, Z), (Y', Z') be the solutions to their BSDEs. Suppose that:

$$\xi \;\leq\; \xi' \; a.s. \quad \text{and} \quad f_t(Y_t, Z_t) \;\leq\; f'_t(Y_t, Z_t), \;\; dt \otimes d\mathbb{P} \; a.e.$$

Then,

$$Y_t \;\leq\; Y'_t, \quad 0 \leq t \leq T.$$

Let us next consider another case of interest where the generator $f_t(y, z)$ is convex in z, and is written in the form:

$$f_t(y, z) \;=\; \delta_t y + \sup_{a \in A}\big[a.z + \gamma_t(a)\big], \tag{2.4}$$

for some bounded adapted processes (δ_t), where A is a compact subset of $\mathbb{R}^d$, $\gamma_t(a)$ is a $\mathcal{P}_{\mathbb{F}} \otimes \mathcal{B}(A)$-measurable map s.t. $\operatorname*{ess\,sup}_{a \in A} |\gamma_t(a)| \in \mathbb{H}^2_{\mathbb{F}}$. By using comparison theorem for BSDEs and result for the linear case, one shows that the solution (in Y) to the BSDE is expressed as:

$$Y_t \;=\; \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}} \mathbb{E}^{\mathbb{P}^\alpha}\Big[e^{\int_t^T \delta_s ds}\xi + \int_t^T e^{\int_t^u \delta_u du}\gamma_s(\alpha_s) ds \big| \mathcal{F}_t\Big], \quad 0 \leq t \leq T,$$

where $\mathcal{A}$ is the set of adapted processes α valued in A and $\mathbb{P}^\alpha$ is the probability measure equivalent to $\mathbb{P}$ under which the drifted process W^α defined in (2.3) is a $\mathbb{P}^\alpha$-Brownian motion. Hence, controlled drift problems and risk measures with uncertain drifts are related to BSDE by choosing a generator in the form (2.4). We refer to [8], [9] or [22] for a more detailed review and applications of BSDEs.

2.2 Markov BSDEs and PDEs

We put ourselves in a Markov setting in the sense that we suppose that the terminal data and generator of the BSDE (2.2) are in the form:

$$\xi \;=\; h(X_T), \qquad f_t(\omega, y, z) \;=\; f(X_t(\omega), y, z)$$

where $h(x)$ is some measurable function on $\mathbb{R}^d$, $f(x, y, z)$ is some measurable function on $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ (we kept the same notation f by misuse), and X is a forward diffusion process of dynamics:

$$dX_s \;=\; b(X_s)ds + \sigma(X_s)dW_s \quad \text{in } \mathbb{R}^d. \tag{2.5}$$

Under standard Lipschitz assumptions on the coefficients $b : \mathbb{R}^d \mapsto \mathbb{R}^d$, and $\sigma : \mathbb{R}^d \mapsto \mathbb{R}^{d\times d}$, there exists a unique strong solution to (2.5) given some initial condition, and we have the standard estimate:

$$\mathbb{E}\big[\sup_{0\le t\le T} |X_t|^2\big] \;\le\; C(1 + |X_0|^2).$$

A forward BSDE is then written as:

$$Y_t \;=\; h(X_T) + \int_t^T f(X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad 0 \le t \le T, \tag{2.6}$$

and under Lipschitz condition on f, and linear growth condition on h, there exists a unique solution (Y, Z) to the Markov BSDE (2.6). Denoting by $(Y^{t,x}, Z^{t,x})_{t\le s\le T}$ the solution to the BSDE (2.6) when $X = (X^{t,s})_{t\le s\le T}$ is the solution to (2.5) starting from x at time t, we notice that

$$v(t, x) \;:=\; Y_t^{t,x}, \quad (t, x) \in [0, T] \times \mathbb{R}^d, \tag{2.7}$$

is a deterministic function on $[0, T] \times \mathbb{R}^d$, and by the Markov property of the diffusion process, we have:

$$Y_t \;=\; v(t, X_t), \quad 0 \le t \le T.$$

Let us now derive formally the PDE satisfied by the function v. By definition of the Markov BSDE (2.6), we have:

$$Y_s - Y_t \;=\; v(s, X_s) - v(t, X_t) \;=\; -\int_t^s f(X_u, Y_u, Z_u)ds + \int_t^s Z_u dW_u,$$

for all $0 \le t \le s \le T$. Assuming that v is smooth, it follows from Itô's formula:

$$\begin{aligned} &\int_t^s (\frac{\partial v}{\partial t} + \mathcal{L}v)(u, X_u)du + \int^s \sigma^\intercal(X_u) D_x v(u, X_u) dW_u \\ =\; &-\int_t^s f(X_u, Y_u, Z_u)du + \int_t^s Z_u dW_u, \end{aligned}$$

where $\mathcal{L}$ is the Dynkin operator associated to the diffusion X, and given in (1.4). Identifying the finite variation terms in "dt" and the Brownian martingale terms in "dW", we see that

$$Z_t \;=\; \sigma^\intercal(X_t) D_x v(t, X_t)$$

and v should satisfy the semi-linear parabolic PDE:

$$\left\{ \begin{array}{rcl} \dfrac{\partial v}{\partial t} + \mathcal{L}v + f(x, v, \sigma^\intercal D_x v) &=& 0, \quad \text{on } [0,T) \times \mathbb{R}^d, \\ v(T,.) &=& h, \quad \text{on } \mathbb{R}^d. \end{array} \right. \tag{2.8}$$

The main issue in this derivation comes from the fact that in general, the function v is not smooth, and this is overcome with the notion of viscosity solution: it is proved in [21] that the function v in (2.7) is the unique viscosity solution to (2.8). Therefore, the BSDE (2.6) provides a probabilistic representation to the solution of the semi-linear PDE (2.8). This extends the Feynman-Kac formula (1.5) to the case where $f = f(x, y, z)$ depends on y, z, and we shall see in the next paragraph how it can be used to design a probabilistic numerical scheme for computing the solution to the semi-linear PDE (2.8).

2.3 Numerical issues

The first step in the numerical scheme for the resolution of the BSDE (2.6) is the discrete-time approximation. It is constructed as follows.

• *Euler scheme for the forward process.* We are given a time grid $\pi := \{t_0 = 0 < t_1 < \ldots < t_n = T\}$ of $[0,T]$, with modulus $|\pi| := \max_{i=1,\ldots,n} \Delta t_i$, $\Delta t_i := t_{i+1} - t_i$, and approximate the forward diffusion process X by its Euler scheme X^π defined as

$$X^\pi_{t_{i+1}} \;:=\; X^\pi_{t_i} + b(X^\pi_{t_i})\Delta t_i + \sigma(X^\pi_{t_i})\Delta W_{t_i}, \quad i < n, \quad X^\pi_0 \;=\; X_0,$$

where $\Delta W_{t_i} := W_{t_{i+1}} - W_{t_i}$.

• *Euler scheme for the backward process.* We first approximate the terminal condition $Y_T = h(X_T)$ by simply replacing X by its Euler scheme: $Y_T \simeq h(X^\pi_T)$. Then, from the formal Euler backward discretization:

$$\begin{array}{rcl} Y_{t_i} &=& Y_{t_{i+1}} + \displaystyle\int_{t_i}^{t_{i+1}} f(X_s, Y_s, Z_s) ds - \int_{t_i}^{t_{i+1}} Z_s dW_s \\ &\simeq& Y_{t_{i+1}} + f(X^\pi_{t_i}, Y_{t_i}, Z_{t_i})\Delta t_i - Z_{t_i}\Delta W_{t_i}. \end{array}$$

we define the discrete-time approximation of the BSDE as follows:

(1) taking expectation conditionally on $\mathcal{F}_{t_i}$ on both sides yields

$$Y_{t_i} \;\simeq\; \mathbb{E}\big[Y_{t_{i+1}} | \mathcal{F}_{t_i}\big] + f(X^\pi_{t_i}, Y_{t_i}, Z_{t_i})\Delta t_i$$

(2) Multiplying by ΔW_{t_i} and then taking conditional expectation gives

$$0 \;\simeq\; \mathbb{E}\big[Y_{t_{i+1}} \Delta W_{t_i} | \mathcal{F}_{t_i}\big] - Z_{t_i}\Delta t_i.$$

This formal approximation argument leads to a backward Euler scheme (Y^π, Z^π) of the form:

$$\left\{ \begin{array}{rcl} Z^\pi_{t_i} & = & \mathbb{E}\Big[Y^\pi_{t_{i+1}} \frac{\Delta W_{t_i}}{\Delta t_i} \big| \mathcal{F}_{t_i}\Big], \\ Y^\pi_{t_i} & = & \mathbb{E}\big[Y^\pi_{t_{i+1}} | \mathcal{F}_{t_i}\big] + f(X^\pi_{t_i}, Y^\pi_{t_i}, Z^\pi_{t_i})\Delta t_i, \quad i < n, \end{array} \right. \tag{2.9}$$

with terminal condition $Y^\pi_{t_n} = h(X^\pi_{t_n})$.

Remark 2.1 The above scheme is implicit as $Y^\pi_{t_i}$ appears in both sides of the equation. Since f is assumed to be Lipschitz and since it is multiplied by Δt_i, intended to be small, the equation can be solved numerically very quickly by standard fixed point methods. Alternatively, we could also consider an explicit scheme by replacing the second equation in (2.9) by

$$Y^\pi_{t_i} \quad = \quad \mathbb{E}\big[Y^\pi_{t_{i+1}} + f(X^\pi_{t_i}, Y^\pi_{t_{i+1}}, Z^\pi_{t_i})\Delta t_i | \mathcal{F}_{t_i}\big].$$

This will not change the convergence rate. □

The discrete-time approximation is measured by the squared error:

$$\mathcal{E}(\pi)^2 \quad := \quad \max_{i\le n} \mathbb{E}\big[|Y_{t_i} - Y^\pi_{t_i}|^2\big] + \sum_{i=0}^{n-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - Z^\pi_{t_i}|^2 dt\Big].$$

By using Itô's formula, Gronwall's lemma and Young inequality, it is proved in [7] and [27] that under the Lipschitz assumption on the driver f, there exists a constant C independent of π such that:

$$\begin{array}{rcl} \mathcal{E}(\pi)^2 & \le & C\Big(\mathbb{E}\big|h(X_T) - h(X^\pi_T)|^2 + \max_{i\le n} \mathbb{E}\big|X_{t_i} - X^\pi_{t_i}\big|^2 \\ & & \quad + \sum_{i=0}^{n-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt\Big]\Big), \end{array} \tag{2.10}$$

where $\bar{Z}_{t_i} = \frac{1}{\Delta t_i}\mathbb{E}\big[\int_{t_i}^{t_{i+1}} Z_t dt | \mathcal{F}_{t_i}\big]$. In other words, we have three different error contributions:

1. Strong approximation of the terminal condition, which depends on the terminal data and the forward Euler scheme
2. Strong approximation of the forward SDE, which depends on the forward Euler scheme, but not on the BSDE problem
3. L^2-regularity of Z, which is intrinsic to the BSDE problem.

It is well-known, see e.g. [16], that the strong approximation of the forward Euler scheme: $\max_{i\le n} \mathbb{E}\big|X_{t_i} - X^\pi_{t_i}\big|^2$ provides an error of order $|\pi|$. Consequently, when the terminal data h is Lipschitz, this also gives an error for $\mathbb{E}\big|h(X_T) - h(X^\pi_T)|^2$ of order $|\pi|$. Finally, it is proved in [27] that under Lipschitz condition on f, we have the L^2-regularity of Z:

$\sum_{i=0}^{n-1} \mathbb{E}\big[\int_{t_i}^{t_{i+1}} |Z_t - Z_{t_i}^\pi|^2 dt\big] = O(|\pi|)$. Therefore, under Lipschiz assumption on f and h, the rate of convergence of the discrete-time approximation error $\mathcal{E}(\pi)$ is of order $|\pi|^{\frac{1}{2}}$:

$$\mathcal{E}(\pi) \quad \leq \quad C|\pi|^{\frac{1}{2}}.$$

This convergence rate is clearly optimal and similar to the one obtained for forward SDEs.

The practical implementation of the numerical scheme (2.9) requires the computation of conditional expectations. The key observation in our Markovian context is that all these conditional expectations are regressions, i.e.

$$\mathbb{E}\big[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\big] \; = \; \mathbb{E}\big[Y_{t_{i+1}}^\pi | X_{t_i}^\pi\big], \qquad \mathbb{E}\big[Y_{t_{i+1}}^\pi \Delta W_{t_i} | \mathcal{F}_{t_i}\big] \; = \; \mathbb{E}\big[Y_{t_{i+1}}^\pi \Delta W_{t_i} | X_{t_i}^\pi\big],$$

which can be approximated by methods from statistics:

- *Quantization.* Each $X_{t_i}^\pi$ is replaced by a quantized version, i.e. a projection on a finite grid, which is computed in some optimal way based on stochastic algorithm and Monte-Carlo simulations of X^π (Kohonen). The conditional expectation is then reduced to a discrete sum with weights also computed off line as the grid points. We refer to [1], [19] and the references therein.

- *Integration by parts.* The conditional expectation is approximated via an integration by parts formula and Malliavin calculus, see [5].

- *Least-square regression.* The conditional expectation is approximated by non-parametric regression methods, and the most popular one, known as Longstaff-Schwartz method [18], consists in the projection on a set of basis functions, with optimal coefficients computed from empirical least-square based on Monte-Carlo simulations of $X_{t_i}^\pi$. We refer to [11] for more details and analysis of convergence rate of this approach.

The advantage of these probabilistic methods, based on Monte-Carlo simulations, is that the convergence rate does not depend a priori on the dimension of the problem, and therefore should less suffer from the curse of dimensionality encountered in deterministic procedures.

3 Randomization approach for fully nonlinear HJB equation

3.1 Motivating example

Let us consider the following controlled diffusion example arising from uncertain volatility model in finance:

$$X_s^{t,x,\alpha} \quad = \quad x + \int_t^s \alpha_u dW_u, \quad 0 \leq t \leq s \leq T, \;\; x \in \mathbb{R},$$

where α is an adapted process valued in $A = [\underline{a}, \bar{a}]$, $0 < \underline{a} \leq \bar{a} < \infty$, denoted $\alpha \in \mathcal{A}$, interpreted as the uncertain volatility of the stock price X. We define the value function of the stochastic control problem:

$$v(t,x) \quad := \quad \sup_{\alpha \in \mathcal{A}} \mathbb{E}\big[h(X_T^{t,x,\alpha})\big],$$

which is interpreted as the super-replication cost of an option payoff h under uncertain volatility. The dynamic programming equation (also called Hamilton-Jacobi-Bellman, HJB in short) for this stochastic control problem (see e.g. [24]) is a fully nonlinear PDE in the form:

$$\frac{\partial v}{\partial t} + G(D_x^2 v) \;=\; 0, \tag{3.1}$$

with terminal condition $v(T,.) = h$, and where

$$G(M) \;:=\; \frac{1}{2}\sup_{a\in A}[a^2 M] \;=\; \bar{a}^2 M^+ - \underline{a}^2 M^-, \quad M \in \mathbb{R}.$$

The equation (3.1) can be viewed as a G-heat equation (reducing to the classical heat equation when A is a singleton), and based on this observation, Peng [23] has developed a theory of G-stochastic calculus with G-Brownian motion, extending the classical Itô calculus with Brownian motion, and leading to the concept of nonlinear expectation. Denoting by $B_t^\alpha = \int_0^t \alpha_s dW_s$, and $\mathbb{P}^\alpha$ the law of B^α under $\mathbb{P}$, we notice that $(\mathbb{P}^\alpha)_\alpha$ is a family of non dominated probability measures, and this contrasts with the framework of controlled drift problem in (2.3), which gave rise to equivalent probability measures by Girsanov's theorem. Recalling that Brownian motion and Itô calculus are the basic tools for defining BSDE, Soner, Touzi and Zhang [25] have developed the theory of 2BSDE in connection with G-Brownian motion by using notions from quasi-sure analysis in a singular measures framework. However, the main concerns with the theory of G-expectation and 2BSDE is that (i) it does not cover the general case of HJB equation where control appears both on the drift and diffusion, (ii) it requires uniform ellipticity condition on the diffusion coefficient, (iii) it does not lead clearly to an implementable numerical scheme since one cannot simulate a G-Brownian motion. In the rest of this paper, we shall present an alternative approach for overcoming these issues.

3.2 BSDE with nonpositive jumps

Let us consider the fully nonlinear PDE of HJB type:

$$\left\{\begin{array}{rcl} \dfrac{\partial v}{\partial t} + \sup\limits_{a\in A}\Big[b(x,a).D_x v + \dfrac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x,a)D_x^2 v) & & \\ \qquad + f(x,a,v,\sigma^\intercal(x,a)D_x v)\Big] &=& 0, \quad \text{on } [0,T)\times\mathbb{R}^d \\ v(T,.) &=& h, \quad \text{on } \mathbb{R}^d, \end{array}\right. \tag{3.2}$$

where A is a compact metric space, $b = b(x,a)$ is an $\mathbb{R}^d$-valued Lipschitz continuous function, $\sigma = \sigma(x,a)$ is an $\mathbb{R}^{d\times d}$-valued (possibly degenerate) Lipschitz continuous function, $f = f(x,a,y,z)$, $h = h(x)$ are Lipschitz continuous functions. Under these conditions, there exists a unique viscosity solution with linear growth condition to (3.2), see [12]. An important particular case is when $f = f(x,a)$ does not depend on y,z, and then the PDE (3.2) corresponds to the dynamic programming equation for the stochastic control problem:

$$v(t,x) \;=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[h(X_T^{t,x,\alpha}) + \int_t^T f(X_s^{t,x,\alpha},\alpha_s)ds\Big], \tag{3.3}$$

with the controlled diffusion process in $\mathbb{R}^d$:

$$X_s^{t,x,\alpha} = x + \int_t^s b(X_u^{t,x,\alpha}, \alpha_u)du + \int_t^s \sigma(X_u^{t,x,\alpha}, \alpha_u)dW_u, \quad t \le s \le T, \tag{3.4}$$

where W is a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, and $\alpha \in \mathcal{A}$ is the control, i.e. an $\mathbb{F}$-adapted process valued in A. HJB type equations (3.2) include the G-heat equation (3.1) and arise in many applications in finance, like portfolio optimization, option pricing and risk measures under model uncertainty, etc. We refer to [10] or [24] for an expository treatment of the theory of stochastic control and its applications.

The main issue for a Feynman-Kac type formula of the fully nonlinear PDE (3.2) comes from the fact that the controlled forward process X^α in (3.4) cannot be simulated for all values of the control α, and one cannot remove the control process as in the controlled drift case by Girsanov's theorem. We present here a control randomization approach, whose basic idea is to replace the control process by an (uncontrolled) auxiliary state variable process running over the control set A, hence simulatable, and under which one can apply Girsanov's theorem in order to recover all possible values of the original control process. As we shall see, this method allows us to provide a BSDE representation of general HJB equation (3.2) in terms of a simulatable forward process formulated under a single probability measure, hence a non linear Feynman-Kac formula. An important feature of our approach is that it does not require any ellipticity condition on the diffusion coefficient. Moreover, by using a randomization with jumps, we are able to derive a practical probabilistic numerical scheme, which can take advantage of Monte-Carlo methods for dealing with high dimensional problems, both in state and control space.

Let us introduce a Poisson random measure $\mu(dt, da)$ on $\mathbb{R}_+ \times A$ (hence independent of W), with jump times (T_k), and marks (ζ_k), and intensity measure $\lambda(da)dt$ where λ is a finite measure supporting the whole set A. We denote by $\tilde{\mu}(dt, da) = \mu(dt, da) - \lambda(da)dt$ the compensated martingale measure of μ, and associate to μ the pure-jump process I defined by:

$$I_t = \zeta_i, \quad T_k \le t < T_{k+1}, \quad k \in \mathbb{N},$$

which is also written in differential form as:

$$dI_t = \int_A (a - I_{t^-})\mu(dt, da), \quad t \ge 0.$$

We then consider the regime-switching process of dynamics:

$$dX_t = b(X_t, I_t)dt + \sigma(X_t, I_t)dW_t, \quad t \ge 0.$$

In other words, we have replaced in the dynamics (3.4), the control α by the exogenous pure jump process I. Notice that the pair (X, I) is a Markov process valued in $\mathbb{R}^d \times A$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the Brownian-Poisson filtration $\mathbb{G} = \mathbb{F}^{W,\mu} = (\mathcal{G}_t)_{0 \le t \le T}$. We next consider the BSDE with jumps, consisting in the search for a triple

(Y, Z, U) satisfying:

$$\begin{aligned} Y_t &= h(X_T) + \int_t^T f(X_s, I_s, Y_s, Z_s)ds \\ &\quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a)\tilde{\mu}(ds, da), \quad 0 \le t \le T. \end{aligned} \tag{3.5}$$

Here, the pair (Y, Z) lie in $\mathbb{S}^2_{\mathbb{G}} \times \mathbb{L}^2_{\mathbb{G}}(W)$, and with respect to the Brownian framework, there is in addition the component U lying in $\mathbb{L}^2_{\mathbb{G}}(\tilde{\mu})$, the set of $\mathcal{P}_{\mathbb{G}} \otimes \mathcal{B}(A)$-measurable maps $(U_t(a))_{0\le t\le T}$ such that $\mathbb{E}[\int_0^T \int_A |U_t(a|^2 \lambda(da)dt] < \infty$. Existence and uniqueness of a triple solution $(Y, Z, U) \in \mathbb{S}^2_{\mathbb{G}} \times \mathbb{L}^2_{\mathbb{G}}(W) \times \mathbb{L}^2_{\mathbb{G}}(\tilde{\mu})$ to the BSDE with jumps (3.5) is proved in [26], extending the result of [20]. Moreover, by the Markov property of the forward regime-switching process (X, I), the component solution Y is written in the form:

$$Y_t = v(t, X_t, I_t),$$

for some deterministic function v on $[0, T] \times \mathbb{R}^d \times A$, which satisfies (in the viscosity sense) the semi-linear integro-partial differential equation (IPDE):

$$\frac{\partial v}{\partial t} + \bar{\mathcal{L}}v + \mathcal{M}v + f(x, a, v, \sigma^\intercal(x, a)D_x v) = 0, \quad (t, x, a) \in [0, T) \times \mathbb{R}^d \times A, \tag{3.6}$$

where

$$\begin{aligned} \bar{\mathcal{L}}v(t, x, a) &= b(x, a).D_x v + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x, a)D_x^2 v), \\ \mathcal{M}v(t, x, a) &= \int_A \big(v(t, x, a') - v(t, x, a)\big)\lambda(da'). \end{aligned}$$

In other words, the BSDE with jumps (3.5) provides a Feynman-Kac formula for the semi-linear IPDE (3.6), as proved in [2], thus extending the result of [21]. Moreover, when v is smooth, we have by Itô's formula, the connection:

$$\begin{aligned} Y_t &= v(t, X_t, I_t), \qquad Z_t = \sigma^\intercal(X_t, I_{t^-})D_x v(t, X_t, I_{t^-}), \\ &\text{and} \qquad U_t(a) = v(t, X_t, a) - v(t, X_t, I_{t^-}). \end{aligned}$$

The issue is now to go from the semi-linear IPDE (3.6) to the fully nonlinear PDE (3.2), and the idea is to constrain the jump component U of the BSDE with jumps (3.5) to be nonpositive. Let us formally derive the arguments of this approach. By constraining the jump component, this means in terms of the function v (when it is smooth) that:

$$U_t(a) = v(t, X_t, a) - v(t, X_t, I_{t^-}) \le 0, \quad \text{for all } (t, a) \in [0, T] \times A,$$

which would imply that $v = v(t, x)$ does not depend actually on $a \in A$. Thus, the integral term $\mathcal{M}v$ in (3.6) is removed, and the variable a becomes now a parameter in the PDE satisfied by $v(t, x)$ on $[0, T) \times \mathbb{R}^d$:

$$\frac{\partial v}{\partial t} + b(x, a).D_x v + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x, a)D_x^2 v) + f(x, a, v, \sigma^\intercal(x, a)D_x v) = 0, \quad \text{on } [0, T) \times \mathbb{R}^d,$$

which should hold for any parameter value $a \in A$. By taking supremum over $a \in A$, we formally expect to retrieve the fully nonlinear HJB equation (3.2).

The rigorous derivation of the above argument is formulated in [15] by means of the class of BSDE with nonpositive jumps: this consists in a triple $(Y, Z, U) \in \mathbb{S}^2_{\mathbb{G}} \times \mathbb{L}^2_{\mathbb{G}}(W) \times \mathbb{L}^2_{\mathbb{G}}(\tilde{\mu})$ supersolution to:

$$\begin{aligned} Y_t &\geq h(X_T) + \int_t^T f(X_s, I_s, Y_s, Z_s) ds \\ &\quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da), \quad 0 \leq t \leq T. \end{aligned} \tag{3.7}$$

such that

$$U_t(a) \leq 0, \qquad d\mathbb{P} \otimes dt \otimes \lambda(da) \ \ a.e. \text{ on } \Omega \times [0,T] \times A. \tag{3.8}$$

By supersolution in (3.7), we mean that inequality $\geq$ holds instead of $=$ as in (3.5), and this relaxation is done in order to get flexibility for satisfying the non positivity constraint in (3.8). We are then looking for a minimal supersolution (Y, Z, U) to (3.7)-(3.8) in the sense that for any other triple $(\tilde{Y}, \tilde{Z}, \tilde{U}) \in \mathbb{S}^2_{\mathbb{G}} \times \mathbb{L}^2_{\mathbb{G}}(W) \times \mathbb{L}^2_{\mathbb{G}}(\tilde{\mu})$ satisfying (3.7)-(3.8), we have:

$$Y_t \leq \tilde{Y}_t, \quad 0 \leq t \leq T, \ a.s.$$

Existence and uniqueness of a minimal supersolution (Y, Z, U) to (3.7)-(3.8) is shown in [15] by penalization methods. Moreover, it is proved that the solution Y is in the form

$$Y_t = v(t, X_t), \tag{3.9}$$

for some deterministic function v on $[0,T] \times \mathbb{R}^d$ (hence not depending on the state variable I_t, and this is the key property), and v is the unique viscosity solution to the nonlinear PDE (3.2). Therefore, we have a nonlinear Feynman-Kac formula for the solution to the fully nonlinear PDE (3.2) in terms of BSDE with nonpositive jumps (3.7)-(3.8). The last paragraph shows how this probabilistic representation provides a numerical scheme for solving the PDE (3.2).

Remark 3.1 It is also proved in [15] that, in the case where $f = f(x, a)$ does not depend on y, z, the minimal solution Y to the BSDE with nonpositive jumps (3.7)-(3.8) admits a dual representation in the form:

$$Y_t = \operatorname*{ess\,sup}_{\nu \in \mathcal{D}} \mathbb{E}^{\mathbb{P}^\nu} \Big[h(X_T) + \int_t^T f(X_s, I_s) ds \big| \mathcal{G}_t \Big], \quad 0 \leq t \leq T,$$

where $\mathcal{D}$ is the set of $\mathcal{P}_{\mathbb{G}} \otimes \mathcal{B}(A)$-measurable maps $\nu = (\nu_t(a))_{0 \leq t \leq T}$, valued in $[1, \infty)$ and bounded, and $\mathbb{P}^\nu$ is the probability measure equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{G}_T)$ whose effect by Girsanov's theorem is to change the compensator $\lambda(da)dt$ of μ under $\mathbb{P}$ to $\nu_t(a)\lambda(da)dt$ under $\mathbb{P}^\nu$, and to leave unchanged the Brownian motion W. Together with (3.9), this shows that the value function of a stochastic control problem (3.3) admits an alternative formulation in terms of intensity control. □

3.3 Numerical scheme

The discrete-time approximation for the minimal supersolution to the BSDE with nonpositive jumps (3.7)-(3.8) is constructed as follows. First, we observe that the pure jump process I from the Poisson random measure $\mu(dt,da)$ with jump times/marks $(T_k,\zeta_k)_k$ and intensity measure $\lambda(da)dt$ is perfectly simulated. Indeed, the inter arrival times $S_k = T_{k+1} - T_k$ are i.i.d and follow an exponential law of parameter $\lambda := \int_A \lambda(da)$, while the marks ζ_k are i.i.d. with distribution $\bar{\lambda}(da) = \lambda(da)/\lambda$, assumed to be simulatable. We thus simulate I by:

$$I_t \;=\; I_0 1_{[0,T_1)} + \sum_{k\geq 1} \zeta_k 1_{[t_k,T_{k+1})}(t), \quad t \geq 0.$$

We are next given a time grid $\pi := \{t_0 = 0 < t_1 < \ldots < t_n = T\}$ of $[0,T]$, with modulus $|\pi|$ $:= \max_{i=1,\ldots,n} \Delta t_i$, $\Delta t_i := t_{i+1} - t_i$, and approximate the forward regime-switching process X by its Euler scheme X^π defined as

$$X^\pi_{t_{i+1}} \;:=\; X^\pi_{t_i} + b(X^\pi_{t_i}, I_{t_i})\Delta t_i + \sigma(X^\pi_{t_i}, I_{t_i})\Delta W_{t_i}, \quad i < n, \quad X^\pi_0 \;=\; X_0,$$

where $\Delta W_{t_i} := W_{t_{i+1}} - W_{t_i}$. We then propose a discrete time approximation explicit scheme in the form:

$$\left\{ \begin{array}{rcl} Y^\pi_T \;=\; \mathcal{Y}^\pi_T &=& g(X^\pi_T) \\ \mathcal{Z}^\pi_{t_i} &=& \mathbb{E}\Big[Y^\pi_{t_{i+1}} \frac{\Delta W_{t_i}}{\Delta t_i} \big| \mathcal{G}_{t_i}\Big] \\ \mathcal{Y}^\pi_{t_i} &=& \mathbb{E}\Big[Y^\pi_{t_{i+1}} \big| \mathcal{G}_{t_i}\Big] + f(X^\pi_{t_i}, I_{t_i}, \mathcal{Y}^\pi_{t_i}, \mathcal{Z}^\pi_{t_i})\Delta t_i \\ Y^\pi_{t_i} &=& \operatorname*{ess\,sup}_{a\in A} \mathbb{E}\Big[\mathcal{Y}^\pi_{t_i} \big| \mathcal{G}_{t_i}, I_{t_i} = a\Big], \quad i = 0,\ldots,n-1. \end{array} \right. \tag{3.10}$$

The interpretation of this scheme is the following. The first three lines in (3.10) correspond to the scheme $(\mathcal{Y}^\pi, \mathcal{Z}^\pi)$ for a discretization of a BSDE with jumps, as in [6], and extending the scheme described in paragraph 2.3 (we omit here the computation of the jump component). The last line in (3.10) for computing the approximation Y^π of the minimal supersolution Y corresponds precisely to the minimality condition for the nonpositive jump constraint and should be understood as follows. By the Markov property of the forward process (X,I), the solution $(\mathcal{Y},\mathcal{Z},\mathcal{U})$ to the BSDE with jumps (without constraint) is in the form $\mathcal{Y}_t = \vartheta(t,X_t,I_t)$ for some deterministic function ϑ. Assuming that ϑ is a continuous function, the jump component of the BSDE, which is induced by a jump of the forward component I, is equal to $\mathcal{U}_t(a) = \vartheta(t,X_t,a) - \vartheta(t,X_t,I_{t^-})$. Therefore, the nonpositive jump constraint means that: $\vartheta(t,X_t,I_{t^-}) \geq \operatorname*{ess\,sup}_{a\in A} \vartheta(t,X_t,a)$. The minimality condition is thus written as:

$$Y_t \;=\; v(t,X_t) \;=\; \operatorname*{ess\,sup}_{a\in A} \vartheta(t,X_t,a) \;\;=\;\; \operatorname*{ess\,sup}_{a\in A} \mathbb{E}[\mathcal{Y}_t | X_t, I_t = a],$$

whose discrete time version is the last line in scheme (3.10). Notice that the scheme (3.10) is a dynamic programming type algorithm. The novel feature is that conditional expectation is taken with respect to the uncontrolled randomized extended state process (X,I), and supremum with respect to the auxiliary state variable I.

The discrete-time approximation error is measured by:

$$\mathrm{Err}^{\pi}_{\pm}(Y) \;:=\; \max_{i\le n}\Big(\mathbb{E}\Big[\big(Y_{t_i}-\bar{Y}^{\pi}_{t_i}\big)^2_{\pm}\Big]\Big)^{\frac{1}{2}}$$

It is proved in [13] that

$$\mathrm{Err}^{\pi}_{-}(Y) \;\le\; C|\pi|^{\frac{1}{2}},$$

and under additional conditions on b, σ, f and h, namely: b, σ bounded, $f = f(x, a, y)$ does not depend on z, and is convex in y, and $f(.,.,0)$, g are bounded, we have

$$\mathrm{Err}^{\pi}_{+}(Y) \;\le\; \begin{cases} C|\pi|^{\frac{1}{6}} & \text{when } f = f(x,a) \\ C|\pi|^{\frac{1}{10}} & \text{otherwise.} \end{cases}$$

In particular,

$$-C|\pi|^{\frac{1}{2}} \le v(0,X_0) - Y^{\pi}_0 \;\le\; \begin{cases} C|\pi|^{\frac{1}{6}} & \text{when } f = f(x,a) \\ C|\pi|^{\frac{1}{10}} & \text{otherwise.} \end{cases}$$

The above error bounds are non symmetric as in deterministic methods, and are proved by using shaking coefficients method of Krylov [17] and switching system approximation of Barles and Jacobsen [3].

The last step towards an implementable scheme consists in the approximation of the conditional expectations in (3.10). Here, due to the supremum operation, there is a strong advantage of using least-square regression methods. Let us briefly recall the basic principle of this method. From the definition-property of conditional expectation:

$$\mathbb{E}[H|\mathcal{G}_{t_i}] \;=\; \arg\inf_{V\in L^2(\mathcal{G}_{t_i})} \mathbb{E}|H-V|^2,$$

we approximate it by $\hat{\mathbb{E}}_{t_i}[H] := \hat{\varphi}_i(X^{\pi}_{t_i}, I_{t_i})$ with empirical regression function:

$$\hat{\varphi}_i \;:=\; \arg\inf_{\varphi\in\Phi} \frac{1}{M}\sum_{m=1}^{M}\big(H^m - \varphi(X^{\pi,m}_{t_i}, I^m_{t_i})\big)^2$$

where

- $(X^{\pi,m}_{t_i}, I^m_{t_i})_m$ and $(H^m)_m$ are i.i.d. realizations of $(X^{\pi}_{t_i}, I_{t_i})$ and H
- $\Phi = \mathrm{Span}\{\phi^\ell : 1 \le \ell \le L_\Phi\}$, ϕ^ℓ basis functions on $\mathbb{R}^d \times A$.

Then, $\operatorname{ess\,sup}_{a\in A} \mathbb{E}[H|\mathcal{G}_{t_i}, I_{t_i} = a]$ is approximated by:

$$\operatorname*{ess\,sup}_{a\in A} \hat{\mathbb{E}}_{t_i,a}[H] \;:=\; \operatorname*{ess\,sup}_{a\in A} \hat{\varphi}_i(X^{\pi}_{t_i}, a) \;=\; \hat{\varphi}_i(X^{\pi}_{t_i}, \hat{a}_i(X^{\pi}_{t_i}))$$

where $\hat{a}_i$ is determined by:

$$\hat{a}_i(x) \;:=\; \arg\max_{a\in A} \hat{\varphi}_i(x,a),$$

hence in nonparametric form given the choice of the basis functions in Φ. The advantage of this regression-projection method is that we don't need to run over the set A in the optimization over a, e.g. by Newton method, as in finite-difference methods, and this quite interesting especially in high dimension for the control space A. Moreover, we get an approximate optimal control in feedback form, i.e. a deterministic function $\hat{a}_i(x)$ of the state value x at any date t_i. Error analysis and numerical illustrations of this algorithm are studied and performed in [14].

References

[1] Bally V. and G. Pagès (2002): "A quantization algorithm for solving discrete time multidimensional optimal stopping problems", *Bernoulli*, 9, 100-1049

[2] Barles G., R. Buckdahn, and E. Pardoux (1997): "Backward stochastic differential equations and integral-partial differential equations", *Stochastics and Stochastics Reports*, 60, 57-83.

[3] Barles G. and E.R. Jacobsen (2007): "Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations", *Math. Computation*, 76, 1861-1893.

[4] Bismut J.M. (1973): Analyse convexe et probabilités, thèse, Faculté des sciences de Paris.

[5] Bouchard B., Ekeland I. and N. Touzi (2004): "On the Malliavin approach to Monte Carlo approximation of conditional expectations", *Finance and Stochastics*, 8, 45-71.

[6] Bouchard B. and R. Elie (2008): "Discrete time approximation of decoupled FBSDE with jumps", *Stochastic Processes and Applications*, 118 (1), 53-75.

[7] Bouchard B. and N. Touzi (2004): "Discrete time approximation and Monte-Carlo simulation of BSDEs", *Stochastic processes and their applications*, 111, 175-206.

[8] El Karoui, N., Mazliak L.(1997): Backward stochastic differential equations, *Pitman research notes*, 364.

[9] El Karoui, N., Peng S., and M.C. Quenez (1997): "Backward stochastic differential equations in finance" *Mathematical Finance*, 7, 1-71.

[10] Fleming W. H. and H. M. Soner (2006): Controlled Markov Processes and Viscosity Solutions, 2nd edition, Springer-Verlag.

[11] Gobet E., Lemor J.P. and X. Warin (2006): "Rate of convergence of empirical regression method for solving generalized BSDE", *Bernoulli*, 12, 889-916.

[12] Ishii H. (1989): "On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDE's", *Communications on Pure and Applied Mathematics*, 42, 15-45.

[13] Kharroubi I., Langrené N. and H. Pham (2013): "Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps", to appear in *Annals of Applied Probability*.

[14] Kharroubi I., Langrené N. and H. Pham (2014): "A numerical algorithm for fully nonlinear HJB equations: an approch by control randomization", *Monte-Carlo methods and applications*, 20(2), 145-165.

[15] Kharroubi I. and H. Pham (2012): "Feynman-Kac representation for Hamilton-Jacobi-Bellman IPDE", to appear in *Annals of Probability*.

[16] Kloeden P. and E. Platen (1992): Numerical Solution of Stochastic Differential Equations, Springer, Series SMAP.

[17] Krylov N. V. (2000): "On the rate of convergence of finite difference approximations for Bellman's equations with variable coefficients", *Probability Theory and Related Fields*, 117, 1-16.

[18] Longstaff F. and E. Schwartz (2001): "Valuing american options by simulation : a simple least-square approach", *Review Of Financial Studies*, 14, 113-147.

[19] Pagès G., Pham H. and J. Printems (2004): Optimal quantization and applications to numerical problems in finance, in *Handbook of Computational and Numerical Methods in Finance*, , ed. S.T. Rachev, Birkhauser, Boston.

[20] Pardoux E. and S. Peng (1990): "Adapted solution of a backward stochastic differential equation", *Systems Control Lett.*, 14(1), 55-61.

[21] Pardoux E. and S. Peng (1992): "Backward stochastic differential equations and quasilinear parabolic partial differential equations", *Stochastic partial differential equations and their applications*, (B. L. Rozovskii and R. B. Sowers, eds.), Lect. Notes in Control and Inform. Sci., vol. 176, Springer, Berlin, pp. 200-217.

[22] Peng S. (2003): Nonlinear expectations and risk measures, in *Proceedings of the CIME-EMS summer school Bressanone*.

[23] Peng S. (2006): "G-expectation, G-Brownian motion and related stochastic calculus of Ito type", Proceedings of 2005, Abel symposium, Springer.

[24] Pham H. (2009): Continuous time stochastic control and optimization with financial applications, Springer, Series SMAP.

[25] Soner M., Touzi N. and J. Zhang (2011): "The wellposedness of second order backward SDEs", *Probability Theory and Related Fields*, 153, 149-190.

[26] Tang S. and X. Li (1994): "Necessary conditions for optimal control of stochastic systems with jumps", *SIAM J. Control and Optimization*, 32, 1447-1475.

[27] Zhang J. (2004): "A numerical scheme for BSDEs", *Annals of Applied Probability*, 14, 459-488.