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

USC University of Southern California

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.
- θ: trading strategy
- X_t^θ: Wealth process at time t based on the self–financing strategy θ
- $\mathbb{E}_{t,s} = \mathbb{E}[\cdot|S_t = s]$

USC University of Southern California

Pricing American option I

Markov derivatives

Pay–off: $\phi(S_T)$
Price at time t: $V(t,s) = \sup_\tau \tilde{\mathbb{E}}_{t,s}[e^{-\int_t^\tau r_s ds}\phi(S_\tau)]$.

PDE

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

Longstaff-Schwartz

No analytical solution for the PDE in higher dimensions:
$\hat{V}(t_k, s) := \max\{\tilde{\mathbb{E}}_{t,s}[e^{-\int_{t_k}^{t_{k+1}} r_s ds}\hat{V}(S_{t_{k+1}})], \phi(s)\}$.
In the above, $\tilde{\mathbb{E}}_{t,s}$ is approximated by projection on a set of polynomials, using only one set of sample paths.

Pricing American option II

Monte Carlo Hedging and Greeks

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

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

USC University of Southern California

Portfolio constraint I

Interest rate spread ($R > r$)

$$\begin{aligned} 0 &= -V_t - rsDV - s^2\sigma^2 D^2 V + rV + (R-r)(V - sDV)_- \\ V(T,\cdot) &= \phi(\cdot). \end{aligned}$$

Semi-linear parabolic PDE.

Super hedging under $\Gamma-$constraint

$$\begin{aligned} 0 &= \min\left\{-V_t - rsDV - s^2\sigma^2 D^2 V + rV, \Gamma^* - s^2 D^2 V, s^2 D^2 V - \Gamma_*\right\} \\ V(T,\cdot) &= \phi(\cdot). \end{aligned}$$

Fully non-linear parabolic PDE.

Indifference pricing I

Expected utility maximization

Let U be a utility function.

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

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!

Southern California

Indifference pricing II

B–S Model

Let $W_\cdot$ be a d–dimensional BM ($r = 0$).

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

$$dX_t^\theta = \theta \cdot (\mu dt + \sigma dW_t)$$

Utility Maximization

$$v(t,x) := \sup_{\substack{\theta_s \\ t \le s \le T}} \mathbb{E}\left[U\left(X_T^\theta\right) | X_t^\theta = x\right]$$

Indifference pricing III

HJB equation

$$
\begin{aligned}
v(T,x) &= U(x) \\
0 &= -v_t - \sup_{\theta\in\mathbb{R}}\left(\frac{1}{2}\theta^2\sigma^2 v_{xx} + \theta\mu v_x\right) \\
&= -v_t + \frac{1}{2}\mu^{\mathrm{t}}(\sigma\sigma^{\mathrm{t}})^{-1}\mu\frac{(v_x)^2}{v_{xx}}.
\end{aligned}
$$

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.

USC University of Southern California

Indifference pricing IV

Heston model

$$
\begin{aligned}
dS_t &= \mu S_t dt + \sqrt{Y_t} S_t dW_t^{(1)} \\
dY_t &= k(m - Y_t)dt + c\sqrt{Y_t}\left(\rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)}\right),
\end{aligned}
$$

Indifference pricing V

HJB equation

$$
\begin{aligned}
v(T,x,y) &= U(x)\\
0 &= -v_t - k(m-y)v_y - \frac{1}{2}c^2 y v_{yy} - \sup_{\theta\in\mathbb{R}}\left(\frac{1}{2}\theta^2 y v_{xx} + \theta(\mu v_x + \rho c y v_{xy})\right)\\
&= -v_t - k(m-y)v_y - \frac{1}{2}c^2 y v_{yy} + \frac{(\mu v_x + \rho c y v_{xy})^2}{2 y v_{xx}}.
\end{aligned}
$$

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

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

USC University of Southern California

Indifference pricing VII

HJB equation

$$
\begin{aligned}
U(x) &= v(T, r, x, s_1, y_1, y_2) \\
0 &= -v_t - (\mathbf{L}^r + \mathbf{L}^Y + \mathbf{L}^{S^1})v - rxv_x \\
&\quad + \frac{((\mu_1 - r)v_x + \sigma_1^2 y_1 s_1^{2\beta_1 - 1} v_{xs_1})^2}{2\sigma_1^2 y_1 s_1^{2\beta_1 - 2} v_{xx}} + \frac{((\mu_2 - r)v_x)^2}{2\sigma_2^2 y_2 v_{xx}}
\end{aligned}
$$

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

$$
\text{and} \quad \mathbf{L}^{S^1} v = \mu_1 s_1 v_{s_1} - \frac{1}{2}\sigma_1^2 s_1 y_1 v_{s_1 s_1}.
$$

USC University of Southern California

Monte Carlo methods for PDE

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

$$\begin{aligned} -\partial_t v - \bar{F}(t, x, v(t,x), Dv(t,x), D^2 v(t,x)) &= 0, [0,T) \times \mathbb{R}^d \\ v(T, \cdot) &= g. \end{aligned}$$

Definition:

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

$$\begin{aligned} -\mathcal{L}^X v - F(t,x,v(t,x),Dv(t,x),D^2v(t,x)) &= 0, [0,T)\times\mathbb{R}^d \\ v(T,\cdot) &= g. \end{aligned}$$

where $\mathcal{L}^X\varphi := \frac{\partial\varphi}{\partial t} + \mu\cdot D\varphi + \frac{1}{2}\sigma^{\mathrm{T}}\sigma\cdot D^2\varphi$ 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 μ and σ

$u_t + \frac{1}{2}\triangle u = \{u_t + \frac{1}{4}\triangle u\} + \{\frac{1}{4}\triangle u\} = \{u_t + \frac{1}{8}\triangle u\} + \{\frac{3}{8}\triangle u\}$ but not $\{u_t + \frac{3}{4}\triangle u\} + \{\frac{-1}{4}\triangle u\}$

Southern California

Fully nonlinear Parabolic PDEs III

A backward numerical scheme

$h = \frac{T}{n}$ and $t_i = ih$. $\hat{X}_h$ 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:

$$\begin{aligned}\hat{D}\hat{v}(t_{i+1},x) &= \mathbb{E}_{t_i,x}[D\hat{v}(t_{i+1},\hat{X}_h^x)(t_k,\hat{X}_h^x)]\\ \hat{D}^2\hat{v}(t_{i+1},x) &= \mathbb{E}_{t_i,x}[D^2\hat{v}(t_{i+1},\hat{X}_h^x)(t_k,\hat{X}_h^x)]\end{aligned}$$

USC University of Southern California

Fully nonlinear Parabolic PDEs IV

Key Lemma: Integration by part

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

$$\mathbb{E}[D^i \varphi(x + \sigma W_h)] \quad = \quad \mathbb{E}[\varphi(x + \sigma W_h) H_i^h(W_h)],$$

where $H_h = (H_0^h, H_1^h, H_2^h)$ and

$$H_0^h(x) = 1, H_1^h(x) = \frac{1}{h}\sigma'(x)^{-1} W_h,$$

$$H_2^h(x) = \frac{1}{h^2}\sigma'(x)^{-1}(W_h W_h' - h\mathbf{I}_{d\times d})\sigma(x)^{-1}$$

One dimensional case

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.
- $\mathbb{E}[\psi(x+\sigma W_h)H_2^h(W_h)] \approx \frac{\psi(x+\sigma\sqrt{3h})+\psi(x-\sigma\sqrt{3h})-\psi(x)}{3\sigma^2 h^2} \approx \psi''(x)$

Southern California

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.
- $\mathbb{E}[\psi(x+\sigma W_h)H_2^h(W_h)] \approx \frac{\psi(x+\sigma\sqrt{3h})+\psi(x-\sigma\sqrt{3h})-\psi(x)}{3\sigma^2 h^2} \approx \psi''(x)$

Southern California

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

Linear PDEs

$$0 = -v_t - \mathcal{L}^X v + kv \quad \text{and} \quad v(T,\cdot) = g(\cdot).$$

$\mathcal{L}^X = \frac{1}{2}\sigma^{\mathsf{T}}\sigma \cdot D^2 v - +mu \cdot Dv$.

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

where $dX_t = \mu dt + \sigma \cdot dW_t$.

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 $0 = -v_t - \mathcal{L}^X v$ and $\Phi_h^{(2)}[g](t,x) = \exp(-\int_t^{t+h} k_s ds)g(x)$ the semi–group generated by $0 = -v_t + kv$. Then, $(h \to 0)$

$$v(t,x) \approx \Phi_h^{(1)} \circ \Phi_h^{(2)} \circ \cdots \circ \Phi_h^{(1)} \circ \Phi_h^{(2)}[g](t,x)$$

Hopefully, two semi–groups commute:

$$\begin{aligned} v(t,x) &\approx \Phi_{T-t}^{(1)} \circ \Phi_{T-t}^{(2)}[g](t,x) \\ &\to \mathbb{E}_{t,x}\left[\exp\left(-\int_t^T k(X_s)ds\right) g(X_T)\right]. \end{aligned}$$

But, when the equation is non–linear, they don't commute.

Semi–linear PDEs I

Semi–linear equations

$$\begin{aligned} 0 &= -v_t - \frac{1}{2}\sigma^{\mathsf{T}}\sigma \cdot D^2 v - \mu \cdot Dv - F(v, Dv) \\ v(T,\cdot) &= g(\cdot). \end{aligned}$$

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 $dX_t = \mu dt + \sigma dW_t$.

$$\begin{aligned} dY_t &= F(Y_t, Z_t)dt - Z_t dX_t \\ Y_T &= g(X_T). \end{aligned}$$

Relation with PDE

If v is the classical solution of semi–linear PDE, $Y_t = v(t, X_t)$, $Z_t = Dv(t, X_t)$.
Theory: [Bismut 78], [El Karoui–Peng–Quenez 97], [Pardoux–Peng92]

USC University of Southern California

Semi–linear PDEs III

Discretization of BSDE

$$\begin{aligned}\hat{Y}_i &= \mathbb{E}_i[\hat{Y}_{i+1} + hF(t_i, \hat{X}_i, \hat{Y}_i, \hat{Z}_i)] \quad \hat{Y}_T = g(X_T) \\ \hat{Z}_i &= \frac{1}{h}\mathbb{E}_i[\hat{Y}_{i+1}\Delta W_{i+1}] \approx \Delta_t.\end{aligned}$$

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

Fully non–linear PDEs I

Fully non–linear equations

$$
\begin{aligned}
0 &= -\frac{\partial v}{\partial t} - \bar{F}(t, x, v, Dv, D^2 v) \\
v(T, \cdot) &= g(\cdot).
\end{aligned}
$$

Parabolicity: $\bar{F}(t, x, r, p, \gamma)$ is increasing with respect to γ.
The solution definitely should be considered in viscosity sense.
Application: Merton portfolio selection model, Super–hedging under Γ constrain
No Monte Carlo method is known for the general above type.

Fully non–linear PDEs II

2BSDE

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

If v is the classical solution of semi–linear PDE, $Y_t = v(t, X_t)$, $Z_t = Dv(t, X_t)$, $\Gamma_t = D^2 v(t, X_t)$, $A_t = \mathcal{L}^X Dv(t, X_t)$.
Theory is recently developed its first steps by [Cheridito–Soner–Touzi–Victoir 07] and [Soner–Touzi–Zhang 10]$\times$4.

USC University of Southern California

Fully non–linear PDEs III

Discretization of 2BSDE

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

Alternative scheme.

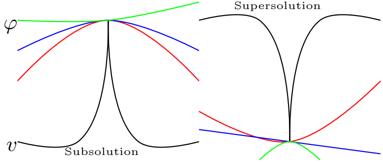
Main results (F–Touzi–Warin) I

Viscosity solution

An upper-semicontinuous (resp. lower semicontinuous) function $\underline{v}$ (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 φ satisfying

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

we have: $\quad -\mathcal{L}^X \varphi - F(t, x, \mathcal{D}\varphi(t, x)) \leq$ (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, γ) uniformly in t.
2. $|F(\cdot, \cdot, 0, 0, 0)|_\infty < \infty$.
3. F is elliptic (increasing w.r.t. γ).
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. $F_p \in \mathrm{Image}(F_\gamma)$ and $\left|F_p^\mathrm{T} F_\gamma^- F_p\right|_\infty < +\infty$.

Main results (F–Touzi–Warin) III

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

$$v^h \longrightarrow v \qquad \text{locally uniformly.}$$

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).

Main results (F–Touzi–Warin) IV

Monotonicity

Should be: If $\varphi \leq \psi$ then $\mathbf{T}_h[\varphi] \leq \mathbf{T}_h[\psi]$.
But it is: If $\varphi \leq \psi$ then $\mathbf{T}_h[\varphi] \leq \mathbf{T}_h[\psi] + Ch\mathbb{E}[(\psi - \varphi)(t+h, \hat{X}_h^x)]$.

Stability

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

Consistency

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

$$\frac{1}{h}(\psi + c - \mathbf{T}_h[\psi + c])(t', x') \to -v_t - \mathcal{L}^X v - F(t, x, \mathcal{D}v(t, x)).$$

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:

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

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

where the functions μ, a, σ^α, b^α, c^α and f^α satisfy:

$$|\mu|_\infty + |a|_\infty + \sup_{\alpha \in \mathcal{A}} \left(|\sigma^\alpha|_1 + |b^\alpha|_1 + |c^\alpha|_1 + |f^\alpha|_1\right) \quad < \quad \infty.$$

USC University of Southern California

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 $\alpha \in \mathcal{A}$:

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

Thm: F–Touzi–Warin

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

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

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).

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_ε and a smooth super solution v^ε 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) $|\partial_t^k D^k v_\varepsilon^\varepsilon| \leq C\varepsilon^{2-2k-|l|}$.

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

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

$|\hat{v} - v| \leq |\hat{v} - v_\varepsilon| + |v_\varepsilon - v| \leq C\varepsilon^{-3}h + C\varepsilon$

Backward implementation I

Scheme reminded

$$v^h(T,.) = g \quad \text{and} \quad v^h(t_i,x) = \mathbf{T}_h[v^h(t_{i+1},\cdot)](x).$$

$$\mathbf{T}_h\psi(x) := \mathbb{E}[\psi(\hat{X}_h^x)] + hF(x, \mathcal{D}_h\psi(x))$$

$$\mathcal{D}_h\psi := \left(\mathcal{D}_h^0\psi, \mathcal{D}_h^1\psi, \mathcal{D}_h^2\psi\right) \qquad \mathcal{D}_h^i\psi(x) := \mathbb{E}[\psi(\hat{X}_h^x)H_i^h(W_h)], i = 0, 1, 2.$$

Backward implementation II

Implementation

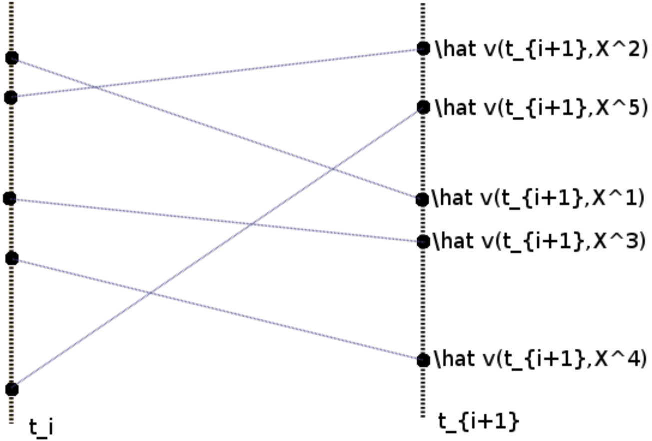
1. $t_i = \frac{iT}{n}$.
2. Generating N sample paths from the Euler discretization of X_t; $\hat{X}_t$.

$$\{(\hat{X}_{t_i}^{(j)}, \Delta W_{t_{i+1}}^{(j)})|0 = t_0, \cdots, t_n = T, j = 1, \cdots, N\}.$$

3. Start from terminal condition $g(\cdot)$. And proceed backward in time.
4. Knowing $v^h(t_{i+1}, \hat{X}_{t_{i+1}}^{(j)})$s for $j = 1, \cdots, N$, then one calculates $v^h(t_i, \hat{X}_{t_i}^{(k)})$s for $k = 1, \cdots, N$ using the scheme.

USC University of Southern California

Backward implementation III



Backward implementation IV

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

$Y = v(t+h, \hat{X}_{t+h})H_h^i$ 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 $\varepsilon_\kappa \to \delta_0$. 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

$Y = v(t+h, \hat{X}_{t+h})H_h^i$ and $X = \hat{X}_t$. Formally;

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

Coefficients c_ks should be determined so that the L^2 approximation error be minimum.

Approximation of conditional expectations III

Malliavin methods

$Y = v(t+h, \hat{X}_{t+h})H_h^i$ and $X = \hat{X}_t$. Formally;

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

$\mathbb{E}[Y\delta_x(X)] = \mathbb{E}[Y\mathbb{1}_{\{X>x\}}\bar{\delta}_0^t]$ where $\bar{\delta}_0^t$ is Skorokhod integral which depends only on the path of X from 0 to t.

Approximation of conditional expectations IV

Stochastic scheme

$$\tilde{\mathbf{T}}_h^N[\psi](t,x) \quad := \quad \hat{\mathbb{E}}^N\left[\psi(t+h,\hat{X}_h^x)\right] + hF\left(\cdot,\hat{\mathcal{D}}_h\psi\right)(t,x),$$
$$\hat{\mathbf{T}}_h^N[\psi](t,x) \quad := \quad -K_h[\psi] \vee \tilde{\mathbf{T}}_h^N[\psi](t,x) \wedge K_h[\psi]$$

where

$$\hat{\mathcal{D}}_h\psi(t,x) := \hat{\mathbb{E}}^N\left[\psi(t+h,\hat{X}_h^{t,x})H_h(t,x)\right],\ K_h[\psi] := \|\psi\|_\infty(1+C_1h)+C_2h,$$

and

$$C_1 = \frac{1}{4}|F_p^{\mathrm{T}}F_\gamma^- F_p|_\infty + |F_r|_\infty \quad \text{and} \quad C_2 = |F(t,x,0,0,0)|_\infty.$$

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 ψ is a function by b and H_i's are the Hermite polynomials:

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

There exist constants $C_b, \lambda, \nu > 0$ such that $\left\|\hat{\mathbb{E}}^N[R] - \mathbb{E}[R]\right\|_p \leq C_b h^{-\lambda} N^{-\nu}$ 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 $\nu = \frac{1}{2p}$.

Asymptatic 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

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

Assume that the final condition g is bounded Lipschitz, and the coefficients μ and σ are bounded. Then, for almost every ω:

$$\hat{v}^h_{N_h}(\cdot,\omega) \longrightarrow v \qquad \text{locally uniformly,}$$

where v is the unique viscosity solution of equation.

Asymptatic properties II

Rate of convergence result

Let the nonlinearity F be as in Assumption **HJB**, and consider a regression operator satisfying Assumption **E**. Let the sample size N_h be such that

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

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

$$\|v - \hat{v}^h\|_p \quad \leq \quad Ch^{1/10}.$$

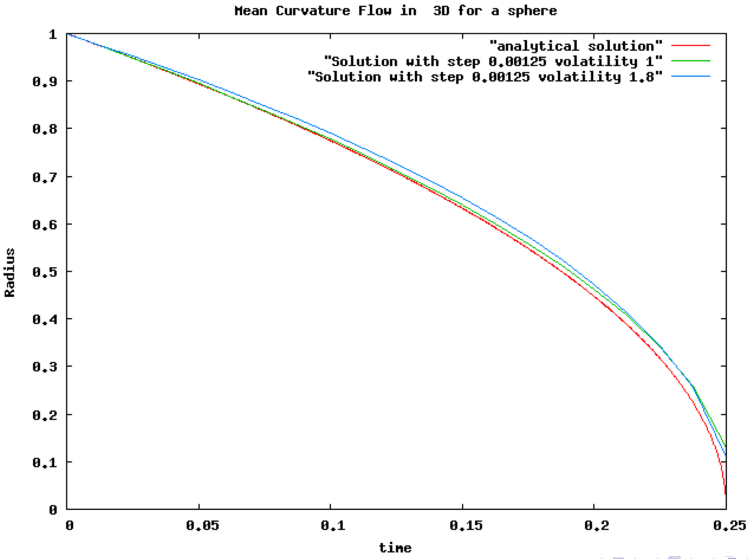
Problem consider in numerical experiments

- 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)

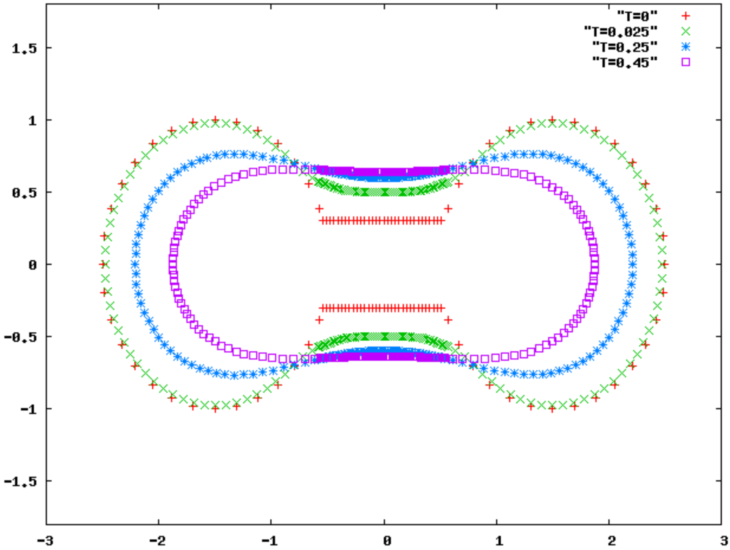
Alternative schemes

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

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
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15
0.2
0.25

2-d dumbbell



Portfolio selection in dimension 2 I

Heston model

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{Y_t} S_t dW_t^{(1)} \\ dY_t &= k(m - Y_t)dt + c\sqrt{Y_t}\left(\rho dW_t^{(1)} + \sqrt{1-\rho^2} dW_t^{(2)}\right), \end{aligned}$$

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

$$
\begin{aligned}
v(T,x,y) &= -e^{-\eta x} \\
0 &= -v_t - k(m-y)v_y - \frac{1}{2}c^2 y v_{yy} \\
&\quad - \sup_{\theta\in\mathbb{R}}\left(\frac{1}{2}\theta^2 y v_{xx} + \theta(\mu v_x + \rho c y v_{xy})\right) \\
&= -v_t - k(m-y)v_y - \frac{1}{2}c^2 y v_{yy} + \frac{(\mu v_x + \rho c y v_{xy})^2}{2 y v_{xx}}.
\end{aligned}
$$

Zariphopoulou semi-explicit solution

$$
v(t,x,y) = -e^{-\eta x}\left\|\exp\left(-\tfrac{1}{2}\int_t^T \frac{\mu^2}{\tilde{Y}_s} ds\right)\right\|_{\mathbb{L}^{1-\rho^2}}
$$

Portfolio selection in dimension 2 III

Separation into linear and fully non–linear part

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

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

Portfolio selection in dimension 2 IV

Truncation of the non–linearity

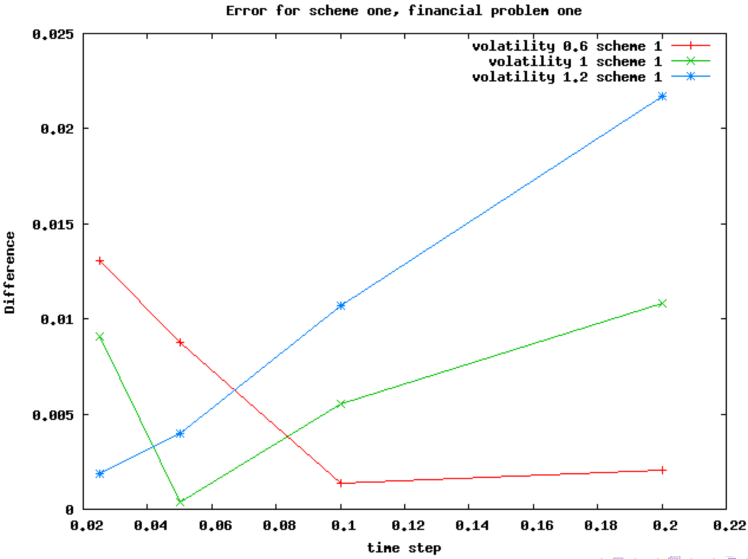
$$F_{\varepsilon,M}(y,z,\gamma) \quad := \quad \frac{1}{2}\sigma^2\gamma_{11} - \sup_{\varepsilon\leq\theta\leq M}\left(\frac{1}{2}\theta^2(y\vee\varepsilon)\gamma_{11} + \theta(\mu z_1 + \rho c(y\vee\varepsilon)\gamma_{12}\right),$$

Choice of diffusion

$$dX_t^{(1)} = \sigma dW_t^{(1)}, \quad \text{and} \quad dX_t^{(2)} = k(m - X_t^{(2)})dt + c\sqrt{X_t^{(2)}}dW_t^{(2)}.$$

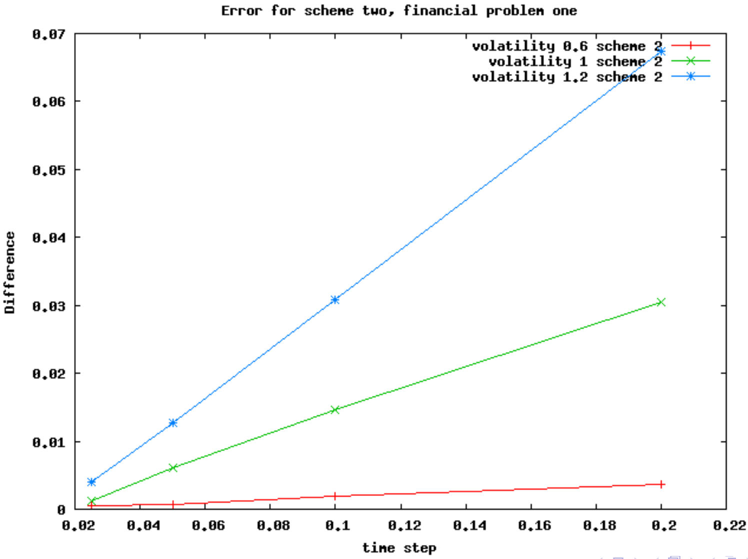
$\mu = 0.15$, $c = 0.2$, $k = 0.1$, $m = 0.3$, $Y_0 = m$, $\rho = 0$, $x_0 = 1$, $T = 1$ then $v_0 = -0.3534$.

2 dimensional 1



University of Southern California

2 dimensional 2



Portfolio selection in dimension 5 I

Vasicek, Heston and CEV-SV models

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

Portfolio selection in dimension 5 II

HJB equation

$$
\begin{aligned}
0 = & -v_t - (\mathbf{L}^r + \mathbf{L}^Y + \mathbf{L}^{S^1})v - rxv_x \\
& + \frac{((\mu_1 - r)v_x + \sigma_1^2 y_1 s_1^{2\beta_1 - 1} v_{xs_1})^2}{2\sigma_1^2 y_1 s_1^{2\beta_1 - 2} v_{xx}} + \frac{((\mu_2 - r)v_x)^2}{2\sigma_2^2 y_2 v_{xx}}
\end{aligned}
$$

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

$$
\text{and } \mathbf{L}^{S^1} v = \mu_1 s_1 v_{s_1} - \frac{1}{2}\sigma_1^2 s_1 y_1 v_{s_1 s_1}.
$$

Southern California

Portfolio selection in dimension 5 III

Separation into linear and fully non–linear part

$$-v_t - (\mathbf{L}^r + \mathbf{L}^Y + \mathbf{L}^{S^1})v - \frac{1}{2}\sigma^2 v_{xx} + F\left((x, r, s_1, y_1, y_2), Dv, D^2 v\right) = 0,$$
$$v(T, x, r, s_1, y_1, y_2) = -e^{-\eta x},$$

$$F(u, z, \gamma) = \frac{1}{2}\sigma^2\gamma_{11} - x_1 x_2 z_1 + \frac{((\mu_1 - x_2)z_1 + \sigma_1^2 x_4 x_3^{2\beta_1 - 1}\gamma_{1,3})^2}{2\sigma_1^2 x_4 x_3^{2\beta_1 - 2}\gamma_{11}} + \frac{((\mu_2 - x_2)z_1)^2}{2\sigma_2^2 x_5 \gamma_{11}},$$

where $u = (x_1, \cdots, x_5)$.

Portfolio selection in dimension 5 IV

Truncation of the non–linearity

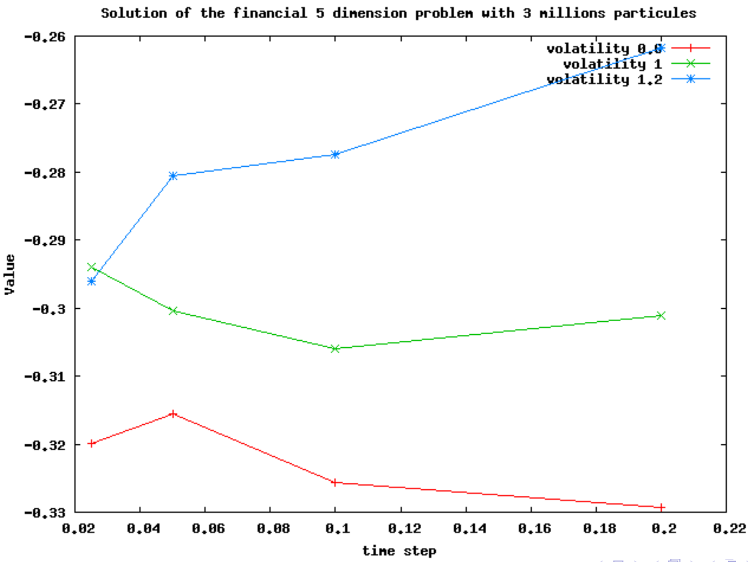
$$
\begin{aligned}
F_{\varepsilon,M}(u,z,\gamma) \quad := \quad & \frac{1}{2}\sigma^2\gamma_{11} - x_1x_2z_1 + \sup_{\varepsilon\le|\theta|\le M}\Big\{(\theta\cdot(\mu - r\mathbb{1})z_1 \\
& +\theta_1\sigma_1^2(x_4\vee\varepsilon)(x_3\vee\varepsilon)^{2\beta_1-1}\gamma_{13} \\
& +\frac{1}{2}(\theta_1^2\sigma_1^2(x_3\vee\varepsilon)(x_4\vee\varepsilon)^{2\beta_1-2} + \theta_2^2\sigma_2^2(x_5\vee\varepsilon))\gamma_{11}\Big\},
\end{aligned}
$$

Portfolio selection in dimension 5 V

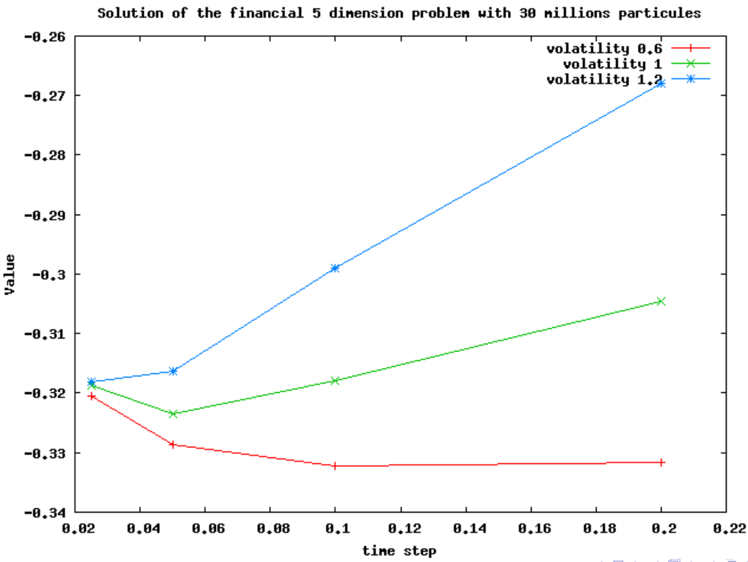
Choice of diffusion

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

5 dimensional 1



5 dimensional 2



Non–local Parabolic PDEs I

Fully non–linear non–local parabolic PDEs

$$
\begin{aligned}
&-\mathcal{L}^X v(t,x) - F\left(t,x,v(t,x),Dv(t,x),D^2 v(t,x),v(t,\cdot)\right) = 0,\\
&v(T,\cdot) = g,
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{L}^X \varphi(t,x) &:= \left(\frac{\partial \varphi}{\partial t} + \mu \cdot D\varphi + \frac{1}{2} a \cdot D^2 \varphi\right)(t,x)\\
&+ \int_{\mathbb{R}^d_*} \Big(\varphi(t,x+\eta(t,x,z)) - \varphi(t,x) - \mathbb{1}_{\{|z|\le 1\}} D\varphi(t,x)\eta(t,x,z)\Big) d\nu(z).
\end{aligned}
$$

[F 10]

Southern California

Parabolic PDE in more general domains I

Fully non–linear parabolic PDEs in general domains

$$
\begin{aligned}
&-\mathcal{L}^X v(t,x) - F\left(t, x, v(t,x), Dv(t,x), D^2 v(t,x)\right) = 0, \text{ on } \mathcal{O} \\
&v(t,\cdot) = \varphi(\cdot), \text{ on } \partial\mathcal{O}, \\
&v(T,\cdot) = g,
\end{aligned}
$$

Thank you for your attention.