Monte Carlo Methods for Nonlinear PDEs

Arash Fahim, University of Michigan
Joint work with Nizar Touzi, Xavier Warin
Joint work with Erhan Bayraktar
Sketch of the presentation

1. Motivations
2. Linear and Nonlinear Monte Carlo Methods
3. Fully nonlinear Monte Carlo
Terminology I

- \((\Omega, \{F_t\}, \mathbb{P})\): Filtered probability space, \(\mathbb{Q}\): the martingale measure, \(\mathbb{E}\) or \(\tilde{\mathbb{E}}\): the expectation, \(\{W_t\}_{t \geq 0}\), \(d\)-dimensional Brownian motion

- Risky assets \((S_t^{(i)})_{i=1}^d\)

- Money market account with interest rate = \(r_t\).

- \(\theta\): trading strategy

- \(X^\theta_t\): Wealth process at time \(t\) based on the self–financing strategy \(\theta\)

- \(\mathbb{E}_{t,s} = \mathbb{E}[\cdot | S_t = s]\)
Motivations

Pricing American option

Markov derivatives

Pay–off: $\phi(S_T)$
Price at time $t$: $V(t, s) = \sup_T \hat{E}_{t, s}[e^{-\int_t^T r_s ds} \phi(S_T)]$.

PDE

$0 = \min\{-V_t - rsDV - s^2 \sigma^2 D^2 V + rV, v - \phi\}$ and $V(T, \cdot) = \phi(\cdot)$.
$\Delta$-Hedging: $\theta_t = DV(t, S_t)$ for $t < \tau$.

Longstaff-Schwartz

No analytical solution for the PDE in higher dimensions:

$\hat{V}(t_k, s) := \max\{\hat{E}_{t, s}[e^{-\int_{t_k}^{t_{k+1}} r_s ds} \hat{V}(S_{t_{k+1}})], \phi(s)\}$.

In the above, $\hat{E}_{t, s}$ is approximated by projection on a set of polynomials, using only one set of sample paths.
Monte Carlo Hedging and Greeks

Euler approximation of $S_t$: $S_h = \sigma s W_h$ and $\Delta W_h := W_{t+h} - W_t$:

$$
\begin{align*}
\Delta_t(s) & \approx \frac{1}{\sigma s h} \tilde{E}_{t,s}[V(t, s + \sigma s W_h)(\Delta W_h)] \\
\Gamma_t(s) & \approx \frac{1}{\sigma^2 s^2 h^2} \tilde{E}_{t,s}[V(t, s + \sigma s W_h)((\Delta W_h)^2 - h)].
\end{align*}
$$
Motivations

Portfolio constraint I

Interest rate spread ($R > r$)

\[
0 = -V_t - rsDV - s^2 \sigma^2 D^2 V + rV + (R - r)(V - sDV) \\
V(T, \cdot) = \phi(\cdot).
\]

Semi-linear parabolic PDE.

Super hedging under $\Gamma$–constraint

\[
0 = \min \left\{ -V_t - rsDV - s^2 \sigma^2 D^2 V + rV, \Gamma^* - s^2 D^2 V, s^2 D^2 V - \Gamma^* \right\} \\
V(T, \cdot) = \phi(\cdot).
\]

Fully non-linear parabolic PDE.
Indifference pricing I

Expected utility maximization
Let $U$ be a utility function.

$$v_0 := \sup_{\theta \in \mathcal{A}} \mathbb{E} \left[ U \left( X^\theta_T \right) \right].$$

General framework
- Assign a diffusion model to the price of each risky assets
- Change the utility maximization problem into a Hamilton–Jacobi–Bellman PDE
- Solving the PDE!
Indifference pricing II

B–S Model
Let $W$ be a $d$–dimensional BM ($r = 0$).

\[
dS_t = \text{diag}(S_t)(\mu dt + \sigma dW_t) \text{ where } \mu \in \mathbb{R}^d \text{ and } \sigma \in \mathbb{R}^{d \times d}.
\]

\[
dx^\theta_t = \theta \cdot (\mu dt + \sigma dW_t)
\]

Utility Maximization

\[
v(t, x) := \sup_{\theta_s \geq t} \mathbb{E} \left[ U(X^\theta_T) \mid X^\theta_t = x \right]
\]
Motivations

Indifference pricing III

HJB equation

\[ v(T, x) = U(x) \]

\[ 0 = -v_t - \sup_{\theta \in \mathbb{R}} \left( \frac{1}{2} \sigma^2 v_{xx} + \theta \mu v_x \right) \]

\[ = -v_t + \frac{1}{2} \mu^t (\sigma^t)^{-1} \mu \left( \frac{v_x}{v_{xx}} \right)^2. \]

This PDE is fully non-linear.
For exponential utility the solution can be find analytically.
The dimension of the equation does not increase with the number of assets.
Indifference pricing IV

Heston model

\[ dS_t = \mu S_t dt + \sqrt{Y_t} S_t dW_t^{(1)} \]
\[ dY_t = k(m - Y_t) dt + c \sqrt{Y_t} \left( \rho dW_t^{(1)} + \sqrt{1 - \rho^2} dW_t^{(2)} \right), \]
Indifference pricing V

HJB equation

\[
\nu(T, x, y) = U(x)
\]

\[
0 = -v_t - k(m - y)v_y - \frac{1}{2} c^2 y v_{yy} - \sup_{\theta \in \mathbb{R}} \left( \frac{1}{2} \theta^2 y v_{xx} + \theta (\mu v_x + \rho c y v_{xy}) \right)
\]

\[
= -v_t - k(m - y)v_y - \frac{1}{2} c^2 y v_{yy} + \frac{(\mu v_x + \rho c y v_{xy})^2}{2 y v_{xx}}.
\]

The dimension of fully non–linear PDE do increases with the number of stochastic volatility models.
Indifference pricing VI

Vasicek, Heston and CEV-SV models

\[
\begin{align*}
    dr_t &= \kappa (b - r_t) dt + \zeta dW_t^{(0)} \\
    dS_t^{(i)} &= \mu_i S_t^{(i)} dt + \sigma_i \sqrt{Y_t^{(i)}} S_t^{(i)} \beta_i dW_t^{(i,1)}, \quad \beta_2 = 1, \\
    dY_t^{(i)} &= k_i \left( m_i - Y_t^{(i)} \right) dt + c_i \sqrt{Y_t^{(i)}} dW_t^{(i,2)}. 
\end{align*}
\]
Indifference pricing VII

HJB equation

\[ U(x) = v(T, r, x, s_1, y_1, y_2) \]

\[ 0 = -v_t - (L^r + L^y + L^{S^1})v - rxv_x \]

\[ + \frac{((\mu_1 - r)v_x + \sigma_1^2 y_1 s_1^{2\beta_1 - 1} v_{xs_1})^2}{2\sigma_1^2 y_1 s_1^{2\beta_1 - 2} v_{xx}} + \frac{((\mu_2 - r)v_x)^2}{2\sigma_2^2 y_2 v_{xx}} \]

\[ L^r v = \kappa (b - r)v_r + \frac{1}{2} \zeta^2 v_{rr}, \quad L^y v = \sum_{i=1}^{2} k_i (m_i - y_i) v_{y_i} + \frac{1}{2} c_i^2 y_i v_{y_i y_i}, \]

and \[ L^{S^1} v = \mu_1 s_1 v_{s_1} - \frac{1}{2} \sigma_1^2 s_1 y_1 v_{s_1 s_1}. \]
The curse of dimensionality

PDEs appear in many areas including finance, image processing,... The analytic solutions usually refuse to exist and we need to approximate the solution. The deterministic approximation methods like FD, FEM, ... are highly sensitive w.r.t. dimension of the space so that they result non efficient algorithms in dimensions $d > 3$. However, the Monte Carlo scheme is less sensitive to dimension and could be used to develop numerical schemes.
Fully nonlinear Parabolic PDEs I

General form

\[-\partial_t v - \bar{F}(t, x, v(t, x), Dv(t, x), D^2v(t, x)) = 0, [0, T) \times \mathbb{R}^d\]

\[v(T, \cdot) = g.\]

Definition:

- Parabolicity means \(\bar{F}(t, x, r, p, \gamma)\) is increasing with respect to \(\gamma\).
- Fully non-linear is due to dependence of non-linearity to the second derivative.
Fully nonlinear Parabolic PDEs II

Separation into linear and fully nonlinear parts

\[-\mathcal{L}^X v - F(t, x, v(t, x), Dv(t, x), D^2v(t, x)) = 0, [0, T) \times \mathbb{R}^d\]

\[v(T, \cdot) = g.\]

where \(\mathcal{L}^X \varphi := \frac{\partial \varphi}{\partial t} + \mu \cdot D\varphi + \frac{1}{2} \sigma^T \sigma \cdot D^2\varphi\) is the infinitesimal generator of

\[dX_t = \mu dt + \sigma dW_t\]

and \(\mathcal{L}^X + F = \frac{\partial}{\partial t} + \bar{F}\) and \(F\) is still parabolic.

Choice of \(\mu\) and \(\sigma\)

\[u_t + \frac{1}{2} \Delta u = \{u_t + \frac{1}{4} \Delta u\} + \{\frac{1}{4} \Delta u\} = \{u_t + \frac{1}{8} \Delta u\} + \{\frac{3}{8} \Delta u\} \text{ but not}\]

\[\{u_t + \frac{3}{4} \Delta u\} + \{\frac{-1}{4} \Delta u\}\]
A backward numerical scheme

\[ h = \frac{T}{n} \text{ and } t_i = ih. \hat{X}_h \text{ is the Euler discretization of } X. \]

\[
\hat{v}(T, x) = g(x) \\
\hat{v}(t_i, x) = \mathbb{E}_{t_i, x}[\hat{v}(t_{i+1}, \hat{X}_h^x)] \\
+ hF \left( t_i, x, \hat{v}(t_{i+1}, x), \hat{D}\hat{v}(t_{i+1}, x), \hat{D}^2\hat{v}(t_{i+1}, x) \right)
\]

\( \hat{D}^i \) is the approximation of derivatives:

\[
\hat{D}\hat{v}(t_{i+1}, x) = \mathbb{E}_{t_i, x}[D\hat{v}(t_{i+1}, \hat{X}_h^x)(t_k, \hat{X}_h^x)] \\
\hat{D}^2\hat{v}(t_{i+1}, x) = \mathbb{E}_{t_i, x}[D^2\hat{v}(t_{i+1}, \hat{X}_h^x)(t_k, \hat{X}_h^x)]
\]
Key Lemma: Integration by part

For every exponentially bounded smooth function \( \varphi : \mathbb{R}^d \to \mathbb{R} \), we have:

\[
\mathbb{E}[D^i \varphi(x + \sigma W_h)] = \mathbb{E}[(x + \sigma W_h) H^h_i(W_h)],
\]

where \( H_h = (H^h_0, H^h_1, H^h_2) \) and

\[
H^h_0(x) = 1, \quad H^h_1(x) = \frac{1}{h} \sigma'(x)^{-1} W_h, \\
H^h_2(x) = \frac{1}{h^2} \sigma'(x)^{-1} (W_h W'_h - h I_{d \times d}) \sigma(x)^{-1}
\]

One dimensional case
Fully non-linear parabolic PDEs

Similarity with Finite Difference

\[ \mathbb{E}[\psi^{(i)}(x + \sigma W_h)] = \mathbb{E}[\psi(x + \sigma W_h)H^h_i(W_h)] \]

First derivative

- \( W_h \approx \sqrt{h}X \) where \( X \) takes \( \pm 1 \) with probability \( \frac{1}{2} \)
- \( \mathbb{E}[\psi(x + \sigma W_h)H^h_1(W_h)] \approx \frac{\psi(x+\sigma\sqrt{h})-\psi(x-\sigma\sqrt{h})}{2\sigma\sqrt{h}} \approx \psi'(x) \)

Second derivative

- \( W_h \approx \sqrt{3h}X \) where \( X \) takes \( \pm 1 \) and 0 with probability \( \frac{1}{6} \) and \( \frac{2}{3} \), resp.
- \( \mathbb{E}[\psi(x + \sigma W_h)H^h_2(W_h)] \approx \frac{\psi(x+\sigma\sqrt{3h})+\psi(x-\sigma\sqrt{3h})-\psi(x)}{3\sigma^2 h^2} \approx \psi''(x) \)
Fully non-linear parabolic PDEs

Similarity with Finite Difference

\[ \mathbb{E}[\psi^{(i)}(x + \sigma W_h)] = \mathbb{E}[\psi(x + \sigma W_h)H_i^h(W_h)] \]

First derivative

- \( W_h \approx \sqrt{h}X \) where \( X \) takes ±1 with probability \( \frac{1}{2} \)
- \( \mathbb{E}[\psi(x + \sigma W_h)H_1^h(W_h)] \approx \frac{\psi(x+\sigma\sqrt{h}) - \psi(x-\sigma\sqrt{h})}{2\sigma\sqrt{h}} \approx \psi'(x) \)

Second derivative

- \( W_h \approx \sqrt{3h}X \) where \( X \) takes ±1 and 0 with probability \( \frac{1}{6} \) and \( \frac{2}{3} \), resp.
- \( \mathbb{E}[\psi(x + \sigma W_h)H_2^h(W_h)] \approx \frac{\psi(x+\sigma\sqrt{3h}) + \psi(x-\sigma\sqrt{3h}) - \psi(x)}{3\sigma^2 h^2} \approx \psi''(x) \)
Probabilistic interpretation for Parabolic PDEs

Why Feynman-Kac doesn’t work for nonlinear PDEs I

Linear PDEs

\[ 0 = -v_t - \mathcal{L}_X v + kv \quad \text{and} \quad v(T, \cdot) = g(\cdot). \]

\[ \mathcal{L}_X = \frac{1}{2} \sigma \sigma^T \cdot D^2 v + mu \cdot Dv. \]

\[ v(t, x) = \mathbb{E} \left[ \exp \left( - \int_t^T k(X_s) ds \right) g(X_T) \bigg| X_t = x \right]. \]

where \( dX_t = \mu dt + \sigma \cdot dW_t. \)
Probabilistic interpretation for Parabolic PDEs

Why Feynman-Kac doesn’t work for nonlinear PDEs II

Totter–Kato

\[ \Phi_h^{(1)}[g](t, x) = \mathbb{E}_{t,x}[g(X_{t+h})] \] the semi–group generated by \( 0 = -\nu_t - \mathcal{L}X \nu \) and

\[ \Phi_h^{(2)}[g](t, x) = \exp(-\int_t^{t+h} k_s ds) g(x) \] the semi–group generated by \( 0 = -\nu_t + k \nu \). Then, \((h \to 0)\)

\[ \nu(t, x) \approx \Phi_h^{(1)} \circ \Phi_h^{(2)} \circ \ldots \circ \Phi_h^{(1)} \circ \Phi_h^{(2)}[g](t, x) \]

Hopefully, two semi–groups commute:

\[ \nu(t, x) \approx \Phi_{T-t}^{(1)} \circ \Phi_{T-t}^{(2)}[g](t, x) \]

\[ \to \mathbb{E}_{t,x} \left[ \exp \left( - \int_t^{T} k(X_s) ds \right) g(X_T) \right]. \]

But, when the equation is non–linear, they don’t commute.
Semi–linear PDEs

**Semi–linear PDEs I**

Semi–linear equations

\[
0 = -v_t - \frac{1}{2} \sigma^T \sigma \cdot D^2 v - \mu \cdot Dv - F(v, Dv)
\]

\[
v(T, \cdot) = g(\cdot).
\]

Possibly no classic solution. The solution should be considered in viscosity sense.

**Example:** Interest rate spread.
Semi–linear PDEs II

Backward Stochastic Differential Equations

The linear part gives us a diffusion process $dX_t = \mu dt + \sigma dW_t$.

\[
\begin{align*}
    dY_t &= F(Y_t, Z_t)dt - Z_t dX_t \\
    Y_T &= g(X_T).
\end{align*}
\]

Relation with PDE

If $\nu$ is the classical solution of semi–linear PDE, $Y_t = \nu(t, X_t)$, $Z_t = D\nu(t, X_t)$.

Theory: [Bismut 78], [El Karoui–Peng–Quenez 97], [Pardoux–Peng92]
Semi-linear PDEs III

Discretization of BSDE

\[ \hat{Y}_i = \mathbb{E}_i[\hat{Y}_{i+1} + hF(t_i, \hat{X}_i, \hat{Y}_i, \hat{Z}_i)] \quad \hat{Y}_T = g(X_T) \]

\[ \hat{Z}_i = \frac{1}{h} \mathbb{E}_i[\hat{Y}_{i+1} \Delta W_{i+1}] \approx \Delta_t. \]

Discretization and numerical aspects: [Touzi–Bouchard 03], [Zhang 03], [Gober–Lemor–Warin 04]
Fully non–linear PDEs

Fully non–linear equations

\[
0 = -\frac{\partial v}{\partial t} - \bar{F}(t, x, v, Dv, D^2v)
\]

\[
v(T, \cdot) = g(\cdot).
\]

Parabolicity: \( \bar{F}(t, x, r, p, \gamma) \) is increasing with respect to \( \gamma \).

The solution definitely should be considered in viscosity sense.

Application: Merton portfolio selection model, Super–hedging under \( \Gamma \) constrain

No Monte Carlo method is known for the general above type.
Fully non–linear PDEs II

2BSDE

\[
\begin{align*}
dY_t &= F(Y_t, Z_t, \Gamma_t)dt - Z_t dX_t \\
dZ_t &= A_t dt + \Gamma_t dX_t \\
Y_T &= g(X_T).
\end{align*}
\]

If \( v \) is the classical solution of semi–linear PDE, \( Y_t = v(t, X_t) \), \( Z_t = Dv(t, X_t) \), \( \Gamma_t = D^2v(t, X_t) \), \( A_t = \mathcal{L}^X Dv(t, X_t) \).

Discretization of 2BSDE

\[
\begin{align*}
\hat{Y}_i &= \mathbb{E}_i[\hat{Y}_{i+1} + hF(t_i, \hat{X}_i, \hat{Y}_i, \hat{Z}_i, \hat{\Gamma}_i)] \quad \hat{Y}_T = g(X_T) \\
\hat{Z}_i &= \frac{1}{h}\mathbb{E}_i[\hat{Y}_{i+1} \Delta W_{i+1}] \quad \hat{Z}_T = Dg(X_T) \\
\hat{\Gamma}_i &= \frac{1}{h}\mathbb{E}_i[\hat{Z}_{i+1} \Delta W_{i+1}]
\end{align*}
\]

Alternative scheme.
Viscosity solution

An upper-semicontinuous (resp. lower semicontinuous) function $v$ (resp. $\overline{v}$) on $[0, T] \times \mathbb{R}^d$, is called a viscosity subsolution (resp. supersolution) of the PDE if for any $(t, x) \in [0, T] \times \mathbb{R}^d$ and any smooth function $\varphi$ satisfying

$$0 = (v - \varphi)(t, x) = \max_{[0, T] \times \mathbb{R}^d} (v - \varphi) \quad \text{(resp.} \quad 0 = (\overline{v} - \varphi)(t, x) = \min_{[0, T] \times \mathbb{R}^d} (\overline{v} - \varphi) \text{)},$$

we have: $-\mathcal{L}^X \varphi - F(t, x, D\varphi(t, x)) \leq (\text{resp.} \geq)0$.
Main results (F–Touzi–Warin) II

Comparison principle

We say that fully non-linear equation has comparison for bounded functions if for any bounded upper semicontinuous subsolution \( \underline{v} \) and any bounded lower semicontinuous supersolution \( \overline{v} \) on \([0, T) \times \mathbb{R}^d\), satisfying \( \underline{v}(T, \cdot) \leq \overline{v}(T, \cdot) \), we have \( \underline{v} \leq \overline{v} \).

Assumption F

1. \( F \) is Lipschitz-continuous with respect to \((x, r, p, \gamma)\) uniformly in \(t\).
2. \( |F(\cdot, \cdot, 0, 0, 0)|_\infty < \infty \).
3. \( F \) is elliptic (increasing w.r.t. \( \gamma \)).
4. \( \nabla_\gamma F.a^{-1} \leq 1 \) where \( a := \sigma' \sigma \) on \( \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d \).
5. \( F_p \in \text{Image}(F_\gamma) \) and \( |F_p F_\gamma F_p|_\infty < +\infty \).
Convergence Thm: F–Touzi–Warin

Assume $F$ and comparison for the PDE. For every bounded Lipschitz function $g$, there exists a bounded function $v$ so that

$$v^h \longrightarrow v$$

locally uniformly.

In addition, $v$ is the unique bounded viscosity solution of fully non-linear problem.

Proof of Convergence

The proof of convergence relies on the method of Barles and Souganidis for viscosity solutions (Not directly applicable).
Asymptotics

Main results (F–Touzi–Warin) IV

Monotonicity

Should be: If $\varphi \leq \psi$ then $T_h[\varphi] \leq T_h[\psi]$.
But it is: If $\varphi \leq \psi$ then $T_h[\varphi] \leq T_h[\psi] + C h^\varepsilon [h](\psi - \varphi)(t + h, \hat{X}_h^x)$.

Stability

The family $\{v^h\}$ is uniformly bounded.

Consistency

When $h \to 0$, $c \to 0$ and $(t', x') \to (t, x)$:

$$\frac{1}{h}(\psi + c - T_h[\psi + c])(t', x') \to -v_t - \mathcal{L}^X v - F(t, x, \mathcal{D}v(t, x)).$$
Main results (F–Touzi–Warin) V

Final condition

When $h \to 0$ and $(t', x') \to (T, x)$: $v^h(t', x') \to g(x)$. (This result is neither necessary for nor provided by Barles–Souganidis but very crucial in this context, because of the form of the equation.)

Regularity of approximate solution

$v^h$ is Lipschitz in $x$ and $\frac{1}{2}$-Hölder on $t$. What we need for above result is the later. In F-Touzi-Warin, we used $x$ Lipschitz continuity to show $t \frac{1}{2}$-Hölder continuity. Later on, this step was skipped in the future work on PDEs on general domains.
Rate of convergence I

**HJB (convexity)**

The nonlinearity $F$ satisfies Assumption **F3–5**, and is of the Hamilton-Jacobi-Bellman type:

\[
\frac{1}{2} a \cdot \gamma + b \cdot p + F(t, x, r, p, \gamma) = \inf_{\alpha \in \mathcal{A}} \{ \mathcal{L}^\alpha(t, x, r, p, \gamma) \}
\]

\[
\mathcal{L}^\alpha(t, x, r, p, \gamma) := \frac{1}{2} \operatorname{Tr}[\sigma^\alpha \sigma^\alpha^\top(t, x) \gamma] + b^\alpha(t, x)p + c^\alpha(t, x)r + f^\alpha(t, x)
\]

where the functions $\mu$, $a$, $\sigma^\alpha$, $b^\alpha$, $c^\alpha$ and $f^\alpha$ satisfy:

\[
|\mu|_\infty + |a|_\infty + \sup_{\alpha \in \mathcal{A}} (|\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1) < \infty.
\]
Rate of convergence II

HJB+

The nonlinearity $F$ satisfies HJB, and for any $\delta > 0$, there exists a finite set \( \{\alpha_i\}_{i=1}^{M_\delta} \) such that for any $\alpha \in \mathcal{A}$:

$$\inf_{1 \leq i \leq M_\delta} \left| \sigma^\alpha - \sigma^{\alpha_i} \right|_\infty + \left| b^\alpha - b^{\alpha_i} \right|_\infty + \left| c^\alpha - c^{\alpha_i} \right|_\infty + \left| f^\alpha - f^{\alpha_i} \right|_\infty \leq \delta.$$ 

Thm: F–Touzi–Warin

Assume that $g$ is Lipschitz and let HJB and HJB+

$$-Ch^{1/10} \leq v - v^h \leq Ch^{1/4}.$$ 

Rate of conv.

The proof of rate of convergence is obtained through Krylov, Barles and Jakobsen method of shaking coefficients and switching system approximation of Barles and Jakobsen (Not directly applicable).
Asymptotics

Rate of convergence III

Consistency estimate

By HJB and HJB+, \( v \), the solution of PDEs, is unique (in a suitable class) and is Lipschitz in \( x \) and \( \frac{1}{2} \)-Hölder on \( t \). There exists a smooth sub solution \( v_{\varepsilon} \) and a smooth super solution \( v^{\varepsilon} \) for the PDE with the properties:

1) \( |v_{\varepsilon} - v| \leq C\varepsilon \) (based on convexity) and \( |v^{\varepsilon} - v| \leq C\varepsilon^{\frac{1}{3}} \) (not optimal).
2) \( \left| \frac{\partial^k}{\partial t^k} D^k v_{\varepsilon} \right| \leq C\varepsilon^{2-2k-|l|} \).

\[ |PDE(\phi) - h^{-1}\text{scheme}(\phi)| \leq C\varepsilon^{-3}h \]

Comparison for scheme

If \( h^{-1}\text{scheme}(\phi) \geq g \) and \( h^{-1}\text{scheme}(\psi) \leq h \), then \( \phi - \psi \leq (g - h)_+ \).

RHS bound

\[ |\hat{v} - v| \leq |\hat{v} - v_{\varepsilon}| + |v_{\varepsilon} - v| \leq C\varepsilon^{-3}h + C\varepsilon \]
Implementation

Backward implementation I

Scheme reminded

\[ \nu^h(T, .) = g \quad \text{and} \quad \nu^h(t, x) = T_h[\nu^h(t_{i+1}, .)](x). \]

\[ T_h \psi(x) := \mathbb{E}[\psi(\hat{X}_h^x)] + hF(x, D_h \psi(x)) \]

\[ D_h \psi := (D^0_h \psi, D^1_h \psi, D^2_h \psi) \quad D^i_h \psi(x) := \mathbb{E}[\psi(\hat{X}_h^x)H^i_h(W_h)], \ i = 0, 1, 2. \]
Implementation

1. \[ t_i = \frac{iT}{n}. \]

2. Generating \( N \) sample paths from the Euler discretization of \( X_t; \hat{X}_t \).

\[ \{(\hat{X}^{(j)}_{t_i}, \Delta W^{(j)}_{t_{i+1}}) | 0 = t_0, \ldots, t_n = T, j = 1, \ldots, N\}. \]

3. Start from terminal condition \( g(\cdot) \). And proceed backward in time.

4. Knowing \( v^{h}(t_{i+1}, \hat{X}^{(j)}_{t_{i+1}}) \)s for \( j = 1, \ldots, N \), then one calculates \( v^{h}(t_{i}, \hat{X}^{(k)}_{t_{i}}) \)s for \( k = 1, \ldots, N \) using the scheme.
Backward implementation III

\[ \hat{v}(t_{i+1}, X^2) \]
\[ \hat{v}(t_{i+1}, X^5) \]
\[ \hat{v}(t_{i+1}, X^1) \]
\[ \hat{v}(t_{i+1}, X^3) \]
\[ \hat{v}(t_{i+1}, X^4) \]

\[ t_i \]
\[ t_{i+1} \]
4th step

To compute $v^h(t_i, x)$, one needs to approximate:

\[
\mathbb{E}[v^h(t_{i+1}, \hat{X}_{t_{i+1}})|\hat{X}_{t_i} = x]
\]
\[
\mathbb{E}[v^h(t_{i+1}, \hat{X}_{t_{i+1}})\Delta W_{t_{i+1}}|\hat{X}_{t_i} = x]
\]
\[
\mathbb{E}[v^h(t_{i+1}, \hat{X}_{t_{i+1}})((\Delta W_{t_{i+1}})^2 - h)|\hat{X}_{t_i} = x]
\]
Approximation of conditional expectations I

Kernel methods

\( Y = v(t + h, \hat{X}_{t+h})H^i_h \) and \( X = \hat{X}_t \). Informally;

\[
\mathbb{E}[Y|X = x] = \frac{\mathbb{E}[Y \delta_x(X)]}{\mathbb{E}[\delta_x(X)]} \approx \frac{\mathbb{E}[Y \varepsilon_\kappa(x - X)]}{\mathbb{E}[\varepsilon_\kappa(x - X)]}
\]

where \( \varepsilon_\kappa \rightarrow \delta_0 \). In terms of a sample \( \{(X^i, Y^i)\}_{i=1}^N \);

\[
\frac{\sum_i \varepsilon_\kappa(X^i - x) Y^i}{\sum_i \varepsilon_\kappa(X^i - x)}
\]
Approximation of conditional expectations II

Projection methods

\[ Y = v(t + h, \hat{X}_{t+h}) H^i_h \] and \( X = \hat{X}_t \). Formally;

\[
\mathbb{E}[Y | X = x] \approx \sum_k c_k \psi^k(x)
\]

Coefficients \( c_k \)'s should be determined so that the \( L^2 \) approximation error be minimum.
Malliavin methods

\[ Y = \nu(t + h, \hat{X}_{t+h})H^i_h \text{ and } X = \hat{X}_t. \text{ Formally;} \]

\[
\mathbb{E}[Y|X = x] = \frac{\mathbb{E}[Y\delta_x(X)]}{\mathbb{E}[\delta_x(X)]}
\]

\[
\mathbb{E}[Y\delta_x(X)] = \mathbb{E}[Y \mathbb{1}_{\{X>x\}} \delta^t_0] \text{ where } \delta^t_0 \text{ is Skorokhod integral which depends}
\]
\[
\text{only on the path of } X \text{ from 0 to } t.
\]
Approximation of conditional expectations IV

Stochastic scheme

\[
\tilde{T}_h^N[\psi](t, x) := \tilde{E}^N \left[ \psi(t + h, \hat{X}_h^t) \right] + hF \left( \cdot, \hat{D}_h\psi \right)(t, x),
\]
\[
\hat{T}_h^N[\psi](t, x) := -K_h[\psi] \lor \tilde{T}_h^N[\psi](t, x) \land K_h[\psi]
\]

where

\[
\hat{D}_h\psi(t, x) := \tilde{E}^N \left[ \psi(t + h, \hat{X}_h^t)H_h(t, x) \right] , \quad K_h[\psi] := \|\psi\|_\infty (1 + C_1 h) + C_2 h,
\]

and

\[
C_1 = \frac{1}{4} |F_p^\top F^-_{\gamma} F_p|_\infty + |F_r|_\infty \quad \text{and} \quad C_2 = |F(t, x, 0, 0, 0)|_\infty.
\]
Approximation of conditional expectations V

Assumption \(E\)

Let \(\mathcal{R}_b\) be the family of random variables \(R\) of the form \(\psi(W_h)H_i(W_h)\) where \(\psi\) is a function by \(b\) and \(H_i\)’s are the Hermite polynomials:

\[
H_0(x) = 1, \quad H_1(x) = x \quad \text{and} \quad H_2(x) = x^T x - h \quad \forall x \in \mathbb{R}^d.
\]

There exist constants \(C_b, \lambda, \nu > 0\) such that

\[
\left| \mathbb{E}^N[R] - \mathbb{E}[R] \right|_p \leq C_b h^{-\lambda} N^{-\nu}
\]

for every \(R \in \mathcal{R}_b\), for some \(p \geq 1\).

Regression approximation based on the Malliavin integration introduced in [Lions and Reigner], [Bouchard, Ekeland and Touzi], and analyzed in the context of the simulation of backward stochastic differential equations by [Bouchard and Touzi]. Then Assumption is satisfied for every \(p > 1\) with the constants \(\lambda = \frac{d}{4p}\) and \(\nu = \frac{1}{2p}\).
Asymptotic properties I

Convergence result

Let Assumptions E and F hold true, and assume that the fully nonlinear PDE has comparison with growth $q$. Suppose in addition that

$$\lim_{h \to 0} h^{\lambda + 2} N_h^\nu = \infty.$$ 

Assume that the final condition $g$ is bounded Lipschitz, and the coefficients $\mu$ and $\sigma$ are bounded. Then, for almost every $\omega$:

$$\hat{v}_{Nh}^h(\cdot, \omega) \longrightarrow v \quad \text{locally uniformly,}$$

where $v$ is the unique viscosity solution of equation.
Rate of convergence result

Let the nonlinearity $F$ be as in Assumption $\textbf{HJB}$, and consider a regression operator satisfying Assumption $\textbf{E}$. Let the sample size $N_h$ be such that

$$\lim_{h \to 0} h^{\lambda + \frac{21}{10}} N_h^\nu > 0.$$ 

Then, for any bounded Lipschitz final condition $g$, we have the following $\mathbb{L}^p$—bounds on the rate of convergence:

$$\| v - \hat{v}^h \|_p \leq Ch^{1/10}.$$
Problem consider in numerical experiments

- Mean curvature flow of a sphere in dimension 3
- Mean curvature flow of a dumbbell shaped area in dimension 2
- Portfolio selection in dimension 2 (an asset with stochastic volatility)
- Portfolio selection in dimension 5 (stochastic interest rate and two assets with stochastic volatility)

Alternative schemes

\[ D^2 v^h(t + h, x) \approx \mathbb{E}[v^h(t + h, \hat{X}_h^x) H^1_h H^1_h] \]
Numerical experiments

3-d sphere

Mean Curvature Flow in 3D for a sphere

"analytical solution"
"Solution with step 0.00125 volatility 1"
"Solution with step 0.00125 volatility 1.8"

Radius

time

University of

hern California
2-d dumbbell
Portfolio selection in dimension 2

Heston model

\[ dS_t = \mu S_t dt + \sqrt{Y_t} S_t dW_t^{(1)} \]
\[ dY_t = k(m - Y_t) dt + c\sqrt{Y_t} \left( \rho dW_t^{(1)} + \sqrt{1 - \rho^2} dW_t^{(2)} \right) , \]

Utility maximization

\[ v_0 := \sup_{\theta \in \mathcal{A}} \mathbb{E} \left[ - \exp \left( -\eta X_T^\theta \right) \right] . \]
Portfolio selection in dimension 2 II

HJB equation

\[ v(T, x, y) = -e^{-\eta x} \]
\[ 0 = -v_t - k(m - y)v_y - \frac{1}{2} c^2 yv_{yy} \]
\[ - \sup_{\theta \in \mathbb{R}} \left( \frac{1}{2} \theta^2 yv_{xx} + \theta(\mu v_x + \rho cyv_{xy}) \right) \]
\[ = -v_t - k(m - y)v_y - \frac{1}{2} c^2 yv_{yy} + \frac{(\mu v_x + \rho cyv_{xy})^2}{2yv_{xx}}. \]

Zariphopoulou semi-explicit solution

\[ v(t, x, y) = -e^{-\eta x} \left\| \exp \left( -\frac{1}{2} \int_t^T \frac{\mu^2}{Y_s^2} ds \right) \right\|_{\mathbb{L}^{1-\rho^2}} \]
Numerical experiments

Portfolio selection in dimension 2 III

Separation into linear and fully non–linear part

\[-v_t - k(m-y)v_y - \frac{1}{2} c^2 yv_{yy} - \frac{1}{2} \sigma^2 v_{xx} + F(y, Dv, D^2 v) = 0, \quad v(T, x, y) = -e^{-\eta x},\]

\[F(y, z, \gamma) = \frac{1}{2} \sigma^2 \gamma_{11} + \frac{(\mu z_1 + \rho c y \gamma_{12})^2}{2y \gamma_{11}}.\]
Numerical experiments

Portfolio selection in dimension 2 IV

Truncation of the non-linearity

\[ F_{\varepsilon,M}(y, z, \gamma) := \frac{1}{2} \sigma^2 \gamma_{11} - \sup_{\varepsilon \leq \theta \leq M} \left( \frac{1}{2} \theta^2 (y \lor \varepsilon) \gamma_{11} + \theta (\mu z_1 + \rho c (y \lor \varepsilon) \gamma_{12}) \right), \]

Choice of diffusion

\[ dX^{(1)}_t = \sigma dW^{(1)}_t, \quad \text{and} \quad dX^{(2)}_t = k (m - X^{(2)}_t) dt + c \sqrt{X^{(2)}_t} dW^{(2)}_t. \]

\( \mu = 0.15, \ c = 0.2, \ k = 0.1, \ m = 0.3, \ Y_0 = m, \ \rho = 0, \ x_0 = 1, \ T = 1 \) then \( \nu_0 = -0.3534. \)
Numerical experiments

2 dimensional 1

Error for scheme one, financial problem one

volatility 0.6 scheme 1
volatility 1 scheme 1
volatility 1.2 scheme 1

Difference

Time step
Numerical experiments

2 dimensional 2

Error for scheme two, financial problem one

- Volatility 0.6 scheme 2
- Volatility 1 scheme 2
- Volatility 1.2 scheme 2

Difference vs. time step graph.
Numerical experiments

Portfolio selection in dimension 5

Vasicek, Heston and CEV-SV models

\[
\begin{align*}
    dr_t &= \kappa (b - r_t) dt + \zeta dW_t^{(0)} \\
    dS_t^{(i)} &= \mu_i S_t^{(i)} dt + \sigma_i \sqrt{Y_t^{(i)}} S_t^{(i)} \beta_i dW_t^{(i,1)}, \quad \beta_2 = 1, \\
    dY_t^{(i)} &= k_i \left( m_i - Y_t^{(i)} \right) dt + c_i \sqrt{Y_t^{(i)}} dW_t^{(i,2)}.
\end{align*}
\]
HJB equation

\[ 0 = -v_t - (L^r + L^Y + L^{S_1})v - rxv_x \]
\[ + \frac{((\mu_1 - r)v_x + \sigma_1^2 y_1 s_1^{2\beta_1 - 1}v_{x_{s_1}})^2}{2\sigma_1^2 y_1 s_1^{2\beta_1 - 2}v_{xx}} + \frac{((\mu_2 - r)v_x)^2}{2\sigma_2^2 y_2 v_{xx}} \]

\[ L^r v = \kappa (b - r) v_r + \frac{1}{2} \zeta^2 v_{rr}, \quad L^Y v = \sum_{i=1}^{2} k_i (m_i - y_i) v_{y_i} + \frac{1}{2} c_i^2 y_i v_{y_i y_i}, \]

and \[ L^{S_1} v = \mu_1 s_1 v_{s_1} - \frac{1}{2} \sigma_1^2 s_1 y_1 v_{s_1 s_1}. \]
Numerical experiments

Portfolio selection in dimension 5 III

Separation into linear and fully non–linear part

\[-v_t - (L^r + L^Y + L^S) v - \frac{1}{2} \sigma^2 v_{xx} + F ((x, r, s_1, y_1, y_2), Dv, D^2 v) = 0,
\]

\[v(T, x, r, s_1, y_1, y_2) = -e^{-\eta x},\]

\[F(u, z, \gamma) = \frac{1}{2}\sigma^2 \gamma_{11} - x_1 x_2 z_1 + \frac{((\mu_1 - x_2)z_1 + \sigma_1^2 x_4 x_3^{2\beta_1 - 1}\gamma_{1,3})^2}{2\sigma_1^2 x_4 x_3^{2\beta_1 - 2}\gamma_{11}}
\]

\[+ \frac{((\mu_2 - x_2)z_1)^2}{2\sigma_2^2 x_5 \gamma_{11}},\]

where \(u = (x_1, \cdots, x_5).\)
Numerical experiments

Portfolio selection in dimension 5 IV

Truncation of the non–linearity

\[
F_{\varepsilon, M}(u, z, \gamma) := \frac{1}{2} \sigma^2 \gamma_{11} - x_1 x_2 z_1 + \sup_{\varepsilon \leq |\theta| \leq M} \left\{ (\theta \cdot (\mu - r 1)) z_1 \\
+ \theta_1 \sigma_1^2 (x_4 \vee \varepsilon) (x_3 \vee \varepsilon)^{2\beta_1 - 1} \gamma_{13} \\
+ \frac{1}{2} (\theta_1^2 \sigma_1^2 (x_3 \vee \varepsilon) (x_4 \vee \varepsilon)^{2\beta_1 - 2} + \theta_2^2 \sigma_2^2 (x_5 \vee \varepsilon)) \gamma_{11} \right\},
\]
Choice of diffusion

\[
\begin{align*}
    dX_t^{(1)} &= \sigma dW_t^{(0)}, \\
    dX_t^{(2)} &= \kappa(b - X_t^{(2)}) dt + \zeta dW_t^{(1)}, \\
    dX_t^{(3)} &= \mu_1 X_t^{(3)} dt + \sigma_1 \sqrt{X_t^{(4)} X_t^{(3)}}^{\beta_1} dW_t^{(1,1)}, \\
    dX_t^{(4)} &= k_1(m_1 - X_t^{(4)}) dt + c_1 \sqrt{X_t^{(4)}} dW_t^{(1,2)}, \\
    dX_t^{(5)} &= k_2(m_2 - X_t^{(5)}) dt + c_2 \sqrt{X_t^{(5)}} dW_t^{(2,2)}. 
\end{align*}
\]
Numerical experiments

5 dimensional 1

Solution of the financial 5 dimension problem with 3 millions particles

- Value vs. time step chart showing the behavior of solutions for different volatility levels.
Numerical experiments

5 dimensional 2

Solution of the financial 5 dimension problem with 30 millions particles

- volatility 0.6
- volatility 1
- volatility 1.2

value vs. time step
Non–local Parabolic PDEs I

Fully non–linear non–local parabolic PDEs

\[-L^X v(t, x) - F(t, x, v(t, x), Dv(t, x), D^2 v(t, x), v(t, \cdot)) = 0,\]
\[v(T, \cdot) = g,\]

\[L^X \varphi(t, x) := \left( \frac{\partial \varphi}{\partial t} + \mu \cdot D\varphi + \frac{1}{2} a \cdot D^2 \varphi \right)(t, x)\]
\[+ \int_{\mathbb{R}^d} \left( \varphi(t, x + \eta(t, x, z)) - \varphi(t, x) - 1_{\{|z| \leq 1\}} D\varphi(t, x) \eta(t, x, z) \right) d\nu(z).\]
Generalizations

Parabolic PDE in more general domains I

Fully non–linear parabolic PDEs in general domains

\[-\mathcal{L}^X v(t, x) - F \left( t, x, v(t, x), Dv(t, x), D^2 v(t, x) \right) = 0, \text{ on } \mathcal{O} \]
\[v(t, \cdot) = \varphi(\cdot), \text{ on } \partial \mathcal{O},\]
\[v(T, \cdot) = g,\]
Thank you for your attention.