Numerical Approximation of BSDEs by using Backward Euler schemes

Bruno Bouchard
Université Paris-Dauphine, PSL Research University.
bouchard@ceremade.dauphine.fr

European Summer School in Mathematical Finance
Pushkin, St. Petersburg, 29 August - 2 September 2016

Contents

Part I

Motivation

1 Backward SDEs: definition and existence

For a complete introduction to BSDEs, see the lecture notes [25, 5] and the book [26].

Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$, W a d-dimensional Brownian motion, $\mathbb{F} = (\mathcal{F}_s, 0 \leq s \leq T)$ the filtration generated by W.

Process spaces:
- $\mathbf{S}^2$: adapted continuous processes Y such that $\|Y\|_{\mathbf{S}^2} := \mathbb{E}[\sup_{[0,T]} |Y|^2]^{\frac{1}{2}} < \infty$.
- $\mathbf{L}^2_{\mathcal{P}}$: predictable processes Z such that $\|Z\|_{\mathbf{L}^2_{\mathcal{P}}} := \mathbb{E}[\int_0^T |Z_t|^2 dt]^{\frac{1}{2}} < \infty$.

BSDE: Given ξ in $\mathbf{L}^2$ and $f : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$, find $(Y,Z) \in \mathbf{S}^2 \times \mathbf{L}^2_{\mathcal{P}}$ such that

$$Y_t = \xi + \int_t^T f_s(Y_s, Z_s)ds - \int_t^T Z_s dW_s,\ t \leq T,\ \mathbb{P} - \text{a.s.}$$

It means that the process Y has the dynamics

$$dY_s = -f_s(Y_s, Z_s)ds + Z_s dW_s \ \text{ with } Y_T = \xi.$$

Martingale representation #1: In the case $f \equiv 0$,

$$Y_t = \xi + \int_t^T f_s(Y_s, Z_s)ds - \int_t^T Z_s dW_s,\ t \le T,\ \mathbb{P} - \text{a.s.}$$

can hold only if

$$Y_t = \mathbb{E}\left[\xi \mid \mathcal{F}_t\right],$$

and Z is uniquely given by the martingale representation theorem

$$\xi = \mathbb{E}\left[\xi\right] + \int_0^T Z_s dW_s, \quad \text{i.e. } \mathbb{E}\left[\xi \mid \mathcal{F}_t\right] = \mathbb{E}\left[\xi\right] + \int_0^t Z_s dW_s.$$

$\Rightarrow$ the component Z is here to ensure that the process Y is adapted. Unlike deterministic ODEs, we can not simply revert time as the filtration goes in one direction.

Martingale representation #2: If $f(\omega, y, z) = \rho(\omega) + \alpha(\omega)y + \beta(\omega)z$, then

$$\begin{aligned} Y_t &= \xi + \int_t^T (\rho_s + \alpha_s Y_s + \beta_s Z_s) ds - \int_t^T Z_s dW_s \\ &= \xi + \int_t^T (\rho_s + \alpha_s Y_s) ds - \int_t^T Z_s dW_s^\beta \end{aligned}$$

with

$$W^\beta = W - \int_0^\cdot \beta_s ds \ \text{ a Brownian motion under } \mathbb{Q}^\beta,$$

so that

$$e^{\int_0^t \alpha_s ds} Y_t = e^{\int_0^T \alpha_s ds} \xi + \int_t^T e^{\int_0^s \alpha_r dr} \rho_s ds - \int_t^T e^{\int_0^s \alpha_r dr} Z_s dW_s^\beta$$

and (with $\Lambda = e^{\int_0^\cdot \alpha_s ds} \mathcal{E}(\int_0^\cdot \beta_s dW_s)$)

$$Y_t = \mathbb{E}^{\mathbb{Q}^\beta}[e^{\int_t^T \alpha_s ds} \xi + \int_t^T e^{\int_t^s \alpha_r dr} \rho_s ds \ |\mathcal{F}_t] = \mathbb{E}[\Lambda_T \xi + \int_t^T \Lambda_s \rho_s ds \ |\mathcal{F}_t] \Lambda_t^{-1}.$$

Z is again uniquely given by the martingale representation theorem under $\mathbb{Q}^\beta$.

General existence result: Assume that there exists K such that

$$|f(y,z) - f(y',z')| \le K(|y-y'| + |z-z'|)\; dt \times d\mathbb{P},$$

and that $f(0) \in \mathbf{L}^2_{\mathcal{P}}$, then the BSDE admits a unique solution.

Existence is obtained by a contraction argument (like for SDEs): Given (Y, Z), let (Y', Z') be defined by

$$Y'_t = \xi + \int_t^T f_s(Y_s, Z_s) ds - \int_t^T Z'_s dW_s,$$

and set $(Y', Z') = \Phi(Y, Z)$. The map Φ is contracting on a suitable weighted version of $\mathbf{S}^2 \times \mathbf{L}^2_{\mathcal{P}}$.

Existence can hold under weaker conditions, see [5] for a survey.

2 Examples of application

2.1 Hedging in finance

Stock price: $dS_t = S_t\mu_t dt + S_t\sigma_t dW_t$,

Weath process: Let π be the amount invested in the stock, then the wealth Y evolves according to

$$dY_t = \frac{\pi_t}{S_t} dS_t + r_t(Y_t - \pi_t)dt = \{\pi_t(\mu_t - r_t) + r_t Y_t\}\, dt + \pi_t\sigma_t dW_t.$$

Hedging: In order to hedge the claim ξ at T, we need to find π such that $Y_T = \xi$, i.e.

$$Y_t = \xi - \int_t^T \{Z_s\lambda_s + r_s Y_s\}\, ds - \int_t^T Z_s dW_s,$$

after setting $Z := \pi\sigma$ and $\lambda := (\mu - r)/\sigma$.

Different interest rates for borrowing and lending:

$$
\begin{aligned}
Y_t &= \xi - \int_t^T \left\{ \pi_s \mu_s + r_s^l (Y_s - \pi_s)^+ - r_s^b (Y_s - \pi_s)^- \right\} ds - \int_t^T \pi_s \sigma_s dW_s. \\
&= \xi - \int_t^T \left\{ Z_s \frac{\mu_s}{\sigma_s} + \frac{r_s^l}{\sigma_s} (\sigma_s Y_s - Z_s)^+ - \frac{r_s^b}{\sigma_s} (\sigma_s Y_s - Z_s)^- \right\} ds - \int_t^T Z_s dW_s.
\end{aligned}
$$

It is no more linear...

2.2 Optimal control problem and stochastic maximum principle

Maximization problem:

$$J(\nu) := \mathbb{E}\left[g(X_T^\nu) + \int_0^T f(X_s^\nu, \nu_t)dt\right],$$

in which X^ν is the solution of the one dimensional sde

$$dX_t^\nu = b(X_t^\nu, \nu_t)dt + \sigma(X_t^\nu)dW_t$$

with ν in the set $\mathcal{U}$ of predictable processes with values in $\mathbb{R}$.

Associated Hamiltonian:

$$\mathcal{H}(x,u,p,q) := b(x,u)p + \sigma(x)q + f(x,u).$$

Adjoint BSDE equation

$$\hat{P}_t \;=\; \partial_x g(\hat{X}_T) + \int_t^T \partial_x \hat{\mathcal{H}}(\hat{X}_s, \hat{P}_s, \hat{Q}_s)ds - \int_t^T \hat{Q}_s dW_s \tag{2.1}$$

where

$$\hat{\mathcal{H}}(x,p,q) := \sup_{u\in\mathbb{R}} \mathcal{H}(x,u,p,q).$$

Maximum principle: Assume that

$$x \mapsto g(x) \quad \text{and} \quad x \mapsto \hat{\mathcal{H}}(x, \hat{P}_t, \hat{Q}_t) := \sup_{u \in \mathbb{R}} \mathcal{H}(x, u, \hat{P}_t, \hat{Q}_t) \quad \text{are } \mathbb{P} - \text{a.s. concave.}$$

Assume further that $\hat{\nu}$ satisfies

$$\begin{aligned} \mathcal{H}(\hat{X}_\tau, \hat{\nu}_\tau, \hat{P}_\tau, \hat{Q}_\tau) &= \hat{\mathcal{H}}(\hat{X}_\tau, \hat{P}_\tau, \hat{Q}_\tau) \\ \partial_x \mathcal{H}(\hat{X}_\tau, \hat{\nu}_\tau, \hat{P}_\tau, \hat{Q}_\tau) &= \partial_x \hat{\mathcal{H}}(\hat{X}_\tau, \hat{P}_\tau, \hat{Q}_\tau) \end{aligned}$$

for all stopping times τ, and that $(\hat{X}, \hat{P}, \hat{Q})$ solves the adjoint equation (2.1). Then, an optimal control is given by $\hat{\nu}$.

Remark:

- σ can also depend on the control, but the formulation is more complex.
- See [5] for exponential utility maximization problems that can be treated differently and lead to quadratic BSDEs.

2.3 Risk measures representation

See Peng [38] for a complete treatment.

Definition A non-linear $\mathbb{F}$-expectation is an operator $\mathcal{E} : \mathbf{L}^2 \mapsto \mathbb{R}$ such that

- $X' \geq X$ implies $\mathcal{E}[X'] \geq \mathcal{E}[X']$ with equality if and only if $X' = X$.
- $\mathcal{E}[c] = c$ for $c \in \mathbb{R}$.
- For each $X \in \mathbf{L}^2$ and $t \leq T$, there exists $\eta_t^X \in \mathbf{L}^2(\mathcal{F}_t)$ such that $\mathcal{E}[X\mathbf{1}_A] = \mathcal{E}[\eta_t^X \mathbf{1}_A]$ for all $A \in \mathcal{F}_t$. We write $\mathcal{E}_t[X]$ for η_t^X.

Remark: η_t^X is uniquely defined. It corresponds to the notion of conditional expectation, in this non-linear framework.

Let us now consider the solution (Y, Z) of

$$Y_t = \xi + \int_t^T f_s(Y_s, Z_s)ds - \int_t^T Z_s dW_s,\ t \leq T,$$

and call the Y component $\mathcal{E}_t^f[\xi]$.

Theorem If f satisfies the Lipschitz and integrability conditions given in the above existence result, then $\mathcal{E}^f$ is a non-linear $\mathbb{F}$-expectation. Conversely, let $\mathcal{E}$ be a non-linear $\mathbb{F}$-expectation such that for all $X, X' \in \mathbf{L}^2$

$$\mathcal{E}[X + X'] \leq \mathcal{E}[X] + \mathcal{E}^{f_\mu}[X'] \qquad (\text{with } f_\mu(y, z) = \mu|z|)$$

and

$$\mathcal{E}_t[X + X'] = \mathcal{E}_t[X] + X' \text{ if } X' \in \mathbf{L}^2(\mathcal{F}_t).$$

Then, there exists a random driver f (which does not depend on y) such that $\mathcal{E} = \mathcal{E}^f$.

2.4 Semi-linear PDEs

Consider the solution X of the sde

$$X = x + \int_0^\cdot b_s(X_s)ds + \int_0^\cdot \sigma_s(X_s)dW_s,$$

in which b and σ are determinist, Lipschitz in space, and continuous in time. Assume that v is a smooth solution of

$$0 = \mathcal{L}v + f(\cdot, v, \partial_x v\sigma) \text{ on } [0,T) \times \mathbb{R}, \text{ with } v(T,\cdot) = g$$

for some continuous g with polynomial growth, and where

$$\mathcal{L}v = \partial_t v + b\partial_x v + \frac{1}{2}\sigma^2\partial^2_{xx}v.$$

Then,

$$Y := v(\cdot, X) \text{ , } Z := \partial_x v(\cdot, X)\sigma(X)$$

solves

$$Y = g(X_T) + \int_\cdot^T f_s(X_s, Y_s, Z_s)ds - \int_\cdot^T Z_s dW_s.$$

Indeed, by Itô's Lemma,

$$
\begin{aligned}
g(X_T) &= v(t, X_t) + \int_t^T \mathcal{L}v(s, X_s)ds + \int_t^T \partial_x v(s, X_s)\sigma_s(X_s)dW_s \\
&= \underbrace{v(t, X_t)}_{Y_t} - \int_t^T f_s(X_s, \underbrace{v(s, X_s)}_{Y_s}, \underbrace{\partial_x v(s, X_s)\sigma_s(X_s)}_{Z_s})ds \\
&\quad + \int_t^T \underbrace{\partial_x v(s, X_s)\sigma_s(X_s)}_{Z_s} dW_s.
\end{aligned}
$$

In general, no smooth solution exists but $Y = v(\cdot, X)$ where v is the unique viscosity solution of the corresponding PDE (see Crandall, Ishii & Lions [21] for the definition of viscosity solutions).

In particular, solving the BSDE or the PDE is equivalent.

3 Extentions (in the Markovian case from now on)

3.1 BSDEs with jumps and IPDEs

Consider the solution $(X, Y, Z, U) \in \mathbf{S}^2 \times \mathbf{S}^2 \times \mathbf{L}^2_{\mathcal{P}} \times \mathbf{L}^2_\lambda$ of

$$\begin{cases} X_t &= X_0 + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dW_r + \int_0^t \int_E \beta(X_{r-}, e)\overline{\mu}(de, dr) \;, \\ Y_t &= g(X_T) + \int_t^T f\left(\Theta_r\right) dr - \int_t^T Z_r dW_r - \int_t^T \int_E U_r(e)\overline{\mu}(de, dr) \end{cases}$$

where $\Theta := (X, Y, \Gamma, Z)$ with $\Gamma := \int_E \rho(e)U(e)\lambda(de)$, and

- b, σ, β, f, g Lipschitz in (x, y, γ, z) uniformly in $e \in E$.
- μ Poisson measure on $E = \mathbb{R}^\ell$ with compensator $\nu(de, dt) = \lambda(de)dt$ and $\overline{\mu} = \mu - \nu$. $\rho : E \mapsto \mathbb{R}^m$ is bounded.
- $\mathbf{L}^2_\lambda$: $\mathcal{P} \otimes \mathcal{B}(E)$ measurable maps $U : \Omega \times [0, T] \times E \to \mathbb{R}$ such that

$$\|U\|_{\mathbf{L}^2_\lambda} := \mathbb{E}\left[\int_0^T \int_E |U_s(e)|^2 \lambda(de) ds\right]^{\frac{1}{2}} < \infty$$

If v solves

$$\begin{aligned} -\mathcal{L}v(t,x) - f\left(t, x, v(t,x), \sigma(t,x)\partial_x v(t,x), \mathcal{I}[v](t,x)\right) &= 0 \\ v(T,x) &= g(x)\,, \end{aligned}$$

where

$$\begin{aligned} \mathcal{L}v(t,x) &:= \partial_t v(t,x) + \partial_x v(t,x) b(x) + \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^\top(x))^{ij} \frac{\partial^2 u}{\partial x^i \partial x^j}(t,x) \\ &\quad + \int_E \left\{ v(t, x+\beta(x,e)) - v(t,x) - \partial_x v(t,x)\beta(x,e) \right\} \lambda(de)\ , \\ \mathcal{I}[v](t,x) &:= \int_E \left\{ v(t, x+\beta(x,e)) - v(t,x) \right\} \rho(e)\, \lambda(de)\ , \end{aligned}$$

then

$$Y = v(\cdot, x),\ Z = \partial_x v(\cdot, X)\sigma(\cdot, X) \text{ and } U = v(\cdot, X_- + \beta(X_-, \cdot)) - v(\cdot, X_-).$$

3.2 BSDEs with jumps and systems of PDEs

Idea coming from Pardoux, Pradeilles & Rao [37].

System of κ PDE's ($i = 0, \ldots, \kappa - 1$)

$$\begin{aligned} 0 &= \partial_t v_i + b_i \partial_x v_i + \frac{1}{2}\mathrm{Tr}[\sigma_i \sigma_i^\top \partial^2_{xx} v_i] + f_i(\cdot, v, \partial_x v_i \sigma_i) \\ g_i &= v_i(T, \cdot)\ . \end{aligned}$$

Define for $i = 0, \ldots, \kappa - 1$

$$\tilde{f}(i,t,x,y,\gamma,z) \;=\; f_i\left(t, x, (\ldots, y+\gamma^{\kappa-2}, y+\gamma^{\kappa-1}, \underbrace{y}_{i}, y+\gamma^1, y+\gamma^2, \ldots), z\right)$$

Set $E = \{1, \ldots, \kappa - 1\}$, $\lambda(de) = \lambda \sum_{k=1}^{\kappa-1} \delta_k(e)$ and

$$M_t \;=\; \int_0^t \int_E e\mu(de, ds) \quad \text{modulo } \kappa.$$

Then, the corresponding BSDE is:

$$\begin{aligned} dX_t &= b_{M_t}(X_t)dt + \sigma_{M_t}(X_t)dW_t \\ -dY_t &= \tilde{f}(M_t, t, X_t, Y_t, \Gamma_t, Z_t)dt - \lambda \sum_{k=1}^{\kappa-1} \Gamma_t^k dt - Z_t dW_t - \int_E U_t(e)\overline{\mu}(de, dt) \\ Y_T &= g_{M_T}(X_T) \end{aligned}$$

with $\rho(k)^i = \lambda^{-1}\mathbf{1}_{k=i}$, i.e $\Gamma^k = U(k)$.

Indeed, it corresponds to

$$0 = \partial_t v(\cdot, i) + b_i \partial_x v(\cdot, i) + \frac{1}{2}\mathrm{Tr}[\sigma_i \sigma_i^\top \partial_{xx}^2 v(\cdot, i)] + \tilde{f}(\cdot, v(\cdot, i), \partial_x v(\cdot, i)\sigma_i, \mathcal{I}[v(\cdot, i)])$$

where

$$\begin{aligned} &\tilde{f}(\cdot, v, \cdot, \mathcal{I}[v(\cdot, i)]) = \\ &f_i(\cdot, (\ldots, \underbrace{v(\cdot, i) + v(\cdot, i-1) - v(\cdot, i)}_{v(\cdot, i-1)}, \underbrace{v(\cdot, i)}_{i}, \underbrace{v(\cdot, i) + v(\cdot, i+1) - v(\cdot, i)}_{v(\cdot, i+1)}, \ldots), \cdot) \end{aligned}$$

Remark: We obtain such systems in the case of option pricing with possible default.

3.3 BSDEs in a domain

Consider the system:

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \\ Y_t &= g(\tau, X_\tau) + \int_{t\wedge\tau}^{\tau} f_s(X_s, Y_s, Z_s)ds - \int_{t\wedge\tau}^{\tau} Z_s dW_s, \end{aligned}$$

where

$$\tau := \inf\{t \geq 0 : X_t \notin \mathcal{O}\} \wedge T,$$

for some open set $\mathcal{O}$.

Then, it corresponds to the Dirichlet problem:

$$\begin{aligned} -\mathcal{L}v - f(\cdot, v, \partial_x v\sigma) = 0 &\quad \text{on } [0,T) \times \mathcal{O} \\ v = g &\quad \text{on } ([0,T) \times \partial\mathcal{O}) \cup (\{T\} \times \overline{\mathcal{O}}). \end{aligned}$$

Remark: Barrier options in finance.

3.4 Reflected BSDEs

Consider the system:

$$
\begin{aligned}
X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \\
Y_t &= g(X_T) + \int_t^T f_s(X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s + K_T - K_t \\
Y_t &\geq g(X_t),
\end{aligned}
$$

where K is non-decreasing adapted continuous and

$$
\int_0^T [Y_t - g(X_t)]dK_t = 0.
$$

Then, it corresponds to the obstacle problem:

$$
\begin{aligned}
\min\{-\mathcal{L}v - f(\cdot, v, \partial_x v\sigma), v - g\} = 0 &\quad \text{on } [0,T) \times \mathbb{R}^d \\
v = g &\quad \text{on } \{T\} \times \mathbb{R}^d.
\end{aligned}
$$

Remark: American options in finance.

Part II

Discrete time approximation

See [8] for a survey.

1 A backward Euler type scheme

From now on, we set $T = 1$ and consider

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s , \\ Y_t &= g(X_1) + \int_t^1 f(X_s, Y_s, Z_s)ds - \int_t^1 Z_s dW_s \end{aligned}$$

with b, σ, f, g Lipschitz.

We construct a discrete time approximation on a grid $\pi := \{t_i := i/n,\ i \le n\}$ of $[0, T]$ (with $t_0 = 0$ and $t_n = T$). The mesh is $|\pi| = n^{-1}$.

1.1 Construction of the scheme

It is based on the simple approximation

$$\begin{aligned} Y_{t_i} &= Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f\left(X_s, Y_s, Z_s\right) ds - \int_{t_i}^{t_{i+1}} Z_t dW_t \\ &\simeq Y_{t_{i+1}} + \frac{1}{n} f\left(X_{t_i}, Y_{t_i}, \overline{Z}_{t_i}\right) - \int_{t_i}^{t_{i+1}} Z_t dW_t \end{aligned}$$

where

$$\overline{Z}_{t_i} := n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} Z_t dt] \simeq n\mathbb{E}_{t_i}[Y_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})],$$

so that

$$\begin{aligned} Y_{t_i} &\simeq \mathbb{E}_{t_i}[Y_{t_{i+1}}] + \frac{1}{n} f\left(X_{t_i}, Y_{t_i}, \overline{Z}_{t_i}\right) \\ \overline{Z}_{t_i} &\simeq n\mathbb{E}_{t_i}[Y_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})] \end{aligned}$$

The discrete time approximation is based on forcing equality in the above.

We define $(\overline{Y}^\pi, \overline{Z}^\pi)$ by

$$\begin{aligned}\overline{Y}^\pi_{t_i} &:= \mathbb{E}_{t_i}[\overline{Y}^\pi_{t_{i+1}}] + \frac{1}{n} f(X^\pi_{t_i}, \overline{Y}^\pi_{t_i}, \overline{Z}^\pi_{t_i})\\ \overline{Z}^\pi_{t_i} &:= n\mathbb{E}_{t_i}[\overline{Y}^\pi_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})],\end{aligned}$$

where $\overline{Y}^\pi_{t_n} = g(X^\pi_{t_n})$, and X^π is the Euler scheme of X.

We want to control

$$\mathrm{Err}^2 := \max_i \mathbb{E}\left[\sup_{t_i \le t \le t_{i+1}} |Y_t - \overline{Y}^\pi_{t_i}|^2\right] + \int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}^\pi_t|^2\right] dt$$

in which $\overline{Z}^\pi_t := \overline{Z}^\pi_{t_i}$ on $[t_i, t_{i+1})$.

Remark: One needs to solve the first equation in $\overline{Y}^\pi_{t_i}$. One could alternatively set

$$\overline{Y}^\pi_{t_i} = \mathbb{E}_{t_i}[\overline{Y}^\pi_{t_{i+1}}] + \frac{1}{n}\mathbb{E}_{t_i}[f(X^\pi_{t_i}, \overline{Y}^\pi_{t_{i+1}}, \overline{Z}^\pi_{t_i})],$$

without changing the nature of the convergence.

1.2 Important quantities and first error bound

By construction

$$\overline{Z}_{t_i} := n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} Z_s ds]$$

is the best $\mathbf{L}^2$-approximation of Z by a step-constant process, and thus

$$\int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}^{\pi}_t|^2\right] dt \geq \mathcal{R}(Z) := \int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt.$$

On the other hand, if $f \equiv 0$, then $Y_{t_i} = \mathbb{E}_{t_i}[Y_t]$ so that $(Y_{t_i})_i$ also provides the best $\mathbf{L}^2$-approximation and we expect that

$$\max_i \mathbb{E}\left[\sup_{t_i \leq t \leq t_{i+1}} |Y_t - \overline{Y}^{\pi}_{t_i}|^2\right] \gtrsim \mathcal{R}(Y) := \max_i \mathbb{E}\left[\sup_{t_i \leq t \leq t_{i+1}} |Y_t - Y_{t_i}|^2\right].$$

Not surprisingly the error should depend on the regularity of (Y, Z)...

Theorem: There exists $C > 0$ such that

$$\mathrm{Err}^2 \leq C\left(|\pi| + \mathcal{R}(Y) + \mathcal{R}(Z)\right)$$

Proof. (case $f(X,Y,Z) = f(Z)$)

• Continuous backward Euler scheme on $[t_i, t_{i+1}]$

$$Y_t^\pi \;=\; \overline{Y}_{t_{i+1}}^\pi + (t_{i+1} - t) f\left(\overline{Z}_{t_i}^\pi\right) - \int_t^{t_{i+1}} Z_s^\pi dW_s \,.$$

• Set $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$. By Itô's Lemma,

$$\begin{aligned}
\mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;&=\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] \\
&\quad + \int_t^{t_{i+1}} \mathbb{E}\left[\delta Y_s (f(Z_s) - f(\overline{Z}_{t_i}^\pi))\right] ds
\end{aligned}$$

Proof. (case $f(X,Y,Z) = f(Z)$)

• Continuous backward Euler scheme on $[t_i, t_{i+1}]$

$$Y_t^\pi \;=\; \overline{Y}^\pi_{t_{i+1}} + (t_{i+1} - t) f\left(\overline{Z}^\pi_{t_i}\right) - \int_t^{t_{i+1}} Z_s^\pi dW_s \,.$$

• Set $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$. By Itô's Lemma,

$$\begin{aligned}\mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;&=\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] \\ &+\; \int_t^{t_{i+1}} \mathbb{E}\left[\delta Y_s (f(Z_s) - f(\overline{Z}^\pi_{t_i}))\right] ds\end{aligned}$$

• Then, using $|\delta y(f(z) - f(z'))| \le \frac{C}{\alpha}|\delta y|^2 + \frac{4\alpha}{C}|f(z) - f(z')|^2$,

$$\begin{aligned}\mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;&\le\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] + \frac{C}{\alpha}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Y_s|^2\right] ds \\ &+\; \alpha \int_t^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}_{t_i}|^2 + |\overline{Z}_{t_i} - \overline{Z}^\pi_{t_i}|^2\right] ds\end{aligned}$$

Since

$$\begin{aligned}
\overline{Z}_{t_i} &:= n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} Z_s ds] \\
\overline{Z}^{\pi}_{t_i} &:= n\mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})] = n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} Z^{\pi}_s ds],
\end{aligned}$$

we deduce from Jensen's inequality that

$$\begin{aligned}
\int_t^{t_{i+1}} \mathbb{E}\left[|\overline{Z}_{t_i} - \overline{Z}^{\pi}_{t_i}|^2\right] ds &\leq \int_t^{t_{i+1}} n \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_r - Z^{\pi}_r|^2\right] dr ds \\
&\leq \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_s - Z^{\pi}_s|^2\right] ds \\
&= \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds
\end{aligned}$$

Proof. (case $f(X,Y,Z) = f(Z)$)

• Continuous backward Euler scheme on $[t_i, t_{i+1}]$

$$Y_t^\pi \;=\; \overline{Y}^\pi_{t_{i+1}} + (t_{i+1}-t) f\left(\overline{Z}^\pi_{t_i}\right) - \int_t^{t_{i+1}} Z_s^\pi dW_s \,.$$

• Set $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$. By Itô's Lemma,

$$\begin{aligned} \mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;&=\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] \\ &+\; \int_t^{t_{i+1}} \mathbb{E}\left[\delta Y_s (f(Z_s) - f(\overline{Z}^\pi_{t_i}))\right] ds \end{aligned}$$

• Then, using $|\delta y(f(z) - f(z'))| \le \frac{C}{\alpha}|\delta y|^2 + \frac{4\alpha}{C}|f(z)-f(z')|^2$,

$$\begin{aligned} \mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;&\le\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] + \frac{C}{\alpha}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Y_s|^2\right] ds \\ &+\; \alpha \int_t^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}_{t_i}|^2 + \underbrace{|\overline{Z}_{t_i} - \overline{Z}^\pi_{t_i}|^2}_{\text{“}\le |\delta Z_s|^2\text{”}} \right] ds \end{aligned}$$

Proof. (case $f(X,Y,Z) = f(Z)$)

• Continuous backward Euler scheme on $[t_i, t_{i+1}]$

$$Y_t^\pi \;=\; \overline{Y}^\pi_{t_{i+1}} + (t_{i+1}-t) f\left(\overline{Z}^\pi_{t_i}\right) - \int_t^{t_{i+1}} Z_s^\pi dW_s\,.$$

• Set $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$. By Itô's Lemma,

$$\begin{aligned}\mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;&=\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] \\ &\quad + \int_t^{t_{i+1}} \mathbb{E}\left[\delta Y_s (f(Z_s) - f(\overline{Z}^\pi_{t_i}))\right] ds\end{aligned}$$

• For $\eta := \frac{1}{2} - \alpha > 0$, and with the help of Gronwall's lemma,

$$\mathbb{E}\left[|\delta Y_{t_i}|^2\right] + \eta \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;\le\; e^{C|\pi|}\mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] + \alpha \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}_{t_i}|^2\right] ds$$

Proof. (case $f(X,Y,Z) = f(Z)$)

• Continuous backward Euler scheme on $[t_i, t_{i+1}]$

$$Y_t^\pi \;=\; \overline{Y}^\pi_{t_{i+1}} + (t_{i+1} - t) f\left(\overline{Z}^\pi_{t_i}\right) - \int_t^{t_{i+1}} Z_s^\pi dW_s \,.$$

• Set $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$. By Itô's Lemma,

$$\mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2}\int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;=\; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] + \int_t^{t_{i+1}} \mathbb{E}\left[\delta Y_s (f(Z_s) - f(\overline{Z}^\pi_{t_i}))\right] ds$$

• By the discrete Gronwall's lemma,

$$\max_i \mathbb{E}\left[|\delta Y_{t_i}|^2\right] + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \;\le\; C \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}_{t_i}|^2\right] ds$$

Proof. (case $f(X, Y, Z) = f(Z)$)

• Continuous backward Euler scheme on $[t_i, t_{i+1}]$

$$Y_t^\pi \; = \; \overline{Y}^\pi_{t_{i+1}} + (t_{i+1} - t) f\left(\overline{Z}^\pi_{t_i}\right) - \int_t^{t_{i+1}} Z_s^\pi dW_s \, .$$

• Set $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$. By Itô's Lemma,

$$\begin{aligned} \mathbb{E}\left[|\delta Y_t|^2\right] + \frac{1}{2} \int_t^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds \; &= \; \mathbb{E}\left[|\delta Y_{t_{i+1}}|^2\right] \\ &+ \; \int_t^{t_{i+1}} \mathbb{E}\left[\delta Y_s (f(Z_s) - f(\overline{Z}^\pi_{t_i}))\right] ds \end{aligned}$$

• Finally, as above,

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}^\pi_{t_i}|^2\right] ds \; &\le \; \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\overline{Z}_{t_i} - \overline{Z}^\pi_{t_i}|^2\right] ds + \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}_{t_i}|^2\right] ds \\ &\le \; \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|\delta Z_s|^2\right] ds + \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_s - \overline{Z}_{t_i}|^2\right] ds. \end{aligned}$$

□

We have "shown" that

Theorem: There exists $C > 0$ such that

$$\text{Err}^2 \leq C\left(|\pi| + \mathcal{R}(Y) + \mathcal{R}(Z)\right)$$

in which

$$\mathcal{R}(Y) := \max_i \mathbb{E}\left[\sup_{t_i \leq t \leq t_{i+1}} |Y_t - Y_{t_i}|^2\right]$$

$$\mathcal{R}(Z) := \int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt \;\; \text{with } \overline{Z}_{t_i} := n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} Z_s ds].$$

It remains to study the modulus of regularity $\mathcal{R}(Y)$ and $\mathcal{R}(Z)$...

1.3 Modulus of regularity of Y

This is the easy part.... Standard estimates (Gronwall + BDG inequalities) lead to $Y_t = v(t, X_t)$ in which v is 1/2-Hölder in time and Lipschitz in X. Then,

$$|Y_t - Y_s|^2 \leq C\left(|t-s| + |X_t - X_s|^2\right).$$

Hence,

$$\mathcal{R}(Y) = \max_i \mathbb{E}\left[\sup_{t_i \leq t \leq t_{i+1}} |Y_t - Y_{t_i}|^2\right] \leq \max_i C(t_{i+1} - t_i) = C|\pi|.$$

Theorem: There exists $C > 0$ such that

$$\mathrm{Err}^2 \leq C\left(|\pi| + \mathcal{R}(Y) + \mathcal{R}(Z)\right) \leq C\left(|\pi| + \mathcal{R}(Z)\right).$$

1.4 Modulus of regularity of Z

This is the difficult part... We use the representation of Z in terms of Malliavin derivatives. The initial proof is due to Ma & Zhang [35].

• Assume that the coefficients are smooth enough. Then, (Y_t, Z_t) admits a Malliavin derivative for all $t \le T$ and (D_sY, D_sZ) solves

$$D_sY_t \;=\; \nabla g(X_T)D_sX_T + \int_t^T \nabla f(\underbrace{\Theta_r}_{(X_r,Y_r,Z_r)})D_s\Theta_r dr - \int_t^T D_sZ_rdW_r$$

• Since

$$Y_t = Y_0 - \int_0^t f(\Theta_r)dr + \int_0^t Z_rdW_r$$

we have (Λ as on slide 10)

$$\begin{aligned} Z_t = D_tY_t \;&=\; \mathbb{E}\left[\Lambda_T\nabla g(X_T)D_sX_T + \int_t^T \Lambda_r\nabla_x f(\Theta_r)D_sX_rdr \mid \mathcal{F}_t\right]\Lambda_t^{-1} \\ &=\; \mathbb{E}\left[\Lambda_T\nabla g(X_T)\nabla X_T + \int_t^T \Lambda_r\nabla_x f(\Theta_r)\nabla X_rdr \mid \mathcal{F}_t\right](\nabla X_t)^{-1}\sigma(X_t)\Lambda_t^{-1} \end{aligned}$$

Another way to put this: $Y = v(\cdot, X)$ and $Z = \partial_x v(\cdot, X)\sigma(X)$ with v solving

$$-\mathcal{L}v - f(\cdot, v, \partial_x v\sigma) = 0.$$

Then, $u := \partial_x v$ solves a semi-linear equation of the form

$$-\tilde{\mathcal{L}}u - \tilde{f}(\cdot, u, \partial_x u\sigma) = 0.$$

By looking at this equation, we obtain that $(u(\cdot, X)\nabla X, \partial_x u(\cdot, X)\sigma(X)) = (Z\sigma^{-1}(X)\nabla X, \partial_x u(\cdot, X)\sigma(X)))$ solves a linear BSDE whose solution is given by the above (for an appropriate process Λ).

- Then, $Z_t = (V_t - \alpha_t)\eta_t$ where

$$\begin{aligned} V_t &= \mathbb{E}\left[\Lambda_T \nabla g(X_T)\nabla X_T + \int_0^T \Lambda_r \nabla_x f(\Theta_r)\nabla X_r dr \mid \mathcal{F}_t\right] \\ \eta_t &= (\nabla X_t)^{-1}\sigma(X_t)\Lambda_t^{-1} \\ \alpha_t &= \int_0^t \Lambda_r \nabla_x f(\Theta_r)\nabla X_r dr \end{aligned}$$

- There exists $\tilde{\sigma} \in \mathbf{L}^2_{\mathcal{P}}$ such that

$$V_t = V_0 + \int_0^t \tilde{\sigma}_s dW_s,$$

and thus

$$\begin{aligned} \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|V_t - V_{t_i}|^2\right] dt &\le C \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \mathbb{E}\left[|\tilde{\sigma}_s|^2\right] ds dt \\ &\le C\,|\pi|\,. \end{aligned}$$

The additional terms behave in $|\pi|^{\frac{1}{2}}$ in any $\mathbf{S}^p$...

Hence,

$$\mathcal{R}(Z) := \int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt \leq C|\pi|.$$

Theorem: There exists $C > 0$ such that

$$\mathrm{Err}^2 \leq C\left(|\pi| + \mathcal{R}(Y) + \mathcal{R}(Z)\right) \leq C|\pi|.$$

1.5 Strong convergence speed (conclusion)

Theorem:(B. & Touzi [11], and Zhang[39])) There exists $C > 0$ such that

$$\mathrm{Err} \leq C|\pi|^{\frac{1}{2}}.$$

This is the convergence speed in the linear case, it can not be improved.

1.6 Weak convergence speed

Under additional regularity assumptions, Gobet and Labart [29] obtain expansions of the form

$$\begin{aligned} Y_t - Y_t^\pi &= \partial_x v(t, X_t)(X_t - X_t^\pi) + O(|\pi|) + O\left(|X_t - X_t^\pi|^2\right) \\ Z_t - \overline{Z}_t^\pi &= \partial_x \left[\partial_x v(t, X_t)\sigma\right]^\top (X_t - X_t^\pi) + O(|\pi|) + O\left(|X_t - X_t^\pi|^2\right) \ , \end{aligned}$$

where $(X_t^\pi, Y_t^\pi)_{t \le 1}$ is the continuous version of the Euler scheme.

In particular, this implies that

$$Y_0 - \overline{Y}_0^\pi = Y_0 - Y_0^\pi = O(|\pi|).$$

Moreover, when the grid is uniform, this allows to deduce from the weak convergence of the process $\sqrt{n}(X - X^\pi)$ that the process $\sqrt{n}(X - X^\pi, Y - Y^\pi, Z - \overline{Z}^\pi)$ weakly converges too.

2 BSDEs with jumps

In the presence of jumps, the discrete time approximation is essentially the same.

We consider the BSDE

$$
\begin{aligned}
X_t &= X_0 + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dW_r + \int_0^t \int_E \beta(X_{r-}, e)\overline{\mu}(de, dr) \ , \\
Y_t &= g(X_1) + \int_t^1 f\left(\Theta_r\right)dr - \int_t^1 Z_r dW_r - \int_t^1 \int_E U_r(e)\overline{\mu}(de, dr)
\end{aligned}
$$

where $\Theta := (X, Y, \Gamma, Z)$ with $\Gamma := \int_E \rho(e)U(e)\lambda(de)$.

2.1 Approximation scheme

Step-constant driver

$$\overline{Y}^{\pi}_{t_i} = \overline{Y}^{\pi}_{t_{i+1}} + \frac{1}{n} f\left(X^{\pi}_{t_i}, \overline{Y}^{\pi}_{t_i}, \overline{\Gamma}^{\pi}_{t_i}, \overline{Z}^{\pi}_{t_i}\right) - \int_{t_i}^{t_{i+1}} Z^{\pi}_t dW_t - \int_{t_i}^{t_{i+1}} \int_E U^{\pi}_t \overline{\mu}(de, dt),$$

Best $\mathbf{L}^2_{\mathcal{P}}$ approximation of Z^{π} and $\Gamma^{\pi} = \int_E U^{\pi}(e)\rho(e)\lambda(de)$ by $\mathcal{F}_{t_i}$-meas. processes

$$\overline{Z}^{\pi}_t := n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} Z^{\pi}_s ds] = n\mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})]$$

$$\overline{\Gamma}^{\pi}_t := n\mathbb{E}_{t_i}[\int_{t_i}^{t_{i+1}} \Gamma^{\pi}_s ds] = n\mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}} \int_E \rho(e)\overline{\mu}(de, (t_i, t_{i+1}])].$$

Discrete scheme: $\overline{Y}^{\pi}_{t_i} = \mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}] + \frac{1}{n} f\left(X^{\pi}_{t_i}, \overline{Y}^{\pi}_{t_i}, \overline{\Gamma}^{\pi}_{t_i}, \overline{Z}^{\pi}_{t_i}\right)$.

We want to control

$$\text{Err}^2 := \max_i \mathbb{E}\left[\sup_{t_i \le t \le t_{i+1}} |Y_t - \overline{Y}^{\pi}_{t_i}|^2\right] + \int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}^{\pi}_t|^2\right] dt + \int_0^1 \mathbb{E}\left[|\Gamma_t - \overline{\Gamma}^{\pi}_t|^2\right] dt.$$

2.2 Error analysis

We find the same initial error bound in terms of the modulus of regularity of Y and Z, + an additional term related to Γ.

But $Y = v(\cdot, X)$ and $U(e) = v(\cdot, X_- + \beta(X_-, e)) - v(\cdot, X_-)$ so that U (and therefore Γ) can be analyzed as Y.

The analysis of Z can be done by using the same martingale arguments as above but it requires and additional invertibility condition on the flow X.

Assumption : For each $e \in E$, the map $x \in \mathbb{R}^d \mapsto \beta(x, e)$ admits a Jacobian matrix $\nabla\beta(x, e)$ such that the function

$$(x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto a(x, \xi; e) := \xi^\top (\nabla\beta(x, e) + I_d)\xi$$

satisfies one of the following condition uniformly in $(x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d$

$$a(x, \xi; e) \ \geq \ |\xi|^2 K^{-1} \quad \text{or} \quad a(x, \xi; e) \ \leq \ -|\xi|^2 K^{-1} \, .$$

Proposition:(B. & Elie [7]) Under the above condition, there exists $C > 0$ such that

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbb{E}\left[|Z_t - Z_{t_i}|^2\right] dt \;\leq\; C\, |\pi|.$$

Theorem:(B. & Elie [7]) Under the above condition, there exists $C > 0$ such that

$$\mathrm{Err} \;\leq\; C\, |\pi|^{\frac{1}{2}}\,.$$

3 Reflected BSDEs

3.1 Approximation scheme

We want to approximate the solution (Y, Z, K) of

$$\begin{aligned} Y_t &= g(X_1) + \int_t^1 f(X_s, Y_s, Z_s)ds - \int_t^1 Z_s dW_s + K_1 - K_t \ , \\ Y_t &\geq h(X_t) \ , \quad 0 \leq t \leq 1 \end{aligned}$$

with K continuous, increasing, such that $K_0 = 0$ and

$$\int_0^1 (Y_t - h(X_t))dK_t = 0 \ .$$

References:

- $f = 0$: Clément, Lamberton and Protter [20]

- f independent of Z : Bally, Pages and Printemps ([1],...), B. and Touzi [11].

- f depends on Z : Ma and Zhang [36].

Approximation scheme

$$
\begin{aligned}
\overline{Z}^{\pi}_{t_i} &= n\, \mathbb{E}_i \left[\overline{Y}^{\pi}_{t_{i+1}} (W_{t_{i+1}} - W_{t_i}) \right] \\
\tilde{Y}^{\pi}_{t_i} &= \mathbb{E}_i \left[\overline{Y}^{\pi}_{t_{i+1}} \right] + n^{-1}\, f(X^{\pi}_{t_i}, \tilde{Y}^{\pi}_{t_i}, \overline{Z}^{\pi}_{t_i}) \\
\overline{Y}^{\pi}_{t_i} &= \mathcal{J}\left(t_i \,,\; X^{\pi}_{t_i} \,,\; \tilde{Y}^{\pi}_{t_i} \right) \;,
\end{aligned}
$$

with the terminal condition

$$
\overline{Y}^{\pi}_{t_n} = g(X^{\pi}_1) \;.
$$

where for $\Re = \{r_j,\ 0 \le j \le \kappa\} \supset \pi$

$$
\mathcal{J}(t,x,y) := y + [h(x) - y]^{+} \mathbf{1}_{\{t \in \Re \setminus \{0,T\}\}} \;,\quad (t,x,y) \in [0,T] \times \mathbb{R}^{d+1} \;.
$$

3.2 Regularity of Z : Ma and Zhang approach

Representation of the Z: The same ideas as above (but more delicate) combined with an integration by parts argument lead to

$$Z_t = \mathbb{E}_t[g(X_T)N_T^t + \int_t^T f(\Theta_s)N_s^t ds + \int_t^T N_s^t dK_s]$$

where

$$N_r^t \ := \ (r-t)^{-1} \int_t^r \sigma(X_s)^{-1} \nabla X_s dW_s \ (\nabla X_t)^{-1} \sigma(X_t).$$

Theorem: (Ma and Zhang [36]) If σ is uniformly elliptic, σ, $b \in C_b^1$ and $h \in C_b^2$, then $\mathcal{R}(Z) \leq |\pi|^{\frac{1}{2}}$. If $\Re = \pi$, then

$$\mathbb{E}\left[\max_{i\leq n}|Y_{t_i} - \overline{Y}^{\pi}_{t_i}|^2\right] + \max_{i\leq n}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}|Y_t - \overline{Y}^{\pi}_{t_{i+1}}|^2\right] + \|Z^\pi - Z\|^2_{\mathbf{L}^2_{\mathcal{P}}} \ \leq \ C\ |\pi|^{\frac{1}{2}}\ .$$

Remark : Requires stronger conditions and converges only in $|\pi|^{\frac{1}{4}}$ (instead of $|\pi|^{\frac{1}{2}}$).

3.3 Regularity of Z : Discretely reflected case

To try to improve the above, one first considers the **discretely reflected** BSDE

$$\begin{aligned} \tilde{Y}^{\Re}_t &= Y^{\Re}_{r_{j+1}} + \int_t^{r_{j+1}} f({\Theta^{\Re}}_s)ds - \int_t^{r_{j+1}} Z^{\Re}_s dW_s \ , \\ Y^{\Re}_t &= \mathcal{J}\left(t \ , \ X_t \ , \ \tilde{Y}^{\Re}_t\right) \qquad \text{on each } [r_j, r_{j+1}] \ , \ j \le \kappa - 1 \ . \end{aligned}$$

with $Y^{\Re}_1 = g(X_1)$ and $\Theta^{\Re} = (X, \tilde{Y}^{\Re}, Z^{\Re})$.

Proposition: Let τ_j be the next reflection time after r_j. There is a version of $Z^{\Re}$ such that for each $j \le \kappa - 1$ and $t \in [r_j, r_{j+1})$:

$$\begin{aligned} Z^{\Re}_t &= \mathbb{E}_t\left[\nabla g(X_T)\Lambda_T \nabla X_T \mathbf{1}_{\{\tau_j = T\}} + \nabla h(X_{\tau_j})(\Lambda \nabla X)_{\tau_j} \mathbf{1}_{\{\tau_j < T\}}\right]\Lambda_t^{-1} \\ &+ \mathbb{E}_t\left[\int_t^{\tau_j} \nabla_x f(\Theta^{\Re}_s)(\Lambda \nabla X)_s ds\right]\Lambda_t^{-1}(\nabla X_t)^{-1}\sigma(X_t) \ . \end{aligned}$$

This works as in the non-reflected case... because the reflection is performed only at fixed discrete times (do the same as above on each time interval).

Main idea to conclude: Assume $h = g$ is C^1_L and $f \equiv 0$. For simplicity: $\Lambda \equiv (\nabla X_t)^{-1}\sigma(X_t) \equiv 1$ (never true unless $X = X_0 + W$...).

$$Z^{\Re}_t = V^j_t := \mathbb{E}_t\left[\nabla h(X_{\tau_j})\nabla X_{\tau_j}\right]$$

is a martingale, thus, with i_j s.t. $t_{i_j} = r_j$,

$$\begin{aligned}
\sum_{i=0}^{n-1}\mathbb{E}\left[\int_{t_i}^{t_{i+1}}|Z^{\Re}_t - Z^{\Re}_{t_i}|^2 dt\right] &= \sum_{j=0}^{\kappa-1}\sum_{k=i_j}^{i_{j+1}-1}\int_{t_i}^{t_{i+1}}\mathbb{E}\left[|V^j_t - V^j_{t_k}|^2\right]dt \\
&\leq |\pi|\sum_{j=0}^{\kappa-1}\sum_{k=i_j}^{i_{j+1}-1}\mathbb{E}\left[|V^j_{t_{k+1}}|^2 - |V^j_{t_k}|^2\right] \\
&= |\pi|\,\{\mathbb{E}\left[|V^{\kappa-1}_{r_\kappa}|^2 - |V^0_{r_0}|^2\right] \\
&\qquad + \sum_{j=1}^{\kappa-2}\mathbb{E}\left[|V^j_{r_{j+1}}|^2 - |V^{j+1}_{r_{j+1}}|^2\right]\}
\end{aligned}$$

where, by the Lipschitz continuity of ∇h,

$$\begin{aligned}\mathbb{E}\left[|V^j_{r_{j+1}}|^2 - |V^{j+1}_{r_{j+1}}|^2\right] &\le \mathbb{E}\left[\eta_{r_{j+1}}|V^j_{r_{j+1}} - V^{j+1}_{r_{j+1}}|\right] \\ &= \mathbb{E}\left[\eta_{r_{j+1}}|\mathbb{E}_{r_{j+1}}[\nabla h(X_{\tau_j})\nabla X_{\tau_j} - \nabla h(X_{\tau_{j+1}})\nabla X_{\tau_{j+1}}]|\right] \\ &\le \mathbb{E}\left[\hat{\eta}(\tau_{j+1} - \tau_j)^{\frac{1}{2}}\right]\end{aligned}$$

so that

$$\sum_{j=1}^{\kappa-2} \mathbb{E}\left[|V^j_{r_{j+1}}|^2 - |V^{j+1}_{r_{j+1}}|^2\right] \le \sqrt{\kappa}\, \mathbb{E}\left[\hat{\eta}(\tau_{\kappa-1} - \tau_1)^{\frac{1}{2}}\right]$$

• Similarly, if $\nabla h \in C^2$ we apply Itô's lemma to obtain

$$\mathbb{E}\left[|V^j_{r_{j+1}}|^2 - |V^{j+1}_{r_{j+1}}|^2\right] \le \mathbb{E}\left[\hat{\eta}(\tau_{j+1} - \tau_j)\right]$$

so that

$$\sum_{j=1}^{\kappa-2} \mathbb{E}\left[|V^j_{r_{j+1}}|^2 - |V^{j+1}_{r_{j+1}}|^2\right] \le \mathbb{E}\left[\hat{\eta}(\tau_{\kappa-1} - \tau_1)\right].$$

Theorem:(B. & Chassagneux [7]) If $\mathbf{H}_1 : h \in C^1_L$, then

$$\sum_{i=0}^{n-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |Z^{\Re}_t - Z^{\Re}_{t_i}|^2 dt\right] \le C\,\sqrt{\kappa}\,|\pi|\ .$$

If $\mathbf{H}_2 : h \in C^2_L$ and $\sigma \in C^1_L$, then

$$\sum_{i=0}^{n-1} \mathbb{E}\left[\int_{t_i}^{t_{i+1}} |Z^{\Re}_t - Z^{\Re}_{t_i}|^2 dt\right] \le C\,|\pi|\ .$$

3.4 Convergence speed : Discretely reflected case

Theorem:(B. & Chassagneux [7]) Let $\mathbf{H}_1$ hold. Then,

$$\max_{i\leq n-1} \| \, |\overline{Y}^{\pi}_{t_i} - Y^{\Re}_{t_i}| + \sup_{t\in(t_i,t_{i+1}]} |\overline{Y}^{\pi}_{t_{i+1}} - Y^{\Re}_{t}| \, \|_{\mathbf{L}^2} \;\leq\; C(\kappa^{\frac{1}{4}} \, |\pi|^{\frac{1}{2}} + |\pi|^{\frac{1}{4}})$$

and

$$\|\overline{Z}^{\pi} - Z^{\Re}\|_{\mathbf{L}^2_{\mathcal{P}}} \;\leq\; C(\kappa^{\frac{1}{2}} \, |\pi|^{\frac{1}{2}} + |\pi|^{\frac{1}{4}})$$

Theorem: (B. & Chassagneux [7]) Let $\mathbf{H}_2$ hold. Then,

$$\max_{i\leq n-1} \| \, |\overline{Y}^{\pi}_{t_i} - Y^{\Re}_{t_i}| + \sup_{t\in(t_i,t_{i+1}]} |\overline{Y}^{\pi}_{t_{i+1}} - Y^{\Re}_{t}| \, \|_{\mathbf{L}^2} \;\leq\; C|\pi|^{\frac{1}{2}}$$

and

$$\|\overline{Z}^{\pi} - Z^{\Re}\|_{\mathbf{L}^2_{\mathcal{P}}} \;\leq\; C(\kappa^{\frac{1}{2}} \, |\pi|^{\frac{1}{2}} + |\pi|^{\frac{1}{2}})$$

Remark: can do better when $X^{\pi} = X$ on π.

3.5 Convergence speed : Continuously reflected case

Theorem:(B. & Chassagneux [7]) Take $\Re = \pi$. Let $\mathbf{H}_1$ hold. Then,

$$\max_{i \leq n-1} \| \, |\overline{Y}^{\pi}_{t_i} - Y_{t_i}| + \sup_{t \in (t_i, t_{i+1}]} |\overline{Y}^{\pi}_{t_{i+1}} - Y_t| \, \|_{\mathbf{L}^2} \; \leq \; C_L \, \alpha(\pi) \; ,$$

with $\alpha(\pi) = |\pi|^{\frac{1}{4}}$ under $\mathbf{H}_1$ and $\alpha(\pi) = |\pi|^{\frac{1}{2}}$ under $\mathbf{H}_2$.
Moreover, under $\mathbf{H}_1$,

$$\|\overline{Z}^{\pi} - Z\|_{\mathbf{L}^2_{\mathcal{P}}} \; \leq \; C_L \, |\pi|^{\frac{1}{4}} \; ,$$

If $\mathbf{H}_2$ holds and $X^{\pi} = X$ on π, then

$$\|\overline{Z}^{\pi} - Z\|_{\mathbf{L}^2_{\mathcal{P}}} \; \leq \; C_L \, |\pi|^{\frac{1}{2}} \; .$$

4 BSDEs in a domain

4.1 Approximation scheme

We consider

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s \\ Y_t &= g(X_\tau) + \int_{t\wedge\tau}^\tau f_s(X_s, Y_s, Z_s)ds - \int_{t\wedge\tau}^\tau Z_s dW_s, \end{aligned}$$

where

$$\tau := \inf\{t \geq 0 : X_t \notin \mathcal{O}\} \wedge T,$$

for some open set $\mathcal{O}$.

We approximate the first exit time τ by

$$\tau^\pi := \inf\{t \in \pi \ : \ X^\pi_t \notin \mathcal{O}\} \wedge 1\ .$$

The Euler scheme is defined as previously with $Y^\pi_{\tau^\pi} = g(X^\pi_{\tau^\pi})$ and

$$\begin{aligned} Z^\pi_{t_i} &= n\, \mathbb{E}_{t_i}[Y^\pi_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})] \\ Y^\pi_{t_i} &= \mathbb{E}_{t_i}[Y^\pi_{t_{i+1}}] + \frac{1}{n}\, f(X^\pi_{t_i}, Y^\pi_{t_i}, Z^\pi_{t_i}). \end{aligned}$$

For Lipschitz continuous coefficients, we get a similar error term + the exit time approximation error:

$$\text{Err} \ \leq \ C \left(\underbrace{|\pi|^{\frac{1}{2}} + \mathcal{R}(Y) + \mathcal{R}(Z)}_{\text{previous terms}} + \text{Err}(\tau - \tau^\pi) \right)$$

4.2 Regularity of (Y, Z)

We can not follow the Malliavin derivative approach anymore because X_τ is not Malliavin differentiable in general...

However, one can follow the PDE approach (say $f \equiv 0$): $\partial_x v(\cdot, X)\nabla X$ is a martingale, which can be read from the pde, and therefore

$$\begin{aligned} Z_t &= \partial_x v(t, X_t)\nabla X_t \mathbf{1}_{t\leq\tau}(\nabla X_t)^{-1}\sigma(X_t) \\ &= \mathbb{E}_t[\partial_x v(\tau, X_\tau)\nabla X_\tau]\mathbf{1}_{t\leq\tau}(\nabla X_t)^{-1}\sigma(X_t) \\ &\neq \mathbb{E}_t[\partial_x g(X_\tau)\nabla X_\tau]\mathbf{1}_{t\leq\tau}(\nabla X_t)^{-1}\sigma(X_t) \end{aligned}$$

4.2 Regularity of (Y, Z)

We can not follow the Malliavin derivative approach anymore because X_τ is not Malliavin differentiable in general...

However, one can follow the PDE approach (say $f \equiv 0$): $\partial_x v(\cdot, X)\nabla X$ is a martingale, which can be read from the pde, and therefore

$$\begin{aligned} Z_t &= \partial_x v(t, X_t)\nabla X_t \mathbf{1}_{t\le\tau}(\nabla X_t)^{-1}\sigma(X_t) \\ &= \mathbb{E}_t[\partial_x v(\tau, X_\tau)\nabla X_\tau]\mathbf{1}_{t\le\tau}(\nabla X_t)^{-1}\sigma(X_t) \end{aligned}$$

If $\partial_x v$ bounded, we can use the same martingale techniques as in the case $\mathcal{O} = \mathbb{R}^d$ to bound $\mathcal{R}(Z)$ (and Z as well, leading to a bound on $\mathcal{R}(Y)$) !

Assume from now on that:

- All coefficients are Lipschitz.
- $\mathcal{O} := \bigcap_{\ell=1}^{m} \mathcal{O}^\ell$ where $\mathcal{O}^\ell$ is C^2 with a compact boundary.
- Exterior sphere condition + interior truncated cone condition.
- The boundary satisfies a non characteristic condition outside a neighborhood of $\mathcal{C} := \bigcap_{\ell \neq k=1}^{m} \partial\mathcal{O}^\ell \cap \partial\mathcal{O}^k$ and σ is uniformly elliptic on a neighborhood of $\mathcal{C}$.
- $g \in C^2(\overline{O})$ and $\|\partial_x g\| + \|\partial^2_{xx} g\| \ \leq \ L \quad$ on $\overline{O}$.

Proposition: (B. & Menozzi [10]) Let v be such that $Y = v(\cdot, X)$. Then, v is uniformly Lipschitz in space and $\frac{1}{2}$-Hölder in time.

Corollary: Assume that the above conditions hold. Then,

$$\mathcal{R}(Y) + \mathcal{R}(Z) = O(|\pi|^{\frac{1}{2}}) \ .$$

4.3 Study of the exit time approximation error term (strong)

We have (on $\tau^\pi \le \tau$)

$$Y^\pi_{\tau^\pi} - Y_{\tau^\pi} = g(X^\pi_{\tau^\pi}) - g(X_\tau) - \int_{\tau^\pi}^{\tau} f(\cdots)ds + \int_{\tau^\pi}^{\tau} Z_s dW_s.$$

Thus, because of the Ito integral,

$$\begin{aligned}\mathbb{E}\left[|Y^\pi_{\tau^\pi} - Y_{\tau^\pi}|^2\right] &= O\left(\mathbb{E}\left[|g(X^\pi_{\tau^\pi}) - g(X_\tau)|^2\right] + \mathbb{E}\left[\xi|\tau^\pi - \tau|^2\right]\right) \\ &+ O\left(\mathbb{E}\left[\xi|\tau^\pi - \tau|\right]\right).\end{aligned}$$

This leads to $\mathrm{Err}(\tau - \tau^\pi) = O_\varepsilon\left(\mathbb{E}\left[|\tau^\pi - \tau|\right]^{\frac{1}{2}-\varepsilon}\right)$.

Theorem: (B. & Menozzi [10], B., Geiss & Gobet [9])

$$\mathbb{E}\left[|\tau^\pi - \tau|\right] \le O(|\pi|^{\frac{1}{2}}).$$

By combining everything together...

Theorem: (B. & Menozzi [10])

$$\text{Err} \;\leq\; C\Big(|\pi|^{\frac{1}{2}} + \underbrace{\mathcal{R}(Y)+\mathcal{R}(Z)}_{\text{gradient bound : } |\pi|^{\frac{1}{2}}} + \underbrace{\text{Err}(\tau-\tau^{\pi})}_{O\left(\mathbb{E}[\xi|\tau^{\pi}-\tau|]^{\frac{1}{2}}\right)=O_{\varepsilon}\left(|\pi|^{\frac{1}{4}-\varepsilon}\right)}\Big) \;\leq\; C_{\varepsilon}|\pi|^{\frac{1}{4}-\varepsilon}\,.$$

4.4 Study of the exit time approximation error term (weak)

We now stop at $\tau \wedge \tau^\pi$:

$$\begin{aligned} Y^\pi_{\tau\wedge\tau^\pi} - Y_{\tau\wedge\tau^\pi} &= \mathbb{E}_{\tau\wedge\tau^\pi}[g(X^\pi_{\tau^\pi}) - g(X_\tau)] - \mathbb{E}_{\tau\wedge\tau^\pi}[\int_{\tau\wedge\tau^\pi}^{\tau\vee\tau^\pi} f(\cdots)ds] \\ &= O(|\pi|^{\frac{1}{2}}, \omega) + \mathbb{E}_{\tau\wedge\tau^\pi}[\int_{\tau\wedge\tau^\pi}^{\tau\vee\tau^\pi} F(\omega, g, Dg, D^2g)(\cdots)ds] \\ &- \mathbb{E}_{\tau\wedge\tau^\pi}[\int_{\tau\wedge\tau^\pi}^{\tau\vee\tau^\pi} f(\cdots)ds], \end{aligned}$$

which leads to

$$\begin{aligned} \mathbb{E}\left[|Y^\pi_{\tau\wedge\tau^\pi} - Y_{\tau\wedge\tau^\pi}|^2\right] &= O(|\pi|) + O\left(\mathbb{E}\left[\ \mathbb{E}_{\tau\wedge\tau^\pi}[\xi|\tau^\pi - \tau|]^2\ \right]\right) \\ &= O_\varepsilon\left(|\pi|^{1-\varepsilon}\right)\ . \end{aligned}$$

By combining everything together...

Theorem: (B. & Menozzi [10])

$$\max_{i<n}\mathbb{E}\left[\sup_{t\in[t_i,t_{i+1}]}|Y_t-Y^\pi_{t_i}|^2\mathbf{1}_{t\le\tau\wedge\tau^\pi}\right]^{\frac{1}{2}}+\mathbb{E}\left[\sum_i\int_{t_i}^{t_{i+1}}\|Z_t-Z^\pi_{t_i}\|^2\mathbf{1}_{t\le\tau\wedge\tau^\pi}dt\right]^{\frac{1}{2}}$$

$$\le\ C(|\pi|^{\frac{1}{2}}+\underbrace{\mathcal{R}(Y)+\mathcal{R}(Z)}_{\text{gradient bound : }|\pi|^{\frac{1}{2}}}+\underbrace{\mathrm{Err}(\tau-\tau^\pi)}_{O\left(\mathbb{E}\left[\ \mathbb{E}_{\tau\wedge\tau^\pi}[\xi|\tau^\pi-\tau|]^2\ \right]\right)=O_\varepsilon(|\pi|^{\frac{1}{2}-\varepsilon})})$$

$$\le C_\varepsilon|\pi|^{\frac{1}{2}-\varepsilon}\ .$$

In particular, $|Y_0-Y_0^\pi|\le C_\varepsilon(|\pi|^{\frac{1}{2}-\varepsilon})$.

Remark: Should be able to get rid of the ε, at least if $g\in C_b^2$ and f is bounded.

5 Irregular terminal condition

This is the case where g is not Lipschitz, possibly discontinuous. Based on earlier works of S. Geiss, it was first treated by Gobet & Makhlouf [31] (see also Gobet, Geiss & Geiss [28]).

Main idea: The important quantity is

$$\sup_{t<1} \frac{\overbrace{\mathbb{E}[\,|g(X_1) - \mathbb{E}_t[g(X_1)]|^2]}^{V_{t,1}}}{(1-t)^\alpha}.$$

We say that $g \in \mathbf{L}_{2,\alpha}$, $0 < \alpha \leq 1$, if the above is finite and $g(X_1) \in \mathbf{L}^2$ (if $X_1 = W_1$, this corresponds to a Besov space).

When $g \in \mathbf{L}_{2,\alpha}$,

$$\mathbb{E}[|\partial^2_{xx} v(t, X_t)|^2] \leq C V_{t,1}/(1-t)^2.$$

This is obtained under smoothness/uniform ellipticity conditions by estimating the second order derivative (representation + integration by parts).

Because Z is essentially $\partial_x v(\cdot, X)$ (up to nice behaving terms), not surprisingly:

$$\int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt \leq C(\cdots + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (t_{i+1} - s)\mathbb{E}[|\partial_{xx}^2 v(s, X_s)|^2] ds)$$

i.e.

$$\int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt \leq C(|\pi| + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (t_{i+1} - s)\frac{V_{s,1}}{(1-s)^2} ds).$$

If one takes $t_i = 1 - (1 - i/n)^{\frac{1}{\beta}}$, with $\beta < \alpha$, then

$$\begin{aligned}
\sum_i \int_{t_i}^{t_{i+1}} (t_{i+1} - s)\frac{V_{s,1}}{(1-s)^2} ds &= \sum_i \int_{t_i}^{t_{i+1}} \frac{t_{i+1} - s}{(1-s)^{1-\beta}} (1-s)^{1-\beta} \frac{V_{s,1}}{(1-s)^2} ds \\
&\leq C \sup_i \frac{t_{i+1} - t_i}{(1-t_i)^{1-\beta}} \int_0^1 (1-s)^{1-\beta+\alpha-2} ds \\
&= C \sup_i \frac{t_{i+1} - t_i}{(1-t_i)^{1-\beta}} \\
&\leq C \frac{1}{\beta n}.
\end{aligned}$$

Conclusion: For $t_i = 1 - (1 - i/n)^{\frac{1}{\beta}}$, with $\beta < \alpha$, we retrieve

$$\int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt \leq C\ |\pi|.$$

If one takes a uniform grid, then one only has:

$$\int_0^1 \mathbb{E}\left[|Z_t - \overline{Z}_t|^2\right] dt \leq C\ |\pi|^\alpha.$$

Example: If g is α-Hölder with polynomial growth then $g(X_1) \in \mathbf{L}_{2,\alpha}$. If $g(x) = \mathbf{1}_{x \geq x_o}$ and $X_1 = W_1$, then $g(X_1) \in \mathbf{L}_{2,\frac{1}{2}}$.

6 Additional references

More can be done, here are (some) additional contributions:

- Doubly reflected BSDEs: Chassagneux [15].
- Multi-dimensional BSDEs with oblique reflection : Chassagneux, Elie & Kharroubi [18].
- Quadratic BSDEs: Chassagneux & Richou [19].
- Coupled FBSDEs: Delarue & Menozzi [24], Bender & Zhang [4].
- Second order BSDEs: Fahim, Touzi & Warrin [27], see also B., Elie & Touzi [8].
- Linear multi-step or Runge-Kutta type schemes: Chassagneux [16], and Chassagneux & Crisan [17].

Part III

Computation of conditional expectations

1 Motivation

The typical numerical scheme is of the form

$$\begin{aligned}\overline{Y}^{\pi}_{t_i} &= \mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}] + \frac{1}{n} f(X^{\pi}_{t_i}, \overline{Y}^{\pi}_{t_i}, \overline{Z}^{\pi}_{t_i}) \\ \overline{Z}^{\pi}_{t_i} &= n\mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})],\end{aligned}$$

where $\overline{Y}^{\pi}_{t_n} = g(X^{\pi}_{t_n})$.

To compute this quantities, one needs to be able to estimate efficiently the different conditional expectations.

2 Generalities

In practice, we approximate

$$\begin{aligned} Y_1^\pi &= g(X_1^\pi) \\ Y_{t_i}^\pi &= \mathbb{E}\left[Y_{t_{i+1}}^\pi \mid X_{t_i}^\pi\right] + \frac{1}{n} f(X_{t_i}^\pi, Y_{t_i}^\pi, \overline{Z}_{t_i}^\pi) \end{aligned}$$

by

$$\begin{aligned} \hat{Y}_1^\pi &= g(X_1^\pi) \\ \hat{Y}_{t_i}^\pi &= \hat{E}\left[\hat{Y}_{t_{i+1}}^\pi \mid X_{t_i}^\pi\right] + \frac{1}{n} f(X_{t_i}^\pi, \hat{Y}_{t_i}^\pi, \hat{Z}_{t_i}^\pi) \\ \hat{Z}_{t_i}^\pi &= \hat{E}\left[\hat{Y}_{t_{i+1}}^\pi (W_{t_{i+1}} - W_{t_i}) \mid X_{t_i}^\pi\right] \end{aligned}$$

where $\hat{E}\left[\hat{Y}_{t_{i+1}}^\pi \mid X_{t_i}^\pi\right]$ and $\hat{E}\left[\hat{Y}_{t_{i+1}}^\pi (W_{t_{i+1}} - W_{t_i}) \mid X_{t_i}^\pi\right]$ are estimators of $\mathbb{E}\left[\hat{Y}_{t_{i+1}}^\pi \mid X_{t_i}^\pi\right]$ and $\mathbb{E}\left[\hat{Y}_{t_{i+1}}^\pi (W_{t_{i+1}} - W_{t_i}) \mid X_{t_i}^\pi\right]$

Theorem (B. & Touzi [11],...)

$$\|\hat{Y}^{\pi}_{t_i} - Y^{\pi}_{t_i}\|_{\mathbf{L}^p} \;\leq\; nC_p \max_{0\leq j\leq n-1} \mathcal{E}_{j,p}(\hat{E},n)$$

with

$$\begin{aligned}\mathcal{E}_{j,p}(\hat{E},n) \;:=\; & \|\hat{E}\left[\hat{Y}^{\pi}_{t_{j+1}} \mid X^{\pi}_{t_j}\right] - \mathbb{E}\left[\hat{Y}^{\pi}_{t_{j+1}} \mid X^{\pi}_{t_j}\right]\|_{\mathbf{L}^p} \\ & +\|\hat{E}\left[\hat{Y}^{\pi}_{t_{j+1}}(W_{t_{i+1}} - W_{t_i}) \mid X^{\pi}_{t_j}\right] - \mathbb{E}\left[\hat{Y}^{\pi}_{t_{j+1}}(W_{t_{i+1}} - W_{t_i}) \mid X^{\pi}_{t_j}\right]\|_{\mathbf{L}^p}\end{aligned}$$

Remark :

1- At best $\mathcal{E}_{j,p}(\hat{E},n) \sim N^{-1/2}$ if computed by pure Monte-Carlo.

2- $\|Y_{t_i} - Y^{\pi}_{t_i}\|_{\mathbf{L}^p} \sim n^{-1/2}$, hence to get $\|\hat{Y}^{\pi}_{t_i} - Y^{\pi}_{t_i}\|_{\mathbf{L}^p} \sim n^{-1/2}$ we need to take at least $N = n^3$ if the global error is of order $n^{-\frac{1}{2}} + nN^{-\frac{1}{2}}$.

3 Integration by parts technique

See B., Ekeland & Touzi [6], B. & Warin [12], Crisan, Manolarakis & Touzi [23].

3.1 Conditional expectation representation in the Gaussian case

Set $v(W_{t_i}) = Y^{\pi}_{t_i}$ (we omit the dependence of v on the time variable).

Reduction of the problem :

$$\mathbb{E}\left[v(W_{t_{i+1}}) \mid W_{t_i} = w\right] = \frac{\mathbb{E}\left[\delta_w(W_{t_i})v(W_{t_{i+1}})\right]}{\mathbb{E}\left[\delta_w(W_{t_i})\right]}$$

δ_w Dirac mass at w.

- Integration by parts argument (f_δ = density $\mathcal{N}(0,\delta)$):

$$\mathbb{E}[\delta_w(W_{t_i})v(W_{t_{i+1}})]$$
$$=$$
$$\int\int \delta_w(x)\, v(x+y)\, f_{t_i}(x) f_{(t_{i+1}-t_i)}(y) dx dy$$
$$=$$
$$\int\int \mathbf{1}_{\{x\geq w\}} \left[v(x+y)\frac{x}{t_i} - v'(x+y) \right] f_{t_i}(x) f_{(t_{i+1}-t_i)}(y) dx dy$$
$$=$$
$$\int\int \mathbf{1}_{\{x\geq w\}} v(x+y) \left[\frac{x}{t_i} - \frac{y}{t_{i+1}-t_i} \right] f_{t_i}(x) f_{(t_{i+1}-t_i)}(y) dx dy\ ,$$
$$=$$
$$\mathbb{E}\left[\mathbf{1}_{\{W_{t_i}\geq w\}} v\left(W_{t_{i+1}}\right) \left(\frac{W_{t_i}}{t_i} - \frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i} \right) \right]$$

Alternative formulation :

$$\mathbb{E}\left[v(W_{t_{i+1}}) \mid W_{t_i} = w\right] = \frac{\mathbb{E}\left[\mathbf{1}_{\left\{W_{t_i}\geq w\right\}} v\left(W_{t_{i+1}}\right)\left(\frac{W_{t_i}}{t_i} - \frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)\right]}{\underbrace{\mathbb{E}\left[\mathbf{1}_{\left\{W_{t_i}\geq w\right\}}\frac{W_{t_i}}{t_i}\right]}_{\mathbb{E}\left[\delta_w(W_{t_i})\right]=f_{W_{t_i}}(w)}}$$

Monte-Carlo estimator : $\{W^{(\ell)}\}_\ell$, N copies of W

$$\hat{E}\left[v(W_{t_{i+1}}) \mid W_{t_i} = w\right] := \frac{\frac{1}{N}\sum_\ell \mathbf{1}_{\left\{W^{(\ell)}_{t_i}\geq w\right\}} v\left(W^{(\ell)}_{t_{i+1}}\right)\left(\frac{W^{(\ell)}_{t_i}}{t_i} - \frac{W^{(\ell)}_{t_{i+1}}-W^{(\ell)}_{t_i}}{t_{i+1}-t_i}\right)}{\frac{1}{N}\sum_\ell \mathbf{1}_{\left\{W^{(\ell)}_{t_i}\geq w\right\}}\frac{W^{(\ell)}_{t_i}}{t_i}}$$

Variance estimation

$$\mathrm{Var}\left[\frac{W^{(\ell)}_{t_{i+1}} - W^{(\ell)}_{t_i}}{t_{i+1}-t_i}\right]^{\frac{1}{2}} = \frac{(t_{i+1}-t_i)^{\frac{1}{2}}}{t_{i+1}-t_i} \; = \; n^{\frac{1}{2}}$$

leading to

$$\mathrm{Var}\left[\frac{1}{N}\sum_{\ell}\mathbf{1}_{\left\{W^{(\ell)}_{t_i}\geq w\right\}} v\left(W^{(\ell)}_{t_{i+1}}\right)\left(\frac{W^{(\ell)}_{t_i}}{t_i} - \frac{W^{(\ell)}_{t_{i+1}} - W^{(\ell)}_{t_i}}{t_{i+1}-t_i}\right)\right]^{\frac{1}{2}} \sim \frac{n^{\frac{1}{2}}}{N^{\frac{1}{2}}} \, .$$

3.2 Variance Reduction in the Gaussian Case: Control variate

$$\mathbb{E}\left[v\left(W_{t_{i+1}}\right)\left(\frac{W_{t_i}}{t_i}-\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)\right]$$

$$=$$

$$\int\int v\left(x+y\right)\left[\frac{x}{t_i}-\frac{y}{t_{i+1}-t_i}\right]f_{t_i}(x)f_{(t_{i+1}-t_i)}(y)dxdy$$

$$=$$

$$\int\int v'\left(x+y\right)f_{t_i}(x)f_{(t_{i+1}-t_i)}(y)-v'\left(x+y\right)f_{t_i}(x)f_{(t_{i+1}-t_i)}(y)dxdy$$

$$=$$

$$0$$

We can then replace $\mathbf{1}_{\{W_{t_i}\geq w\}}$ by $(\mathbf{1}_{\{W_{t_i}\geq w\}}-c(w))$:

$$\mathbb{E}\left[v(W_{t_{i+1}})\mid W_{t_i}=w\right] = \frac{\mathbb{E}\left[\left(\mathbf{1}_{\{W_{t_i}\geq w\}}-c(w)\right)v\left(W_{t_{i+1}}\right)\left(\frac{W_{t_i}}{t_i}-\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)\right]}{\mathbb{E}\left[\left(\mathbf{1}_{\{W_{t_i}\geq w\}}-\tilde{c}(w)\right)\frac{W_{t_i}}{t_i}\right]}$$

3.3 Variance Reduction in the Gaussian case : Localization

Take φ smooth in $\mathbf{L}^2(\mathbb{R})$ with $\varphi(0) = 1$

$$E[\delta_w(W_{t_i})v(W_{t_{i+1}})]$$

$$=$$

$$\int\int \delta_w(x)\,\varphi(x-w)v(x+y)\,f_{t_i}(x)f_{(t_{i+1}-t_i)}(y)dxdy$$

$$=$$

$$\dots$$

$$=$$

$$\mathbb{E}\left[\mathbf{1}_{\{W_{t_i}\geq w\}}v\left(W_{t_{i+1}}\right)\left\{\varphi\left(W_{t_i}-w\right)\left(\frac{W_{t_i}}{t_i}-\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)-\varphi'\left(W_{t_i}-w\right)\right\}\right]$$

This shows that

$$\mathbb{E}\left[v(W_{t_{i+1}}) \mid W_{t_i} = w\right]$$
$$=$$
$$\frac{\mathbb{E}\left[\mathbf{1}_{\{W_{t_i} \geq w\}} v\left(W_{t_{i+1}}\right)\left\{\varphi\left(W_{t_i}-w\right)\left(\frac{W_{t_i}}{t_i}-\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)-\varphi'\left(W_{t_i}-w\right)\right\}\right]}{\mathbb{E}\left[\mathbf{1}_{\{W_{t_i} \geq w\}}\left\{\tilde{\varphi}\left(W_{t_i}-w\right)\frac{W_{t_i}}{t_i}-\tilde{\varphi}'\left(W_{t_i}-w\right)\right\}\right]}$$

One can then try to minimize an indicator of the variance

$$\min_{\varphi \in \mathbf{L}^2(\mathbb{R}),\, \varphi(0)=1} \int_{\mathbb{R}} \mathbb{E}\left[\mathbf{1}_{\{W_{t_i} \geq w\}} \left[F\varphi\left(W_{t_i} - w\right) - G\varphi'\left(W_{t_i} - w\right)\right]^2\right] dw$$

with

$$\begin{aligned} F &= v\left(W_{t_{i+1}}\right)\left(\frac{W_{t_i}}{t_i} - \frac{W_{t_{i+1}} - W_{t_i}}{t_{i+1} - t_i}\right) \\ G &= v\left(W_{t_{i+1}}\right) \end{aligned}$$

Calculus of Variation : φ is optimal iif for all smooth ϕ with $\phi(0) = 0$ and compact support, and $\varepsilon > 0$

$$\int_{\mathbb{R}} \mathbb{E}\left[\mathbf{1}_{\{W_{t_i} \geq w\}} \left[F\varphi\left(W_{t_i} - w\right) - G\varphi'\left(W_{t_i} - w\right)\right]^2\right]$$
$$\leq \int_{\mathbb{R}} \mathbb{E}\left[\mathbf{1}_{\{W_{t_i} \geq w\}} \left[F(\varphi \pm \varepsilon\phi)\left(W_{t_i} - w\right) - G(\varphi' \pm \varepsilon\phi')\left(W_{t_i} - w\right)\right]^2\right]$$

Sending $\varepsilon \to 0$

$$\begin{aligned} 0 &= \int_{\mathbb{R}} \mathbb{E}\Big[\mathbf{1}_{\{W_{t_i} \geq w\}} \left(F\varphi\left(W_{t_i} - w\right) - G\varphi'\left(W_{t_i} - w\right)\right) \\ &\qquad \left(F\phi\left(W_{t_i} - w\right) - G\phi'\left(W_{t_i} - w\right)\right)\Big]\, dw \\ &= \mathbb{E}\left[\int_0^\infty \left(F\varphi\left(y\right) - G\varphi'\left(y\right)\right)\left(F\phi\left(y\right) - G\phi'\left(y\right)\right) dy\right] \\ &\qquad \text{Fubini + change of variable } y = W_{t_i}(\omega) - w \\ &= \mathbb{E}\left[\int_0^\infty \phi(y)\left(F^2\varphi\left(y\right) - G^2\varphi''\left(y\right)\right) dy\right] \\ &\qquad \text{integration by parts} \\ &= \int_0^\infty \phi(y)\left(\mathbb{E}\left[F^2\right]\varphi\left(y\right) - \mathbb{E}\left[G^2\right]\varphi''\left(y\right)\right) dy \end{aligned}$$

$$\Rightarrow \mathbb{E}\left[F^2\right]\varphi\left(y\right) - \mathbb{E}\left[G^2\right]\varphi''\left(y\right) = 0.$$

Optimal Localizing Function : $\varphi(y) = e^{-\hat{\eta}y}$ with

$$\begin{aligned}
\hat{\eta}^2 &= \mathbb{E}\left[F^2\right]/\mathbb{E}\left[G^2\right] \\
&= \frac{\mathbb{E}\left[v\left(W_{t_{i+1}}\right)^2\left(\frac{W_{t_i}}{t_i} - \frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)^2\right]}{\mathbb{E}\left[v\left(W_{t_{i+1}}\right)^2\right]} \\
&\sim n
\end{aligned}$$

For $\varphi(y) = e^{-\hat{\eta}y}$ **we have** $\varphi'(y) = -\hat{\eta}\varphi(y)$

$$\mathbb{E}\left[\mathbf{1}_{\{W_{t_i}\geq w\}} v\left(W_{t_{i+1}}\right)\left\{\varphi\left(W_{t_i}-w\right)\left(\frac{W_{t_i}}{t_i} - \frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right) - \varphi'\left(W_{t_i}-w\right)\right\}\right]$$

$$=$$

$$\mathbb{E}\left[\mathbf{1}_{\{W_{t_i}\geq w\}} v\left(W_{t_{i+1}}\right)\left\{\varphi\left(W_{t_i}-w\right)\left(\frac{W_{t_i}}{t_i} - \frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i} + \hat{\eta}\right)\right\}\right]$$

where

$$\mathbb{E}\left[\left(\frac{W_{t_{i+1}}-W_{t_i}}{t_{i+1}-t_i}\right)^2\right]^{\frac{1}{2}} = \sqrt{n} \sim \hat{\eta}$$

3.4 Full numerical scheme (f indep. of Z for simplicity)

We consider N copies $(X^{\pi^{(1)}}, \ldots, X^{\pi^{(N)}})$ of X^π.

Initialization : For all j : $\hat{Y}_1^{\pi^{(j)}} = g\left(X_1^{\pi^{(j)}}\right)$.

Backward induction : For $i = n-1, \ldots, 1$, we set, for all j :

$$\hat{Y}_{t_i}^{\pi^{(j)}} = \hat{E}\left[\hat{Y}_{t_{i+1}}^\pi \mid X_{t_i}^\pi = X_{t_i}^{\pi^{(j)}}\right] + \frac{1}{n} f(X_{t_i}^{\pi^{(j)}}, \hat{Y}_{t_i}^{\pi^{(j)}})$$

where

$$\hat{E}\left[\hat{Y}_{t_{i+1}}^\pi \mid X_{t_i}^\pi = X_{t_i}^{\pi^{(j)}}\right] = \frac{\sum_{\ell \le N} \mathbf{1}_{X_{t_{i-1}}^{\pi^{(j)}} \le X_{t_{i-1}}^{\pi^{(\ell)}}} \hat{Y}_{t_{i+1}}^{\pi^{(\ell)}} \mathcal{S}^{h^{(\ell)}} \left(\varphi(X_{t_{i-1}}^{\pi^{(\ell)}} - X_{t_{i-1}}^{\pi^{(j)}})\right)}{\sum_{\ell \le N} \mathbf{1}_{X_{t_{i-1}}^{n^{(j)}} \le X_{t_{i-1}}^{\pi^{(\ell)}}} \mathcal{S}^{h^{(\ell)}} \left(\varphi(X_{t_{i-1}}^{\pi^{(\ell)}} - X_{t_{i-1}}^{\pi^{(j)}})\right)}$$

(up to an additional truncation given by a-priori bounds).

Remark: Only needs $O(N \ln N)$ operations (first sort the simulations) and not N^2.

In dimension d, we can do the same by performing d integration by parts. If f depends on Z, the corresponding conditional expectation is computed similarly.

Theorem (B. & Touzi [11])

$$\|\hat{Y}^{\pi}_{t_i} - Y^{\pi}_{t_i}\|_{\mathbf{L}^p} \;\leq\; nC_p \max_{0\leq j\leq n-1} \mathcal{E}_{j,p}(\hat{E}, n)$$

If we choose $\varphi^n = \varphi(\sqrt{n}x)$. Then,

$$\max_{1\leq j\leq n-1} \mathcal{E}_{j,p}(\hat{E}, n) \;\leq\; Cn^{\frac{d}{4p}} N^{-\frac{1}{2p}} \; .$$

Global error :

$$\max_{0\le i\le n} \|\hat{Y}^{\pi}_{t_i} - Y_{t_i}\|_{\mathbf{L}^p} \;\le\; C_p\left(n^{-\frac{1}{2}} + n\frac{n^{\frac{d}{4p}}}{N^{\frac{1}{2p}}}\right)$$

• $n^{-\frac{1}{2}}$: discretization error

• n : number of regression estimations

• $N^{-\frac{1}{2p}}$: convergence rate of the regression estimator in terms of the number of simulations N.

• $n^{\frac{d}{4p}}$: "variance" of the regression operator.

For an $\mathbf{L}^p$ error of order of $n^{-\frac{1}{2}}$: $N \sim n^{3p+\frac{d}{2}}$.

3.5 The general case

- The conditional expectations can be written similarly by using a Malliavian calculus integration by parts argument. The terms corresponding to the $\mathcal{S}^h$ are now Skorohod integrals that can be decomposed into Itô and Lebesgue integrals. See B., Ekeland & Touzi [6]. The analysis of the error is the same. See B. & Touzi [11].

- The sums in the estimators can be computed efficiently by using the equivalent of a sorting method in dimension d. Its complexity is of order $O(N(\ln N)^{(d-1)\vee 1})$. See B. & Warin [12].

- An alternative formulation with reduced complexity in the computation of the Malliavin weights is given in Crisan, Manolarakis & Touzi [23]. It consists in neglecting a (neglectable) term whose computation complexity is not neglectable.

3.6 Importantly

- The localization by the φ functions is crucial in practice.
- Variance reduction by control variate is also important. One can replace Y^π by $Y^\pi - \tilde{Y}$ such that the conditional expectations of $\tilde{Y}_{t_{i+1}}$ and $\tilde{Y}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})$ are explicit.
- Truncation is also crucial: a-priori bounds for Y and Z are often easy to compute (recall the conditional expectation expression for Z obtained above).

4 Non-parametric regression approach

Initiated by Carrière [14] and Longstaff & Schwartz [34] in the context of American option pricing (made first rigorous by Clément, Lamberton & Protter [20]).

General idea: replace $\mathbb{E}[\hat{Y}^{\pi}_{t_{i+1}}|X^{\pi}_{t_i}]$ by an estimator of its projection on the space generated by $(\psi_m(X^{\pi}_{t_i}))_{m\leq M}$ for some deterministic "basis functions" $(\psi_m)_{m\leq M}$.

Given simulations $(\xi^{(j)}, X^{\pi^{(j)}}_{t_i})_{j\leq N}$ of $(\xi, X^{\pi}_{t_i})$, we set

$$\hat{E}[\xi|X^{\pi}_{t_i} = X^{\pi^{(j)}}_{t_i}] := \sum_{m=1}^{M} \hat{\alpha}_{t_i,m}\psi_m(X^{\pi^{(j)}}_{t_i})$$

where $(\hat{\alpha}_{t_i,m})_{m\leq M}$ minimizes

$$\sum_{j=1}^{N} |\xi^{(j)} - \sum_{m=1}^{M} \alpha_m\psi_m(X^{\pi^{(j)}}_{t_i})|^2$$

over $(\alpha_m)_{m\leq M} \in \mathbb{R}^M$.

4.1 Full numerical scheme (f indep. of Z for simplicity)

We consider N copies $(X^{\pi^{(1)}}, \ldots, X^{\pi^{(N)}})$ of X^π.

Initialization : For all j : $\hat{Y}_1^{\pi^{(j)}} = g\left(X_1^{\pi^{(j)}}\right)$.

Backward induction (for the explicit scheme): For $i = n-1, \ldots, 1$, we set, for all j :

$$\hat{Y}_{t_i}^{\pi^{(j)}} \;=\; \hat{E}\left[\hat{Y}_{t_{i+1}}^{\pi} + \frac{1}{n} f(X_{t_i}^{\pi}, \hat{Y}_{t_{i+1}}^{\pi}) \mid X_{t_i}^{\pi} = X_{t_i}^{\pi^{(j)}}\right]$$

4.2 Choice of the basis functions

Ideally, one should use a first element $\psi_{t_i,1}$ that is supposed to be close to the real conditional expectation and then complete the basis so that it is orthogonal for the law of $X^\pi_{t_i}$ (or a proxy).

Very often, people use polynomials. The choice is very difficult in practice in "high" dimension.... and the induced error difficult to quantify.

It turns out to be more stable to use local polynomials: e.g. piecewise linear maps on a space partition. When the size of the partition vanishes, we are certain that convergence holds (and we can quantify it, think at v being Lipschitz).

This partition should be consistent with the law of $X^\pi_{t_i}$. It can be constructed in an adaptative way: simulate copies of $X^\pi_{t_i}$, then built up a partition so that each part contains approximately the same number of points. This can be done efficiently by sorting like methods, see [12].

4.3 Error analysis

The precise error analysis du to estimating the coefficients of the regression has been performed in Gobet, Lemore & Warin [33].

The expression of the error takes half a page... see the paper.

4.4 The multistep forward dynamic programming scheme

The MDP scheme consists in replacing

$$\overline{Y}^{\pi}_{t_i} := \mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}] + \frac{1}{n} f(X^{\pi}_{t_i}, \overline{Y}^{\pi}_{t_i}, \overline{Z}^{\pi}_{t_i}) \ , \ \overline{Z}^{\pi}_{t_i} := n\mathbb{E}_{t_i}[\overline{Y}^{\pi}_{t_{i+1}}(W_{t_{i+1}} - W_{t_i})],$$

by

$$\begin{aligned}
\overline{Y}^{\pi}_{t_i} &:= \mathbb{E}_{t_i}[g(X^{\pi}_{t_n}) + \frac{1}{n}\sum_{k=i}^{n-1} f(X^{\pi}_{t_k}, \overline{Y}^{\pi}_{t_{k+1}}, \overline{Z}^{\pi}_{t_k})] \\
\overline{Z}^{\pi}_{t_i} &:= n\mathbb{E}_{t_i}[\{g(X^{\pi}_{t_n}) + \frac{1}{n}\sum_{k=i+1}^{n-1} f(X^{\pi}_{t_k}, \overline{Y}^{\pi}_{t_{k+1}}, \overline{Z}^{\pi}_{t_k})\}(W_{t_{i+1}} - W_{t_i})].
\end{aligned}$$

In terms of discrete time approximation, it is equivalent to the forward version of our scheme, the convergence is not impacted as $|\pi| \to 0$.

However, it is shown in Gobet & Turkedjiev [32] that it provides a better control on the conditional expectations estimation error : this procedure avoids the propagation of the error estimation along the backward algorithm.

4.5 Parallelized algorithm

A version of the algorithm that can be parallelized on GPU has been proposed in Gobet et al. [30]:

- Fix hyper-cubes $(H_k)_{k \leq K}$.
- For each k, simulate N path of X^π on $[t_i, 1]$ starting from an iid drawn point in H_k according to a conditional logistic distribution.
- Backward induction: For each $j = n-1, \ldots, 0$, use the MDP scheme to compute $\overline{Y}^\pi_{t_i}$ given that $X^\pi \in H_k$ by using the simulated path of X^π on $[t_i, 1]$ starting from H_k and the knowledge of the estimated functional form of $(\overline{Y}^\pi_{t_j})_{i<j\leq n}$.

Each hyper-cube can be treated in parallel at each backward step.

5 Quality test

5.1 The case of American options

For American options, the algorithm becomes:

Initialization : For all j : $\hat{Y}_1^{\pi^{(j)}} = g\left(X_1^{\pi^{(j)}}\right)$.

Backward induction: For $i = n-1, \ldots, 1$, we set, for all j :

$$\hat{Y}_{t_i}^{\pi^{(j)}} = \max\{g(X_{t_i}^{\pi^{(j)}}), \hat{E}\left[\hat{Y}_{t_{i+1}}^{\pi} \mid X_{t_i}^{\pi} = X_{t_i}^{\pi^{(j)}}\right]\}$$

Because of the max operator, one can expect that the estimator is upper-biased $\hat{Y}_0^{\pi^{(1)}} \geq Y_0^{\Re}$, where $Y^{\Re}$ is the price of the Bermudean option with discrete exercise times.

Alternatively, one can use:

Initialization : For all j : $\hat{Y}_1^{\pi^{(j)}} = g\left(X_1^{\pi^{(j)}}\right)$ and $\hat{\tau}^{(j)} = T$.

Backward induction: For $i = n-1, \ldots, 1$, we set, for all j :

- Case 1: If $g(X_{t_i}^{\pi^{(j)}}) < \hat{E}[g(X_{\hat{\tau}}^{\pi}) \mid X_{t_i}^{\pi} = X_{t_i}^{\pi^{(j)}}]$ then $\hat{\tau}^{(j)} \leftarrow \hat{\tau}^{(j)}$.
- Case 2: If $g(X_{t_i}^{\pi^{(j)}}) \geq \hat{E}[g(X_{\hat{\tau}}^{\pi}) \mid X_{t_i}^{\pi} = X_{t_i}^{\pi^{(j)}}]$ then $\hat{\tau}^{(j)} \leftarrow t_i$.

This provides an estimate of the optimal stopping policy on the different path. Then, one computes

$$\tilde{Y}_0^{\pi} := \frac{1}{M} \sum_{j=1}^{M} g(X_{\hat{\tau}^{(j)}}^{\pi^{(j)}})$$

Because $\hat{\tau}$ is suboptimal, one expects that $\tilde{Y}_0^{\pi} \leq Y^{\Re}$.

Quality test: $\tilde{Y}_0^{\pi}$ should be less but close to $\hat{Y}_0^{\pi^{(1)}}$.

5.2 Mean L^2-discrete trajectory error

This was proposed by Bender & Steiner [3] in the context of the non-parametric approach.

It consists in computing:

$$\frac{1}{M}\sum_{j=1}^{M}\max_{k\le n}|\sum_{i=0}^{k-1}\hat{Y}^{\pi^{(j)}}_{t_{i+1}}-\hat{Y}^{\pi^{(j)}}_{t_i}-\frac{1}{n}f(X^{\pi^{(j)}}_{t_i},\hat{Y}^{\pi^{(j)}}_{t_i},\hat{Z}^{\pi^{(j)}}_{t_i})-\hat{Z}^{\pi^{(j)}}_{t_i}(W^{(j)}_{t_{i+1}}-W^{(j)}_{t_i})|^2.$$

By the law of large number, it converges to

$$\mathrm{Err}^2_\psi:=\max_{k\le n}\mathbb{E}[|\sum_{i=0}^{k-1}\hat{Y}^{\pi}_{t_{i+1}}-\hat{Y}^{\pi}_{t_i}-\frac{1}{n}f(X^{\pi}_{t_i},\hat{Y}^{\pi}_{t_i},\hat{Z}^{\pi}_{t_i})-\hat{Z}^{\pi}_{t_i}(W_{t_{i+1}}-W_{t_i})|^2].$$

If the above is computed with the true coefficients corresponding to the projection on the basis $(\psi_{t_i,m})_{m\le M}$, then Err_ψ provides, up to an additional $O(|\pi|)$ term, an upper-bound of the error due to replacing the true conditional expectations by the corresponding projections.

In practice, one estimates the coefficients with a bunch of simulations, and then use independent simulations to compute the above error criteria.

6 More...

One can also use more "deterministic" techniques:

- The quantization approach: Bally, Pagès & Printems [2], Bronstein, Pagès & Portès [13].
- The cubature approach: Crisan & Manolarakis [22].

References

[1] V. Bally and G. Pages. Error analysis of the optimal quantization algorithm for obstacle problems. *Stochastic processes and their applications*, 106(1):1–40, 2003.

[2] V Bally, G Pages, and J Printems. A quantization metod for pricing and hedging multi-dimensional american style options. *Mathematical Finance*, 15(1):119–168, 2005.

[3] C. Bender and J. Steiner. Error criteria for numerical solutions of backward sdes. *Preprint*, 42, 2010.

[4] C. Bender and J. Zhang. Time discretization and markovian iteration for coupled fbsdes. *The Annals of Applied Probability*, 18(1):143–177, 2008.

[5] B. Bouchard. Lecture notes on bsdes main existence and stability results. *Hal preprint hal-01158912*, 2015.

[6] B. Bouchard, I. Ekeland, and N. Touzi. On the malliavin approach to monte carlo approximation of conditional expectations. *Finance and Stochastics*, 8(1):45–71, 2004.

[7] B. Bouchard and R. Elie. Discrete-time approximation of decoupled forward–backward sde with jumps. *Stochastic Processes and their Applications*, 118(1):53–75, 2008.

[8] B. Bouchard, R. Elie, and N. Touzi. Discrete-time approximation of bsdes and probabilistic schemes for fully nonlinear pdes. *Advanced Financial Modelling. Radon Ser. Comput. Appl. Math*, 8:91–124, 2009.

[9] B. Bouchard, S. Geiss, and E. Gobet. First time to exit of a continuous itô process: general moment estimates and l1-convergence rate for discrete time approximations. *arXiv preprint arXiv:1307.4247*, 2013.

[10] B. Bouchard and S. Menozzi. Strong approximations of bsdes in a domain. *Bernoulli*, 15(4):1117–1147, 2009.

[11] B. Bouchard and N. Touzi. Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.

[12] B. Bouchard and X. Warin. Monte-carlo valuation of american options: facts and new algorithms to improve existing methods. In *Numerical methods in finance*, pages 215–255. Springer, 2012.

[13] A. L. Bronstein, G. Pagès, and J. Portès. Multi-asset american options and parallel quantization. *Methodology and Computing in Applied Probability*, 15(3):547–561, 2013.

[14] Jacques F Carriere. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: mathematics and Economics*, 19(1):19–30, 1996.

[15] J.-F. Chassagneux. A discrete-time approximation for doubly reflected bsdes. *Advances in Applied Probability*, 41(1):101–130, 2009.

[16] J.-F. Chassagneux. Linear multistep schemes for bsdes. *SIAM Journal on Numerical Analysis*, 52(6):2815–2836, 2014.

[17] J.-F. Chassagneux and D. Crisan. Runge–kutta schemes for backward stochastic differential equations. *The Annals of Applied Probability*, 24(2):679–720, 2014.

[18] J.-F. Chassagneux, R. Elie, and I. Kharroubi. Discrete-time approximation of multidimensional bsdes with oblique reflections. *The Annals of Applied Probability*, 22(3):971–1007, 2012.

[19] J.-F. Chassagneux and A. Richou. Numerical simulation of quadratic bsdes. *The Annals of Applied Probability*, 26(1):262–304, 2016.

[20] E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression method for american option pricing. *Finance and Stochastics*, 6(4):449–471, 2002.

[21] Michael G Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User?s guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

[22] D. Crisan and K. Manolarakis. Second order discretization of backward sdes and simulation with the cubature method. *The Annals of Applied Probability*, 24(2):652–678, 2014.

[23] D. Crisan, K. Manolarakis, and N. Touzi. On the monte carlo simulation of bsdes: An improvement on the malliavin weights. *Stochastic Processes and their Applications*, 120(7):1133–1158, 2010.

[24] F Delarue and S Menozzi. A forward-backward algorithm for quasi-linear pde. (6):140–184.

[25] N. El Karoui, S. Hamadène, and A. Matoussi. Backward stochastic differential equations and applications. *Indifference pricing: theory and applications*, pages 267–320, 2008.

[26] N. El Karoui and L. Mazliak. *Backward stochastic differential equations*, volume 364. CRC Press, 1997.

[27] A. Fahim, N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic pdes. *The Annals of Applied Probability*, pages 1322–1364, 2011.

[28] C. Geiss, S. Geiss, and E. Gobet. Generalized fractional smoothness and lp-variation of bsdes with non-lipschitz terminal condition. *Stochastic Processes and their Applications*, 122(5):2078–2116, 2012.

[29] E. Gobet and C. Labart. Error expansion for the discretization of backward stochastic differential equations. *Stochastic processes and their applications*, 117(7):803–829, 2007.

[30] E. Gobet, J. Lopez-Salas, P. Turkedjiev, and C. Vasquez. Stratified regression monte-carlo scheme for semilinear pdes and bsdes with large scale parallelization on gpus. *preprint*, 2015.

[31] E. Gobet and A. Makhlouf. Time regularity of bsdes with irregular terminal functions. *Stochastic Processes and their Applications*, 120(7):1105–1132, 2010.

[32] E. Gobet and P. Turkedjiev. Linear regression mdp scheme for discrete backward stochastic differential equations under general conditions. *Mathematics of Computation*, 85(299):1359–1391, 2016.

[33] J.-P. Lemor, E. Gobet, and X. Warin. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, 12(5):889–916, 2006.

[34] F. A. Longstaff and E. S. Schwartz. Valuing american options by simulation: a simple least-squares approach. *Review of Financial studies*, 14(1):113–147, 2001.

[35] J. Ma and J. Zhang. Path regularity for solutions of backward stochastic differential equations. *Probability Theory and Related Fields*, 122(2):163–190, 2002.

[36] J. Ma and J. Zhang. Representations and regularities for solutions to bsdes with reflections. *Stochastic processes and their applications*, 115(4):539–569, 2005.

[37] E. Pardoux, F. Pradeilles, and Z. Rao. Probabilistic interpretation of a system of semi-linear parabolic partial differential equations. In *Annales de l'IHP Probabilités et statistiques*, volume 33, pages 467–490, 1997.

[38] S. Peng. Backward stochastic differential equation, nonlinear expectation and their applications. In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 393–432, 2011.

[39] J. Zhang. A numerical scheme for bsdes. *The Annals of Applied Probability*, 14(1):459–488, 2004.