Branching for PDEs

Xavier Warin

CEMRACS July 2017

Table of contents

A C++ toolbox with python interface

- Regression methods for conditional expectations :
  - Local with Linear, Constant per mesh approximation ,
  - Local adaptive to the distribution with Linear, Constant per mesh approximation ,
  - Global polynomial (Hermite, Canonical, Tchebychev),
  - Sparse grids,
- Interpolation methods (linear, Monotone Legendre, sparse grids)

Provide a framework to solve complex optimization problems

- General HJB equations with deterministic Semi Lagrangian methods,
- Non linear Stochastic Optimization problems with stocks :
  - Regressions with Monte Carlo for non controlled processes,
  - Stochastic Monte Carlo quantization for controlled process,
- Some Linear problems with stocks in high dimension : Stochastic Dual Dynamic Programming Method.

Parallelization :

- Message Passing (MPI),
- Multi-threaded,
- Using vectorized matrix/array library Eigen (INRIA):

An open source library

- Developed during the ANR Caesars.
- Gitlab site :
  `https://gitlab.com/stochastic-control/StOpt`,
- Documentation :
  `https://hal.archives-ouvertes.fr/hal-01361291`
- Python installer (Windows, Linux available at Labo FiME web site :
  `https://www.fime-lab.org/`

Try it and avoid to redevelop (most of the time less efficiently) even if branching not currently available.

Table of contents

The KPP equation McKean's formulation [1]

Equation to solve in $\mathbb{R}^d$:

$$\begin{aligned} \partial_t u + \mathcal{L}u + \beta f(u) = 0, \\ \mathcal{L}u = \mu \cdot Du + \frac{1}{2}\sigma\sigma^\top : D^2 u \\ u(T,.) = g \end{aligned}$$

with $\mu \in \mathbb{R}^d$, $\sigma \in \mathbb{M}^d$, the non linear term :

$$f(u) = u^2 - u$$

and notation $A : B := \text{Trace}(AB^\top)$.

Using Ito..

Consider the process for W_t a d dimensional Brownian motion :

$$dX_t^{0,x} = \mu dt + \sigma dW_t$$
$$X_0^{0,x} = x$$

Supposing regularity of the solution :

$$\mathbb{E}\left[u(T, X_T^{0,x})e^{-\beta T}\right] = u(0,x) + \mathbb{E}\left[\int_0^T e^{-\beta s}\left(\partial_t u + \mathcal{L}u - \beta u\right)(s, X_s^{0,x})ds\right]$$
$$u(0,x) = \mathbb{E}\left[g(X_T^{0,x})e^{-\beta T}\right] + \mathbb{E}\left[\int_0^T \left(\beta e^{-\beta s} u(s, X_s^{0,x})^2\right) ds\right]$$

if τ a R.V. following an exponential law with parameter β :

$$\mathbb{E}\left[1_{\tau > T}\right] = e^{-\beta T} = 1 - \text{cdf}(\tau),$$
$$\rho^\tau(s) = \beta e^{-\beta s}$$

Introducing a Poisson process $\tau^{(1)}$ with intensity β

Considering the integral as an expectation

$$\begin{aligned} u(0,x) =& \mathbb{E}_{0,x}\left[g(X_T^{0,x})1_{\tau^{(1)}>T} + 1_{\tau^{(1)}<T} u(\tau^{(1)}, X_{\tau^{(1)}}^{0,x})^2\right] \\ =& \mathbb{E}_{0,x}\left[\psi(\tau^{(1)}, X_{\tau^{(1)}}^{0,x})\right] \end{aligned} \qquad (1)$$

where

$$\psi(t,x) = g(x)1_{t>T} + 1_{t<T} u(t,x)^2$$

Introduce the Poisson processes $\tau^{(1,1)}$, $\tau^{(1,2)}$ (2 particles) by independence

$$\begin{aligned} u(t,x)^2 =& \mathbb{E}_{t,x}\left[\psi(t+\tau^{(1,1)}, X_{t+\tau^{(1,1)}}^{t,x}\right] \mathbb{E}_{t,x}\left[\psi(t+\tau^{(1,2)}, X_{t+\tau^{(1,2)}}^{t,x}\right] \\ =& \mathbb{E}_{t,x}\left[\psi(t+\tau^{(1,1)}, X_{t+\tau^{(1,1)}}^{t,x})\psi(t+\tau^{(1,2)}, X_{t+\tau^{(1,2)}}^{t,x})\right] \end{aligned} \qquad (2)$$

By recursion

Plugging (2) in (1), introducing

$$T_{(1)} = T \wedge \tau^{(1)},$$
$$T_{(1,j)} = T \wedge (T_{(1)} + \tau^{(1,j)}), \quad j = 1, 2$$

$$u(0,x) = \mathbb{E}_{0,x}\Big[1_{T_{(1)}=T}\, g(X^{0,x}_{T_{(1)}}) + 1_{T_{(1)}<T} \prod_{j=1}^{2} \Big(1_{T_{(1,j)}=T}\, g(X^{0,x}_{T_{(1,j)}}) + 1_{T_{(1,j)}<T}\, u(T_{(1,j)}, X^{0,x}_{T_{(1,j)}})^2 \Big) \Big]$$

Recursion till all particles arrive at date T.

Kpp tree

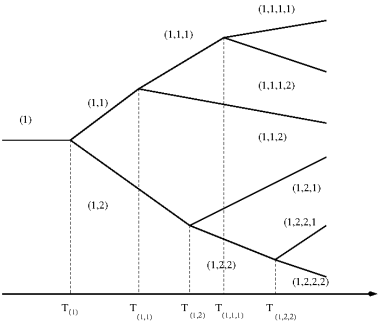


Figure: Galton-Watson tree for KPP

- At date $T_{(1)}$ (1) generates $(1,1)$ and $(1,2)$,
- At date $T_{(1,1)}$, $(1,1)$ generates $(1,1,1)$ and $(1,1,2)$,
- At date $T_{(1,1,1)}$, $(1,1,1)$ generates $(1,1,1,1)$ and $(1,1,1,2)$
- At date $T_{(1,2)}$, $(1,2)$ generates $(1,2,1)$ and $(1,2,2)$,
- At date $T_{(1,2,2)}$, $(1,2,2)$ generates $(1,2,2,1)$ and $(1,2,2,2)$,

Notations

- $k = (k_1, k_2, ..., k_{n-1}, k_n)$, $k_i \in \{1, 2\}$ particle of generation n
- $k- = (k_1, k_2, ..., k_{n-1})$ its ancestor , $(1)- = \emptyset$
- $\mathcal{K}_t^n$ set of all living particles of generation n at date t.
- $\mathcal{K}_t := \cup_{n \geq 1} \mathcal{K}_t^n$ set of all living particles at date t.
- $\overline{\mathcal{K}}_t$ (resp. $\overline{\mathcal{K}}_t^n$) set of all particles (resp. of generation n) alive before time t
- τ^k Poisson process associated to particle $k = (k_1, ..., k_n)$,
- Branching times per particle k

$$\begin{aligned} T_k :=& \left(T_{k-} + \tau^k\right) \wedge T, \\ T_\emptyset =& 0 \end{aligned}$$

Kpp tree

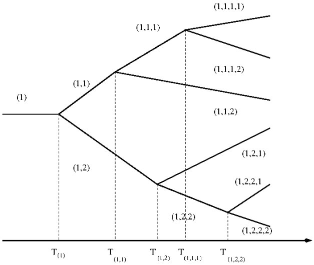


Figure: Galton-Watson tree for KPP

- $(1, 1, 1, 2)$ ancestor :

$$(1, 1, 1, 2)- = (1, 1, 1),$$

- $\mathcal{K}_T^3 = \{(1, 2, 1), (1, 1, 2)\}$
-

$$\begin{aligned}\mathcal{K}_T^4 =& \{(1, 1, 1, 1), (1, 1, 1, 2), \\ & (1, 2, 2, 1), (1, 2, 2, 2)\}\end{aligned}$$

- $\overline{\mathcal{K}}_T$ the 10 particles.

PDE representation

- d-dimensional Brownian motion $(W_t^k)_{k\in\overline{\mathcal{K}}_T}$,
- for $k \in \overline{\mathcal{K}}_T \setminus \mathcal{K}_T$, dynamic for $T_{k-} \leq t < T_k$

$$X_t^k \;=\; X_{T_{k-}}^{k-} + \mu(t - T_{k-}) + \sigma W_{t-T_{k-}}^k$$

- Sample estimator:

$$\hat{u}(0,x) = \prod_{k\in\mathcal{K}_T} g(X_T^k)$$

- The number of particles in $\overline{\mathcal{K}}_T$ is finite a.s
- if $||g||_\infty < 1$, $u \in C^{1,2}([0,T]\times\mathbb{R}^d)$:
  - $\hat{u}(0,x) \in \mathbb{L}^1 \cap \mathbb{L}^2$,
  - $u(0,x) = \mathbb{E}\Big[\hat{u}(0,x)\Big]$

Table of contents

First extension of KPP

Non linear PDE

$$\begin{aligned} \partial_t u + \mathcal{L}u + \beta f(u) = 0, \\ u(T,.) = & g \end{aligned}$$

with

$$f(u) = \sum_{i=0}^{N} p_k u^k - u$$

$$\sum_{i=0}^{N} p_k = 1, \quad 0 \leq p_k \leq 1$$

Feynman Kac :

- Supposing regularity of the solution :

$$u(t,x) = \mathbb{E}\left[g(X_T^x)e^{-\beta(T-t)}\right] + \mathbb{E}\left[\int_t^T \left(\beta e^{-\beta(s-t)} \sum_{i=1}^{N} p_i u(s, X_s^x)^i\right) ds\right]$$

- Introduce for particle k, $(I^k)_{k \in \overline{\mathcal{K}}_T \setminus \mathcal{K}_T}$ random such that

$$P(I^k = l) = p_l.$$

$$u(0,x) = \mathbb{E}_{0,x}\left[g(X_T^{(1)}) 1_{\tau^{(1)} > T} + 1_{\tau^{(1)} < T} u(\tau^{(1)}, X_{\tau^{(1)}}^{(1)})^{I^{(k)}}\right]$$

- same estimator

$$u(0,x) = \mathbb{E}[\prod_{k \in \mathcal{K}_T} g(X_T^k)]$$

Tree generalization : $f(u) = p_0 + p_1 u + p_2 u^2 + p_3 u^3$

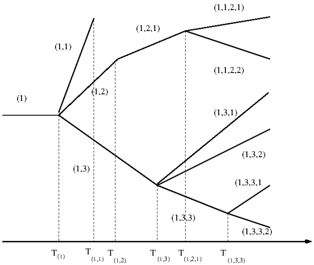


Figure: Galton-Watson tree for generalized KPP

- (1) , (1, 3) generate 3 particles (probability p_3),
- (1, 1) dies without children (prob p_0),
- (1, 2) generates one son (prob p_1),
- (1, 2, 1), and (1, 3, 3) generates 2 sons (prob p_2)

Table of contents

First extension of KPP

Non linear PDE

$$\partial_t u + \mathcal{L}u + \hat{f}(u) = 0, \quad u(T,.) = g$$

with

$$\hat{f}(u) = \sum_{i=0}^{N} \hat{a}_k u^k$$

Rewrite choosing β, $(p_i)_{i=0}^{N}$ with positives values

$$\partial_t u + \mathcal{L}u + \beta f(u) = 0,$$

$$f(u) = \sum_{i=0}^{N} p_i a_i u^i - u$$

where $\beta p_i a_i = \hat{a}_i$, $i \neq 1$, $a_1 = \dfrac{\hat{a}_1 + 1}{\beta p_1}$.

Marked tree

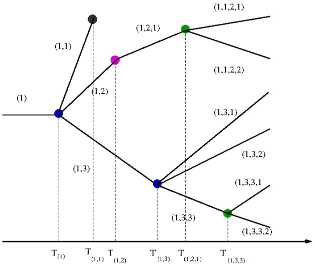


Figure: Marked Galton-Watson tree

- (1) , (1, 3) marked 3 generates 3 particles ,
- (1, 1) marked 0 dies without children,
- (1, 2) marked marked 1 generates one son (prob p_1),
- (1, 2, 1) and (1, 3, 3) marked 2 generates 2 sons.

Feynman Kac :

- $$u(0,x) = \mathbb{E}_{0,x}\left[g(X_T^{(1)})1_{\tau^{(1)}>T} + 1_{\tau^{(1)}<T} a_{I^{(1)}} u(\tau^{(1)}, X_{\tau^{(1)}}^{(1)})^{I^{(1)}}\right]$$
- Same regularity on u, under sufficient condition $\sum_i \frac{\hat{a}_i}{\beta} ||g||_\infty \leq 1$ (not depending on p_k)

$$\begin{aligned} u(0,x) =& \mathbb{E}[\phi] \\ \phi =& \prod_{k \in \mathcal{K}_T} g(X_T^k) \prod_{k \in \overline{\mathcal{K}}_T \setminus \mathcal{K}_T} a_{I(k)} \\ =& \prod_{k \in \mathcal{K}_T} g(X_T^k) \prod_{i=0}^{N} a_i^{w_i} \end{aligned}$$

- w_i number of particles marked i (branching i particles)
- p_k chosen to minimize variance

$$p_k = \frac{a_k ||g||_\infty^k}{\sum_i a_i ||g||_\infty^i}$$

Example

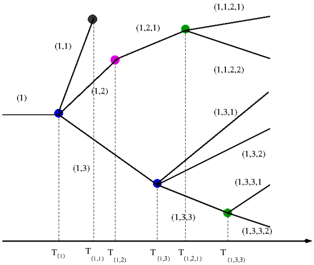


Figure: Sample 0 for Marked Galton-Watson tree

- Sample 0

$$\phi_0 = a_3 \, [a_0] \left[a_1 \, a_2 g(X_T^{(1,1,2,1)}) g(X_T^{(1,1,2,2)}) \right] \left[a_3 g(X_T^{(1,3,1)}) g(X_T^{(1,3,2)}) \left(a_2 g(X_T^{(1,3,3,1)}) g(X_T^{(1,3,3,2)}) \right) \right]$$

Alternative

No probability to choose the power of u

$$u(0,x) = \mathbb{E}_{0,x}\left[g(X_T^{0,x})\mathbb{E}_{0,x}(1_{\tau^{(1)}>T})] + 1_{\tau^{(1)}<T} \sum_i p_i a_i u(\tau^{(1)}, X_{\tau^{(1)}}^{0,x})^i \right]$$

At each branching :

- treat each term $u(\tau^{(1)}, X_{\tau^{(1)}}^{0,x})^i$ generating i particles
- summation : $\sum_i p_i a_i u(\tau^{(1)}, X_{\tau^{(1)}}^{0,x})^i$

Disadvantage

- Explosion of the computer time if many terms on the polynomial or too long maturities

Advantage :

- Reduce the variance,
- As easy to program as for the initial algorithm.

Table of contents

Framework for small maturities

Approximation of the driver by a local polynomial expansion

$$f(x,y) = \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} a_{j,\ell}(x) y^\ell \varphi_j(y), \tag{3}$$

where $(a_{j,\ell}, \varphi_j)_{\ell \le \ell_\circ, j \le j_\circ}$ is continuous and bounded maps satisfying

$$|a_{j,\ell}| \le C_{\ell_\circ} \ , \ |\varphi_j(y'_1) - \varphi_j(y'_2)| \le L_\varphi |y'_1 - y'_2| \text{ and } |\varphi_j| \le 1,$$

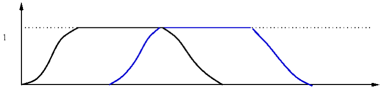


Figure: Example of ϕ functions

Feynman Kac

Feynman Kac, ρ density exponential law with β intensity, F CDF, $\bar{F} = 1 - F$, $\tau^{(1)}$ with density ρ

$$u(0,x) = \mathbb{E}\left[\frac{g(X_T^{0,x})}{\bar{F}(T)}\bar{F}(T) + \int_0^T \sum_{j=1}^{j_\circ}\sum_{\ell=0}^{\ell_\circ} \frac{\left(a_{j,\ell}u^\ell\varphi_j(u)\right)(s, X_s^{0,x})}{\rho(s)}\rho(s)ds\right]$$

$$= \mathbb{E}\left[\frac{g(X_T^{(1)})}{\bar{F}(T)}1_{\tau^{(1)}>T} + 1_{\tau^{(1)}<T}\sum_{j=1}^{j_\circ}\sum_{\ell=0}^{\ell_\circ}\frac{\left(a_{j,\ell}u^\ell\varphi_j(u)\right)(\tau^{(1)}, X_{T_{(1)}}^{0,x})}{\rho(\tau^{(1)})}\right]$$

Idea

- Impossible to use (29) directly in forward : u unknown so relevant φ_j unknown,
- Rewrite as $f(x,y) = \hat{f}(x,y,y)$, choosing probability p_ℓ with $\sum_{\ell=0}^{\ell_\circ} p_\ell = 1$

$$\hat{f}(x,y,y') = \sum_{j=1}^{j_\circ} \sum_{\ell=0}^{\ell_\circ} p_\ell \frac{a_{j,\ell}(x)}{p_\ell} y^\ell \varphi_j(y'),$$

Theoretical algorithm

Use Picard iterations starting with u^0

- Using Feynman Kac

$$
\begin{aligned}
\hat{u}^{n+1}(0,x) =& \mathbb{E}_{0,x}\Bigg[\frac{g(X_T^{(1)})}{\bar{F}(T)}1_{\tau^{(1)}>T} + \\
& 1_{\tau^{(1)}<T}\sum_{j=1}^{j_\circ}\frac{a_{j,l^{(1)}}(\tau^{(1)},X^{(1)}_{T_{(1)}})\varphi_j(u^n(\tau^{(1)},X^{0,x}_{\tau^{(1)}}))}{\rho(\tau^{(1)})p_{l^{(1)}}}\hat{u}^{n+1}(\tau^{(1)},X^{0,x}_{\tau^{(1)}})^{l^{(1)}}\Bigg] \\
\hat{u}^{n+1}(0,x) =& \mathbb{E}_{0,x}\Bigg[\prod_{k\in\mathcal{K}_T}\frac{g(X_T^k)}{\bar{F}(T-T_{k-})}\prod_{k\in\overline{\mathcal{K}}_T\setminus\mathcal{K}_T}\frac{(a_{l^{(k)}}\varphi_j(u^n))(T_{(k)},X^k_{T_k})}{p_{l^{(k)}}\rho(\tau^{(k)})}\Bigg]
\end{aligned}
\tag{4}
$$

- Use a priory bound

$$
u^{n+1} = (\hat{u}^{n+1}\wedge M)\vee -M
$$

Convergence proved in [3].

Effective algorithm for general driver $f(u)$

- Choose a grid $y_i = y_{min} + i\frac{y_{max} - y_{min}}{N}$, $i = 0, N$, φ_j indicator function (not regular)

$$\varphi_j(y) = 1_{y \in [y_j, y_{j+1}[}$$

- Use quadratic or cubic expansion on each mesh for f , with C^1 or C^2 regularity defining f expansion,
- Time discretization $t_i = i\frac{T}{M}$ such that (4) has a bounded variance on $[t_, t_{i+1}]$

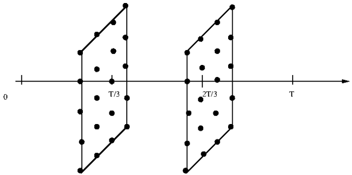


Figure: Resolution with interpolation

- Use interpolator $\hat{I}_i$ at date t_i on a grid G_i
- Use backward resolution : solve with branching on interval

Effective algorithm for general driver $f(u)$

for $x \in G_{M-1}$ **do**

  $T_\emptyset = t_{M-1}$

  $$u(t_{M-1}, x) = \mathbb{E}_{M-1,x}\left[\prod_{k\in\mathcal{K}_T} \frac{g(X_T^k)}{\bar{F}(T - T_{k-})} \prod_{k\in\overline{\mathcal{K}}_T\setminus\mathcal{K}_T} \frac{(a_{I(k)}\varphi_j(g))(T_{(k)}, X_{T_k}^{(k)})}{p_{I(k)}\rho(\tau^{(k)})}\right]$$

end for

for $i = M-2, 0$ **do**

  for $x \in G_i$ **do**

    $T_\emptyset = t_i$

    $$u(t_i, x) = \mathbb{E}_{i,x}\left[\prod_{k\in\mathcal{K}_{t_{i+1}}} \frac{\hat{l}_{i+1}(u(t_{i+1}, X_{t_{i+1}}^k))}{\bar{F}(t_{i+1} - T_{k-})} \prod_{k\in\overline{\mathcal{K}}_{t_{i+1}}\setminus\mathcal{K}_{t_{i+1}}} \frac{(a_{I(k)}\varphi_j(\hat{l}_{i+1}(u(t_{i+1}, .))))(T_{(k)}, X_{T_k}^{(k)})}{p_{I(k)}\rho(\tau^{(k)})}\right.$$

  end for

end for

Remark on the algorithm

- No Picard iteration : pure explicit scheme,
- Interpolation is needed:
  - To compare with general semi-Lagrangian methods [4] where interpolation is used and CFL stability condition (connecting time and spacial discretization)
  - Here CFL replace by variance condition
- Possible to use some general “most of the time high order” monotone interpolator [5] on regular grids
- Subject to the “curse” of dimension.

Does it work ?

$$\partial_t u + \mathcal{L}u + f(u) = 0,$$

- Domain $\mathbf{X} := [0,2]^d$
- SDE coefficient $V = 0.2$, $U = 0.1$

$$\mu(x) = U \times (\mathbf{1} - x) \text{ and } \sigma(x) := V \prod_{i=1}^{d} (2 - x_i) x_i \mathrm{I}d.$$

- Solution not bounded by 1, with $C = \frac{1}{2}$

$$u(t,x) = e^{\frac{C}{d} \sum_{i=1}^{d} x_i + \frac{T-t}{2}},$$

- Use monotone interpolator [5]

First 1D case

$$f(t,y) = y\left(\frac{1}{2} - \frac{V^2}{2C^2}[\phi(t,T,y)(2C - \phi(t,T,y))]^2 - U(C - \phi(t,T,y))\right),$$
$$\phi(t,T,y) = \log(y) - \frac{T-t}{2}.$$

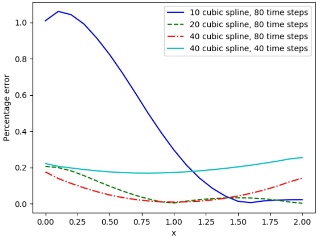


Figure: Cubic spline method.

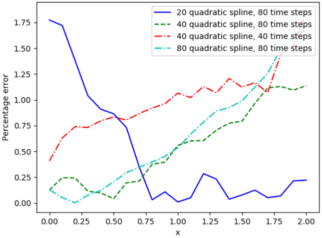


Figure: Quadratic spline method.

Second 1D case

$$
\begin{aligned}
f(x,y) &= f_1(y) + f_2(x),\\
f_1(y) &= \frac{2}{10}\left(y + \sin\left(\frac{\pi}{2} y\right)\right),\\
f_2(x) &= \frac{1}{2} - \left(\frac{2}{10} + C\mu(x)\right) - \frac{\sigma(x)^2 c^2}{2} e^{Cx + \frac{T-t}{2}} - \frac{2}{10}\sin\left(\frac{\pi}{2} e^{cx + \frac{T-t}{2}}\right)
\end{aligned}
$$

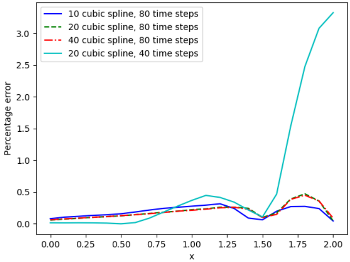


Figure: Cubic spline method.

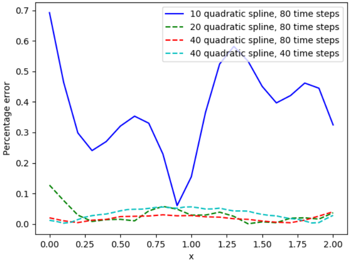


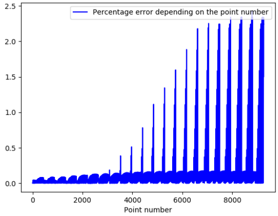
Figure: Quadratic spline method.

Remarks

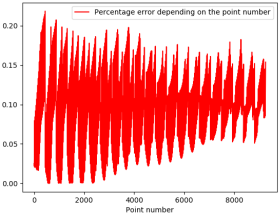
- A small number of splines gives a “large error” :
  - approximation of the driver leads to an error (controlled)
  - large time steps means error on φ_j term : error on the cell meaning large error
- a large number of spline :
  - very small error on the driver
  - larger statistical error on the φ_j term : but an error on the cell number means only use of a polynomial close to the good one.
- high number of time step necessary :
  - limit the variance problem : less Monte Carlo simulation needed meaning less computational time
  - limits the error on φ_j
  - Interpolation error has to be of “second order”

Multidimensional results

$$f_2(x) = \frac{1}{2} - \left(\frac{2}{10} + \frac{C}{d}\sum_{i=1}^{d} x_i\right) - \frac{\sigma_{1,1}(x)^2 c^2}{2d} e^{\frac{C}{d}\sum_{i=1}^{d} x_i + \frac{T-t}{2}} - \frac{2}{10}\sin\left(\frac{\pi}{2} e^{\frac{C}{d}\sum_{i=1}^{d} x_i + \frac{T-t}{2}}\right)$$



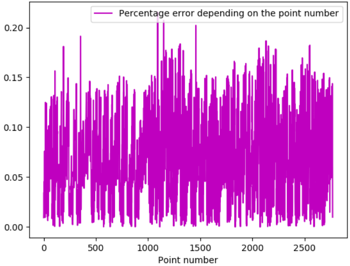
40 splines, 80 time steps.



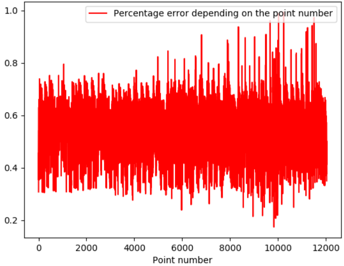
80 splines, 160 time steps

Figure: Error in dimension 3 for different time steps and spline numbers with cubic spline.

Interpolation with sparse grids



4D, 80 splines, 160 time steps.



5D , 80 splines, 160 time steps.

Modified version

For g function bounded by 1 , not to long maturities, small driver coefficients :

for $x \in G_{M-1}$ **do**
    $T_\emptyset = t_{M-1}$
    $$u(t_{M-1}, x) = \mathbb{E}_{M-1,x}\left[\prod_{k \in \mathcal{K}_T} \frac{g(X_T^k)}{\bar{F}(T - T_{k-})} \prod_{k \in \overline{\mathcal{K}}_T \setminus \mathcal{K}_T} \frac{(a_{I(k)}\varphi_j(g))(T_{(k)}, X_{T_k}^{(k)})}{p_{I(k)}\rho(\tau^{(k)})}\right]$$
end for
for $i = M-2, 0$ **do**
    for $x \in G_i$ **do**
        $T_\emptyset = t_i$
        $$u(t_i, x) = \mathbb{E}_{i,x}\left[\prod_{k \in \mathcal{K}_T} \frac{g(X_{T_k}^k)}{\bar{F}(t_{i+1} - T_{k-})} \prod_{k \in \overline{\mathcal{K}}_T \setminus \mathcal{K}_T} \frac{(a_{I(k)}\varphi_j(\hat{I}_{E[\frac{T_{(k)}M}{T}]+1}(u(t_{i+1}, .)))(T_{(k)}, X_{T_k}^{(k)})}{p_{I(k)}\rho(\tau^{(k)})}\right]$$
    end for
end for

Table of contents

General case for first derivative

- $$dX_s^{t,x} \;=\; \mu(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dW_s$$
- Suppose μ, σ continuous, with bounded continuous gradients $D\mu$, $D\sigma$ and σ uniformly elliptic,
- $\phi : \mathbb{R}^d \to \mathbb{R}$ bounded measurable function
- The tangent process is well defined

$$Y_t := \mathbf{I}_d,\; dY_s = D\mu(s, X_s^{t,x})Y_s ds + \sum_{i=1}^{d} D\sigma_i(s, X_s^{t,x})Y_s dW_s^i,\; \text{for } s \in [t, T],\; \mathbb{P}\text{-a.}$$

- We have the automatic differentiation rule :

$$\partial_x \mathbb{E}\left[\phi(X_s^{t,x})\right] \;=\; \mathbb{E}\left[\phi(X_s^{t,x})\frac{1}{s-t}\int_t^s \left[\sigma^{-1}(r, X_r^{t,x})Y_r\right]^{\mathsf{T}} dW_r\right].$$

Case ϕ regular, μ, σ constant, $1D$

$$
\begin{aligned}
\partial_x \mathbb{E}\left[\phi\left(X_s^{t,x}\right)\right] =& \mathbb{E}\left[\phi'\left(X_s^{t,x}\right)\right] \\
=& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi'(x+\mu(s-t)+\sigma\sqrt{s-t}u)e^{-\frac{u^2}{2}} du \\
=& \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(x+\mu(s-t)+\sigma\sqrt{s-t}u)\frac{u}{\sigma\sqrt{s-t}} e^{-\frac{u^2}{2}} du \\
=& E\left[\phi\left(X_s^{t,x}\right)\frac{W_s - W_t}{\sigma(s-t)}\right]
\end{aligned}
$$

- If $(s-t)$ small, high variance ($\approx \frac{C}{s-t}$)

When $s - t$ small

- Variance reduction 1:

$$\partial_x \mathbb{E}[\phi(X_s^{t,x})] = \mathbb{E}\Big[\Big(\phi(X_s^{t,x}) - \phi(x)\Big)\frac{W_s - W_t}{\sigma(s-t)}\Big]$$

- Variance reduction 2 : Define Antithetic :

$$d\bar{X}_s^{t,x} = \mu ds - \sigma dW_s$$

$$\partial_x \mathbb{E}[\phi(X_s^{t,x})] = \frac{1}{2}\mathbb{E}\Big[\Big(\phi(X_s^{t,x}) - \phi(\bar{X}_s^{t,x})\Big)\frac{W_s - W_t}{\sigma(s-t)}\Big]$$

- Variance bounded by $||\phi'||_\infty$ using taylor expansion variance.

Second order derivative

- Suppose μ , σ constant, ϕ regular enough

$$\partial_x^2 \mathbb{E}[\phi(X_s^{t,x})] = E[\phi(X_s^{t,x})\overline{\mathcal{W}}]$$

$$\overline{\mathcal{W}} = (\sigma^{-1})^\top \frac{(W_s - W_t)(W_s - W_t)^\top - (s-t)\mathbb{I}}{(s-t)^2}\sigma^{-1}$$

Proof : double integration by part.

- If $(s-t)$ small , high variance $\approx \frac{1}{(s-t)^2}$,
- Variance reduction :

$$\partial_x^2 \mathbb{E}[\phi(X_s^{t,x})] = E\Big[\Big(\phi(X_s^{t,x}) + \phi(\bar{X}_s^{t,x}) - 2\phi(x)\Big)\frac{\overline{\mathcal{W}}}{2}\Big]$$

Because

$$\phi(X_s^{t,x}) + \phi(\bar{X}_s^{t,x}) - 2\phi(x) \approx \phi''(\xi)(s-t)\overline{\mathcal{W}}$$

First alternative second ordrer scheme

Apply 2 first order derivatives on $[t, \frac{t+s}{2}]$, and $[\frac{t+s}{2}, s]$ with variance reduction

$$\partial_x^2 \mathbb{E}\left[\phi(X_s^{t,x})\right] = \mathbb{E}\left[\psi\left((\sigma^\top)^{-1} \frac{(W_{\frac{t+s}{2}} - W_t)(W_s - W_{\frac{t+s}{2}})^\top}{(s-t)^2} \sigma^{-1}\right)\right]$$

$$\psi = \phi\left(X_s^{t,x}\right) + \phi\left(x + \mu(t-s)\right) - \phi\left(x + \mu(t-s) + \sigma(W_{\frac{t+s}{2}} - W_t)\right) - \phi\left(x + \mu(t-s) + \sigma(W_s - W_{\frac{t+s}{2}})\right)$$

Often more effective [6].

Second alternative second order scheme

Same as before but with Antithetic :

$$\partial_x^2 \mathbb{E}[\phi(X_s^{t,x})] = \mathbb{E}[\psi((\sigma^\top)^{-1}\frac{(W_{\frac{t+s}{2}} - W_t)(W_s - W_{\frac{t+s}{2}})^\top}{(s-t)^2}\sigma^{-1})]$$

$$\psi = \frac{1}{2}[\phi\left(X_s^{t,x}\right) + 2\phi\left(x+\mu(t-s)\right) - \phi\left(x+\mu(t-s)+\sigma(W_{\frac{t+s}{2}} - W_t)\right) -$$
$$\phi\left(x+\mu(t-s)+\sigma(W_s - W_{\frac{t+s}{2}})\right) + \phi\left(\bar{X}_s^{t,x}\right) -$$
$$\phi\left(x+\mu(t-s)-\sigma(W_{\frac{t+s}{2}} - W_t)\right) - \phi\left(x+\mu(t-s)-\sigma(W_s - W_{\frac{t+s}{2}})\right)]$$

Table of contents

Linear problem

Linear problem :

$$\partial_t u + \mathcal{L}u = 0$$
$$u(T, x) = g(x) \ ,$$

$$\begin{cases} dX_t^{0,x} & = & \mu(t, X_t^{0,x})dt + \sigma(t, X_t^{0,x})dW_t, \\ X_0^{0,x} & = & x, \end{cases}$$

$$(\mathcal{L}\varphi)(t,x) = \mu(t,x).D\varphi(t,x) + \frac{1}{2}a(t,x) : D^2\varphi(t,x) \ ,$$
$$a(t,x) := \sigma(t,x)\sigma(t,x)^\top$$

How to solve it without bias (no Euler scheme) with usual condition : μ and a uniformly Lipschitz in space , α Hölder in time

Freezing the coefficient

Operator with coefficient frozen at $(\tilde{t}, \tilde{x})$

$$\mathcal{L}^{\tilde{t},\tilde{x}}\varphi(t,x) = \mu(\tilde{t},\tilde{x}).D\varphi(t,x) + \frac{1}{2}a(\tilde{t},\tilde{x}) : D^2\varphi(t,x) \ ,$$

SDE with frozen coefficients

$$\tilde{X}_t^{\tilde{t},\tilde{x},t_0,x} = x + \mu(\tilde{t},\tilde{x})(t-t_0) + \sigma(\tilde{t},\tilde{x})(W_t - W_{t_0}) \ .$$

Rewriting

$$\partial_t u + \mathcal{L}^{\tilde{t},\tilde{x}} u + H^{\tilde{t},\tilde{x}}(t,x,Du,D^2u) = 0$$

$$H^{\tilde{t},\tilde{x}}(t,x,y,z) = (\mu(t,x) - \mu(\tilde{t},\tilde{x})).y + \frac{1}{2}(a(t,x) - a(\tilde{t},\tilde{x})) : z$$

Feynman Kac for regular u

$$u(t,x) = \mathbb{E}[g(\tilde{X}_T^{\tilde{t},\tilde{x},t,x}) + \int_t^T H^{\tilde{t},\tilde{x}}(s, \tilde{X}_s^{\tilde{t},\tilde{x},t,x}, Du(s,\tilde{X}_s^{\tilde{t},\tilde{x},t,x}), D^2u(s,\tilde{X}_s^{\tilde{t},\tilde{x},t,x}))\, ds]$$

Expression for derivatives

Using Malliavin weights (constant parameters)

$$
\begin{aligned}
Du(t,x) &= \mathbb{E}[g(\tilde{X}_T^{\tilde{t},\tilde{x},t,x})\mathcal{M}_{t,T}^{\tilde{t},\tilde{x}} + \\
&\int_t^T H^{\tilde{t},\tilde{x}}(s, \tilde{X}_s^{\tilde{t},\tilde{x},t,x}, Du(s, \tilde{X}_s^{\tilde{t},\tilde{x},t,x}), D^2u(s, \tilde{X}_s^{\tilde{t},\tilde{x},t,x}))\mathcal{M}_{t,s}^{\tilde{t},\tilde{x}}\, ds] \\
D^2u(t,x) &= \mathbb{E}[g(\tilde{X}_T^{t,x,t,x})\mathcal{V}_{t,T}^{t,x} + \\
&\int_t^T H^{t,x}(s, \tilde{X}_s^{t,x,t,x}, Du(s, \tilde{X}_s^{t,x,t,x}), D^2u(s, \tilde{X}_s^{t,x,t,x}))\mathcal{V}_{t,s}^{t,x}\, ds] \; ,
\end{aligned}
$$

$$
\mathcal{M}_{t,s}^{\tilde{t},\tilde{x}} := (\sigma(\tilde{t},\tilde{x})^{-1})^\top \frac{W_s - W_t}{s-t} \; ,
$$

$$
\mathcal{V}_{t,s}^{\tilde{t},\tilde{x}} := (\sigma(\tilde{t},\tilde{x})^{-1})^\top \frac{(W_s - W_t)(W_s - W_t)^\top - (s-t)\mathbb{I}}{(s-t)^2}\sigma(\tilde{t},\tilde{x})^{-1} \; .
$$

Introducing stochastic mesh

$$\left\{ \begin{array}{lll} T_0 & = & 0 \\ T_{k+1} & = & T_k + \Delta T_{k+1} \text{ , for } k = 0, N_T \text{ where} \\ \Delta T_{k+1} & = & \tau_{k+1} \wedge (T - (T_k + \tau_{k+1}))^+ \text{ ,} \end{array} \right. \quad (5)$$

τ_k i.i.d density ρ , $\bar{F} = 1 - F$, F CDF.
Freezing coefficient between two time steps

$$\left\{ \begin{array}{l} \bar{X}_0 = X^{t_0,x}_{T_0} = x \\ \bar{X}_{k+1} = \bar{X}_k + \mu(T_k, , \bar{X}_k)\Delta T_{k+1} + \sigma(T_k, \bar{X}_k)\Delta W_{k+1} \text{ ,} \end{array} \right.$$

where $\Delta W_{k+1} := W_{T_{k+1}} - W_{T_k}$.

Similar to branching..

$$u(T_k, \bar{X}_k) = \frac{\mathbb{E}[g(\bar{X}_{k+1}) \mathbf{1}_{T_{k+1}=T}]}{\bar{F}(T - T_k)} + \mathbb{E}[H_{k+1} \mathbf{1}_{T_{k+1}<T}]$$

$$H_{k+1} := \frac{H^{T_k, \bar{X}_k}(T_{k+1}, \bar{X}_{k+1}, Du(T_{k+1}, \bar{X}_{k+1}), D^2 u(T_{k+1}, \bar{X}_{k+1}))}{\rho(\Delta T_{k+1})}$$

Need for Du, and D^2U expression to plug in for recursion

$$Du(T_{k+1}, \bar{X}_{k+1}) = \frac{\mathbb{E}[g(\bar{X}_{k+2}) \mathcal{M}_{T_{k+1}, T}^{T_{k+1}, \bar{X}_{k+1}} \mathbf{1}_{T_{k+2}=T}]}{\bar{F}(T - T_{k+1})} + \mathbb{E}[H_{k+2} \mathcal{M}_{T_{k+1}, T_{k+2}}^{T_{k+1}, \bar{X}_{k+1}} \mathbf{1}_{T_{k+2}<T}]$$

$$D^2 u(T_{k+1}, \bar{X}_{k+1}) = \frac{\mathbb{E}[g(\bar{X}_{k+2}) \mathcal{V}_{T_{k+1}, T}^{T_{k+1}, \bar{X}_{k+1}} \mathbf{1}_{T_{k+2}=T}]}{\bar{F}(T - T_{k+1})} + \mathbb{E}[H_{k+2} \mathcal{V}_{T_{k+1}, T_{k+2}}^{T_{k+1}, \bar{X}_{k+1}} \mathbf{1}_{T_{k+2}<T}] .$$

Representation

$$\begin{cases} P_{k+1} & = \dfrac{M_{k+1} + \frac{1}{2} V_{k+1}}{\rho(\Delta T_k)} \,, \\ M_{k+1} & = \Delta\mu_k.(\sigma_k^{-1})^\top \dfrac{\Delta W_{k+1}}{\Delta T_{k+1}}, \quad \text{with} \quad \Delta\mu_k := \mu_k - \mu_{k-1} \\ V_{k+1} & = \Delta a_k : (\sigma_k^{-1})^\top \dfrac{\Delta W_{k+1} \Delta W_{k+1}^\top - \Delta T_{k+1} \mathbb{I}}{(\Delta T_{k+1})^2} \sigma_k^{-1} \,, \quad \text{with} \quad \Delta a_k := a_k - a_{k-1} \,. \end{cases}$$

Using previous equations recursively ($T_{N_T+1} = T$):

$$\begin{aligned} u(0,x) :=& \mathbb{E}[g(X_T^{0,x})] \\ =& \mathbb{E}[\frac{g(\bar{X}_{N_T+1})}{\bar{F}(\Delta T_{N_T+1})} \prod_{k=2}^{N_T+1} P_k] \,, \end{aligned}$$

Second representation with antithetic

Control variate for all gradient weights

$$u(t_0, x_0) = \mathbb{E}[\beta \prod_{k=2}^{N_T} P_k \mathbf{1}_{N_T \geq 1}] + \mathbb{E}[\frac{g(\bar{X}_1)}{\bar{F}(\Delta T_1)} \mathbf{1}_{N_T=0}] \,,$$

where $\beta := \frac{1}{2}(\beta_1 + \beta_2)$ with

$$\begin{cases} \beta_1 & := & \dfrac{g(\bar{X}_{N_T+1}) - g(\bar{X}_{N_T})}{\bar{F}(\Delta T_{N_T+1})} \dfrac{M_{N_T+1} + \frac{1}{2} V_{N_T+1}}{\rho(\Delta T_{N_T})}, \\ \beta_2 & := & \dfrac{g(\hat{X}_{N_T+1}) - g(\bar{X}_{N_T})}{\bar{F}(\Delta T_{N_T+1})} \dfrac{-M_{N_T+1} + \frac{1}{2} V_{N_T+1}}{\rho(\Delta T_{N_T})} \end{cases}$$

Use of control variate necessary for variance issue !

Variance issue (Poisson process)

- As ΔT_k can go to zero do we have variance bounded ?
- Mixing two successive weights , ΔT_k going to 0

$$\Delta a_k \approx O(\Delta T_k^{\frac{1}{2}}),$$
$$\frac{\Delta W_{k+1} \Delta W_{k+1}^\top - \Delta T_{k+1} \mathbb{I}}{(\Delta T_{k+1})^2} \approx O(\Delta T_k^{-1})$$

so in $1D$

$$\frac{\Delta a_k \frac{\Delta W_k \Delta W_k^\top - \Delta T_k \mathbb{I}}{(\Delta T_k)^2}}{\rho(\Delta T_k)} \approx O(\frac{\Delta T_k^{-\frac{1}{2}}}{\rho(\Delta T_k)})$$

- Suppose the branching dates follow a Poisson process :
  - Condition with respect to the number of Branching dates,
  - Conditional law of increment uniform
  - $\left[\frac{\Delta a_k \frac{\Delta W_k \Delta W_k^\top - \Delta T_k \mathbb{I}}{(\Delta T_k)^2}}{\rho(\Delta T_k)}\right]^2 \approx O(\frac{\Delta T_k^{-1}}{\rho(\Delta T_k)^2})$ not integrable

Variance issue : change law for time increments

- Use Gamma law

$$\rho_{\Gamma}^{\kappa,\theta}(s) = \frac{s^{\kappa-1}e^{-s/\theta}}{\Gamma(\kappa)\theta^{\kappa}}, \quad \text{for all } s > 0, \tag{6}$$

  Γ gamma Euler function

- $$\left[\frac{\Delta a_k \frac{\Delta W_k \Delta W_k^{\top} - \Delta T_k \mathbb{I}}{(\Delta T_k)^2}}{\rho(\Delta T_k)}\right]^2 \approx O((\Delta T_k)^{1-2\kappa})$$

  So Sufficient Condition for bounded variance bounded : $\kappa \leq 0.5$
- Rigorous demonstration for bounded variance in [8]
- Variance reduction with interaction particles (“a la Del moral”) in [8].

Results dimension 4

$$\begin{aligned}\sigma(t,x) &= (0.5 + a\min((\sum_{i=1}^{4} x_i)^2, 1))\mathbb{I} \\ g(x) &= (\frac{1}{d}\sum_{i=1}^{d} x_i - 1)^+ \\ \mu(t,x) &= -10 \vee (1-x) \wedge 10 \\ x_0 &= 1 \\ T &= 1\end{aligned}$$

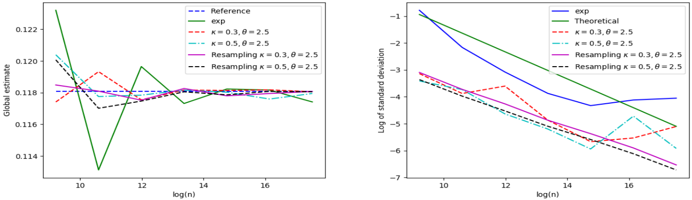


Figure: 4D , $a = 0.4$

Results dimension 4

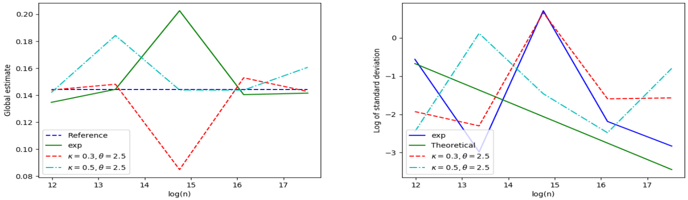


Figure: 4D, $a = 0.6$, no re-sampling

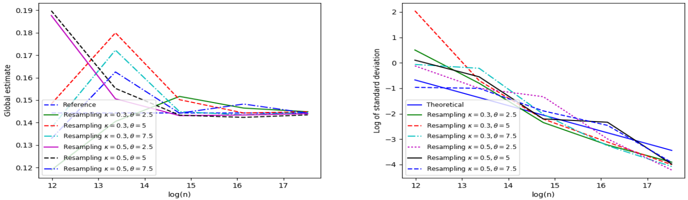


Figure: 4D $a = 0.6$ re-sampling

Conclusion

- Only effective for small maturities, small change in coefficients,
- Permits to avoid time discretization
- Can compete with Euler only for small change in coefficients

Table of contents

An example

$$
\begin{aligned}
-\partial_t u - \mathcal{L}u =& f(u, Du), \\
u_T =& g, \quad t < T, \; x \in \mathbb{R}^d, \\
\mathcal{L}u :=& \frac{1}{2}\Delta u, \\
f(y, z) =& \frac{1}{2}(y^2 + yz).
\end{aligned}
$$

Feynman Kac

$$
\begin{aligned}
u(0,x) =& \mathbb{E}_{0,x}\left[\overline{F}(T)\frac{g(W_T)}{\overline{F}(T)} + \int_0^T \frac{f(u,Du)(t,W_t)}{\rho(t)}\rho(t)dt\right] & (7)\\
=& \mathbb{E}_{0,x}\left[\phi\big(0, T_{(1)}, W^1_{T_{(1)}}\big)\right], & (8)
\end{aligned}
$$

$I^{(1)}$ with values 0 and 1 with equal probability

$$
\phi(s,t,y) := \frac{\mathbf{1}_{\{t \geq T\}}}{\overline{F}(T-s)} g(y) + \frac{\mathbf{1}_{\{t < T\}}}{\rho(t-s)} (uD^{I^{(1)}}u)(t,y).
$$

On the event set $\{I^{(1)} = 0\}$

$$(uD^{I^{(1)}}u)(t,y) = u(t,y)^2 = \mathbb{E}_{t,y}\left[\phi(t, t+\tau^1, W^1_{t+\tau^1})\right]^2.$$

By independence

$$\begin{aligned}(uD^{I^{(1)}}u)(t,y) =& \mathbb{E}_{t,y}\left[\phi\left(t, t+\tau^{(1,1)}, W^{(1,1)}_{t+\tau^{(1,1)}}\right)\right]\mathbb{E}_{t,y}\left[\phi\left(t, t+\tau^{(1,2)}, W^{(1,1)}_{t+\tau^{(1,2)}}\right)\right] \\ =& \mathbb{E}_{t,y}\left[\phi\left(t, t+\tau^{(1,1)}, W^{(1,1)}_{t+\tau^{(1,1)}}\right)\phi\left(t, t+\tau^{(1,2)}, W^{(1,2)}_{t+\tau^{(1,2)}}\right)\right],\end{aligned}$$

On the event set $\{I^{(1)} = 1\}$

$$(uD^{I^{(1)}}u)(t,y) \;=\; \mathbb{E}_{t,y}\big[\phi(t, t+\tau^{(1,1)}, W^{(1,1)}_{t+\tau^{(1,1)}})\big]\partial_y \mathbb{E}_{t,y}\big[\phi(t, t+\tau^{(1,2)}, W^{(1,2)}_{t+\tau^{(1,2)}})\big].$$

Automatic differentiation :

$$\partial_y \mathbb{E}_{t,y}\big[\phi(t, t+\tau^{(1,2)}, W^{(1,2)}_{t+\tau^{(1,2)}})\big] = \mathbb{E}_{t,y}\Big[\frac{W^{(1,2)}_{(t+\tau^{(1,2)})\wedge T} - W^{(1,2)}_t}{\tau^{(1,2)} \wedge (T-t)}\phi(t, t+\tau^{(1,2)}, W^{(1,2)}_{t+\tau^{(1,2)}})\Big],$$

Independance :

$$(uD^{I^{(1)}}u)(t,y) = \mathbb{E}_{t,y}\Big[\frac{W^{(1,2)}_{(t+\tau^{(1,2)})\wedge T} - W^{(1,2)}_t}{\tau^{(1,2)} \wedge (T-t)}\phi(t, \tau^{(1,1)}_t, W^{(1,1)}_{\tau^{(1,1)}_t})\phi(t, \tau^{(1,2)}_t, W^{(1,2)}_{\tau^{(1,2)}_t})\Big],$$

Plugging $uD^{I^{(1)}}u$ into initial (8)

Notation

$$\mathcal{W}^{(1)} := \mathbf{1}_{\{I^{(1)}=0\}} + \mathbf{1}_{\{I^{(1)}=1\}} \frac{\Delta W^{(1,2)}_{T_{(1,2)}}}{\Delta T_{(1,2)}},$$

$$\Delta W^{(1,2)}_{T_{(1,2)}} := W^{(1,2)}_{T_{(1,2)}} - W^{(1,2)}_{T_{(1)}}, \ \Delta T_{(1,2)} := T_{(1,2)} - T_{(1)},$$

so that

$$u(0,x) = \mathbb{E}_{0,x}\Bigg[\mathbf{1}_{\{T_{(1)}=T\}} \frac{g(W_T)}{\overline{F}(T)} + \mathbf{1}_{\{T_{(1)}<T\}} \frac{\mathcal{W}^{(1)}}{\rho(T_{(1)})}$$
$$\prod_{i=1}^{2}\Bigg(\mathbf{1}_{\{T_{(1,i)}=T\}} \frac{g(W^{1,i}_T)}{\overline{F}(\Delta T_{(1,i)})} + \mathbf{1}_{\{T_{(1,i)}<T\}} \frac{(uD^{I^{1,i}}u)\big(T_{(1,i)}, W^{1,i}_{T_{(1,i)}}\big)}{\rho(\Delta T_{(1,i)})}\Bigg)\Bigg].$$

General case

- $\ell = (\ell_0, \ell_1, \cdots, \ell_m) \in L$, $L \subset \mathbb{N}^{m+1}$, $|\ell| := \sum_{i=0}^{m} \ell_i$

$$f(t,x,y,z) := \sum_{\ell=(\ell_0,\ell_1,\cdots,\ell_m)\in L} c_\ell(t,x)\, y^{\ell_0} \prod_{i=1}^{m} \left(b_i(t,x)\cdot z\right)^{\ell_i}.$$

- same Galton Watson tree construction as for $f(u)$
- for a particle k, I_k permits to identify the term to treat in f ; Identify values taken by I_k and element of L
- On the event $I_k = \ell = (\ell_0, \ell_1, \cdots, \ell_m)$, we consider
  $c_\ell(t,x)\, y^{\ell_0} \prod_{i=1}^{m} \left(b_i(t,x)\cdot z\right)^{\ell_i}$
- On the event $I_k = \ell$, $|I_k|$ particles are generated :
  - ℓ_0 are marked 0,
  - ℓ_1 are marked 1 ...

Example marked Galton Watson tree

$$f(t,x,y,z) \;:=\; c_{0,0}(t,x) \;+\; c_{1,0}(t,x)y \;+\; c_{1,1}(t,x)yz$$
$$m = 1, L = \{\bar{\ell}_1 = (1,0), \bar{\ell}_2 = (1,1)\}$$

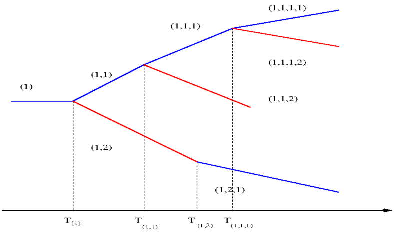


Figure: Galton-Watson tree for KPP

- $T_{(1)}$, (1) branches into two particles $(1,1)$ and $(1,2)$.
- $T_{(1,1)}$, $(1,1)$ branches into $(1,1,1)$ and $(1,1,2)$.
- $T_{(1,2)}$, $(1,2)$ branches into $(1,2,1)$.
- $T_{(1,1,2)}$, $(1,1,2)$ dies out without any offspring particle.
- $T_{(1,1,1)}$, $(1,1,1)$ branches into $(1,1,1,1)$ and $(1,1,1,2)$.
- Particles in blue marked by 0, particles in red marked by 1.

Representation in case σ constant (explicit Malliavin weight to simplify)

$$\mathcal{W}_k := \mathbf{1}_{\{\theta_k=0\}} + \mathbf{1}_{\{\theta_k\neq 0\}} \, b_{\theta_k}(T_{k-}, X^k_{T_{k-}}) \cdot (\sigma_0^\top)^{-1} \frac{\Delta W_k}{\Delta T_k}$$

Weight for u term Du term

$$\psi := \Big[\prod_{k\in\mathcal{K}_T} \frac{g(X^k_T) - g(X^k_{T_{k-}})\mathbf{1}_{\{\theta_k\neq 0\}}}{\overline{F}(\Delta T_k)} \mathcal{W}_k \Big] \Big[\prod_{k\in\overline{\mathcal{K}}_T\setminus\mathcal{K}_T} \frac{c_{I_k}(T_k, X^k_{T_k})}{p_{I_k}} \frac{\mathcal{W}_k}{\rho(\Delta T_k)} \Big]$$

Necessary variance reduction for Du term when reaching T

$$u(0,x) = \mathbb{E}\Big[\psi\Big]$$

Variance consideration

- Suppose
  - $(p_\ell)_{\ell \in L}$ satisfies $p_\ell > 0$ for all $\ell \in L$, and $\sum_{\ell \in L} |\ell| \, p_\ell < \infty$.
  - $\rho(t) \geq C t^{-\frac{q}{2(q-1)}}$ with $q \in (2, \infty)$
  - μ, σ bounded continuous, bounded continuous partial gradients $D\mu$, $D\sigma$, σ is uniformly elliptic.
  - $c_\ell : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ and $b_i : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ bounded continuous and some integration conditions on c_ℓ.
- Then $E(\psi) < \infty$,
- Then $E(\psi^2) < \infty$:
  - Consider ψ^s, $s \geq 2$ and bound its coefficients
  - Show that for one s the representation with bounded coefficient corresponds to the branching representation $\hat{\phi}$ associated an EDO with a solution v bounded,
  - Integrability gives that $E[|\psi^s|] < \mathbb{E}[|\hat{\phi}|] < \infty$
- Convergence towards the viscosity solution.

In practice

Finite variance if :

- small coefficients , small maturities,
- ρ can be chosen as a gamma law with $\kappa < 0.5$.

Small time steps have a high probability meaning sometimes a high number of weights

Variance intuition when gamma law

- Term in product when *Du* term : $\frac{\Delta W_k}{\rho(\Delta T_k)\Delta T_k} \approx C(\Delta T_k)^{0.5-\kappa}$
- Variance bounded for all laws of the branching date distribution conditionally to the number of branching if $\kappa \leq \frac{1}{2}$
- For Full Non Linear, second ordrer Malliavin term

$$\frac{(\Delta W_k)(\Delta W_k)^\top - \Delta T_k \mathbb{I}}{\rho(\Delta T_k)(\Delta T_k)^2} \approx \frac{C}{\rho(\Delta T_k)\Delta T_k}$$

Integrability problem.

Test case

- Gamma law with $\kappa = 0.5$ and $\theta = 2.5$
- $$f(t,x,y,z) = k(t,x) + cy(b \cdot z)$$
  $$k(t,x) := \cos(x_1 + \cdots + x_d)\Big(\alpha + \frac{\sigma^2}{2} + c\sin(x_1 + \cdots + x_d)\frac{3d+1}{2d}e^{\alpha(T-t)}\Big)e^{\alpha(T-t)}$$
- $\alpha = 0.2$, $c = 0.15$, $T = 1$, $x_0 = 0.5\mathbb{1}_d$
- Small non linearity decreasing with the dimension, g bounded by one.
- Solution

$$u(t,x) = \cos(x_1 + \cdots + x_d)e^{\alpha(T-t)}.$$

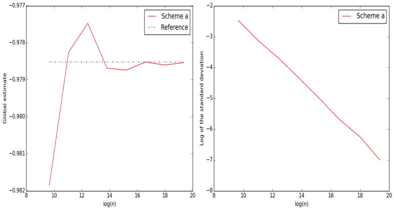
Linear versus non linear results

Non linearity has an impact on solution :

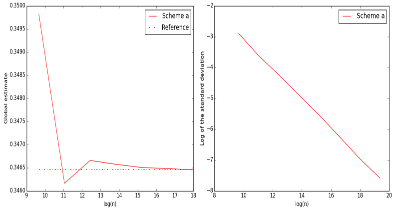
Dimension	5	10	20
Linear Solution	-1.0436	0.3106	-0.9661
Non linear solution	-0.97851	0.34646	-1.0248

Table: Analytic solution linear PDE versus analytic solution for the semi-linear PDE in $d = 5, 10$ and 20.

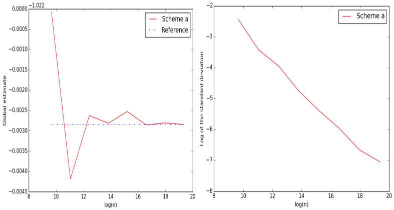
Results



Estimation and standard deviation $d = 5$.



Estimation and standard deviation $d = 10$.



- Estimation and standard deviation $d = 20$

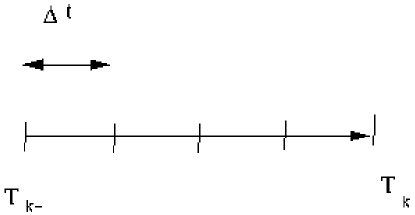
Table of contents

Remark on branching with Gamma laws

- Gamma laws permits to get finite variance methods,
- κ should be taken below 0.5 so high number of small jump :
  - Computation time important,
  - High number of weights terms meaning quite high variance,
- Use of the following ghost method permits to deal with longer maturities with less computation cost.
- Possibility to use nesting : each conditional expectation estimated with a few particles.
- No proof of convergence with ghost even for low maturity ... but we are sure we have integrability and finite variance.

When coupled to a Euler scheme

- Malliavin can be use by integration by part on first step with size Δt



- Gradient weight

$$b_{\theta_k}(T_{k-}, X^k_{T_{k-}}) \cdot (\sigma_0^\top)^{-1} \frac{W_{(T_k+\Delta t)\wedge T_{k+1}} - W_{T_k}}{\Delta T_k \wedge \Delta t}$$

- Variance explodes when taking a small time step

Burgers without ghost

$$u(0,x) = \mathbb{E}_{0,x}\left[\phi(T_{(1)}, X^{(1)}_{T_{(1)}})\right]$$

$$\phi(t,y) := \frac{\mathbf{1}_{\{t \geq T\}}}{\bar{F}(T)} g(y) + \frac{\mathbf{1}_{\{t < T\}}}{\rho(t)} (buDu)(t,y).$$

- On $\{\mathbf{1}_{\{T_{(1)} \geq T\}}\}$ just compute $\dfrac{g(X_T)}{\bar{F}(T)}$,
- On $\{\mathbf{1}_{\{T_{(1)} < T\}}\}$,

$$\frac{buDu(T_{(1)}, X_{T_{(1)}})}{\rho(T_{(1)})} = \frac{b}{\rho(T_{(1)})} \mathbb{E}_{T_{(1)}, X_{T_{(1)}}}\left[\phi(T_{(1,1)}, X^{(1,1)}_{T_{(1,1)}})\right]$$

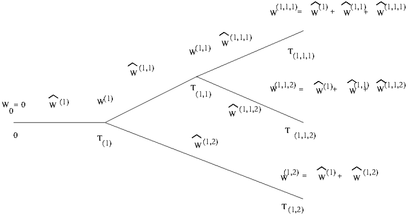
$$\mathbb{E}_{T_{(1)}, X_{T_{(1)}}}\left[\frac{\hat{W}^{(1,2)}_{\Delta T_{(1,2)}}}{\sigma_0 \Delta T_{(1,2)}} \phi(T_{(1,2)}, X^{(1,2)}_{T_{(1,2)}})\right]$$

- Generate 2 particles $(1,1)$ marked $\theta((1,1)) = 0$ and $(1,2)$ marked $\theta((1,2)) = 1$

Re-normalization Labordère et al. [9]

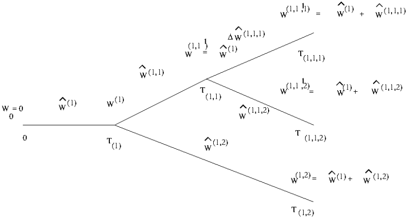
- For gradient term :
$$\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[\frac{\hat{W}^{(1,p)}_{\Delta T_{(1,p)}}}{\sigma_0 \Delta T_{(1,p)}}\left(\phi\left(T_{(1,p)}, X^{(1,p)}_{T_{(1,p)}}\right) - \phi\left(T_{(1,p)}, X^{(1,p^1)}_{T_{(1,p)}}\right)\right)\right],$$
$p = 1, 2$
- $X^{(1,p^1)}$ has the same past as $X^{(1,p)}$ at date $T_{(1)}$,
same future increments between $T_{(1,p)}$ and T ,
no brownian increment between $T_{(1)}$ and $T_{(1,p)}$
- Acts as a control variate.
- $\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[\left(\phi\left(T_{(1,p)}, X^{(1,p)}_{T_{(1,p)}}\right) - \phi\left(T_{(1,p)}, X^{(1,p^1)}_{T_{(1,p)}}\right)\right)^2\right] = O(\Delta T_{(1,p)})$.
- Permits to use all ρ densities (so exponential); finite variance in the linear case. No current result in the semi linear one.
- This ghost method outperforms the original method.

Original Galton-Watson tree and the ghost particles associated for the Brownian.



(a) Original Galton-Watson tree

$W^{(1)} = \hat{W}^{(1)}$
$W^{(1,1)} = \hat{W}^{(1)} + \hat{W}^{(1,1)}$
$W^{(1,2)} = \hat{W}^{(1)} + \hat{W}^{(1,2)}$
$W^{(1,1,1)} = \hat{W}^{(1)} + \hat{W}^{(1,1)} + \hat{W}^{(1,1,1)}$
$W^{(1,1,2)} = \hat{W}^{(1)} + \hat{W}^{(1,1)} + \hat{W}^{(1,1,2)}$



(b) Tree with ghost particle $k = (1, 1^1)$

$W^{(1)} = \hat{W}^{(1)}$
$W^{(1,1^1)} = \hat{W}^{(1)}$
$W^{(1,2)} = \hat{W}^{(1)} + \hat{W}^{(1,2)}$
$W^{(1,1^1,1)} = \hat{W}^{(1)} + \hat{W}^{(1,1,1)}$
$W^{(1,1^1,2)} = \hat{W}^{(1)} + \hat{W}^{(1,1,2)}$

Original re-normalization for burgers Labordère et al. [9]

Backward recursion :

$$\widehat{\psi}_k := \frac{g(X_T^k)}{\overline{F}(\Delta T_k)} \text{ if } T_k = T$$

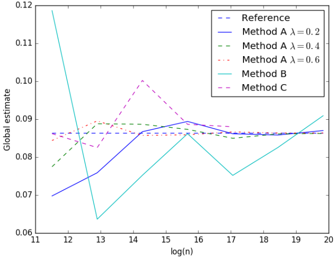
$$\widehat{\psi}_k := \frac{b}{\rho(\Delta T_k)} \prod_{\tilde{k}=\{(k,1),(k,2)\}} \left(\widehat{\psi}_{\tilde{k}} - \widehat{\psi}_{\tilde{k}^1} \mathbf{1}_{\{\theta(\tilde{k}) \neq 0\}}\right) \mathcal{W}_{\tilde{k}}, \quad \text{if } T_k < T$$

$$u(0,x) = \mathbb{E}_{0,x}\left[\widehat{\psi}_{(1)}\right].$$

Re-normalization with antithetic ghosts Warin [11]

- $$\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[(\sigma_0^\top)^{-1}\frac{\hat{W}^{(1,p)}_{\Delta T_{(1,p)}}}{\Delta T_{(1,p)}}\frac{1}{2}\left(\phi\left(T_{(1,p)},X^{(1,p)}_{T_{(1,p)}}\right)-\phi\left(T_{(1,p)},X^{(1,p^1)}_{T_{(1,p)}}\right)\right)\right].$$
- $X^{(1,p^1)}$ has the same past as $X^{(1,p)}$ at date $T_{(1)}$, same future increments between $T_{(1,p)}$ and T and $-\hat{W}^{(1,p)}_{\Delta T_k}$ increment between $T_{(1)}$ and $T_{(1,p)}$.
- Finite variance in the linear case.

Gamma without ghost versus exponential law with original ghost Labordère et al. [9]



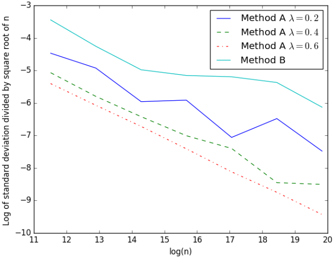


Figure: Analytical case : Estimation, error in $d = 3$, $c = 0.2$, $T = 1$ on the semilinear case depending on the log of the number of particles using a non linearity $cu(Du.b)$, $b := \frac{1}{d}(1 + \frac{1}{d}, 1 + \frac{2}{d}, \cdots, 2)$

Numerical original ghost Labordère et al. [9] versus antithetic ghosts Warin [11] for u calculation.

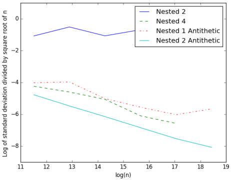


Figure: Error in $d = 6$ $T = 3$ for Burgers

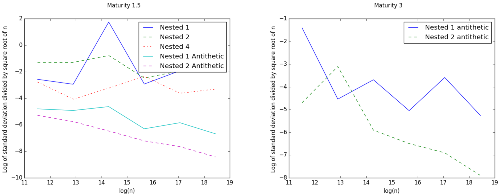


Figure: Error in $d = 6$ for $(Du)^2$ non linearity.

Numerical original ghost versus antithetic ghosts for *Du* calculation.

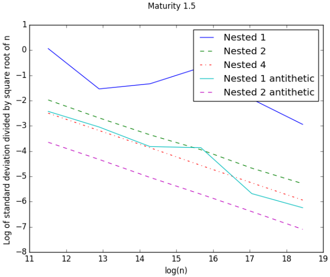


Figure: Error in $d = 6$ for the term $b.Du$ on Burgers test case for $T = 1.5$.

Table of contents

Full non linear $f(u, Du, D^2u) = bu^{l_0}(Du)^{l_1}(D^2u)^{l_2}$: original scheme with 2 ghosts Labordère et al. [9]

$$D^2\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[\phi(T_{(1,p)}, X^{(1,p)}_{T_{(1,p)}})\right] = \mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[(\sigma_0)^{-2}\frac{(\hat{W}^{(1,p)}_{\Delta T_{(1,p)}})^2 - \Delta T_{(1,p)}}{(\Delta T_{(1,p)})^2}\psi\right],$$

$$\psi = \frac{1}{2}\left[\phi(T_{(1,p)}, X^{(1,p)}_{T_{(1,p)}}) + \phi(T_{(1,p)}, X^{(1,p^1)}_{T_{(1,p)}}) - 2\phi(T_{(1,p)}, X^{(1,p^2)}_{T_{(1,p)}})\right].$$

- $X^{(1,p)}_{T_{(1,p)}}$ the original particle
- $X^{(1,p^1)}_{T_{(1,p)}}$ ghost with $-\hat{W}^{(1,p)}_{\Delta T_k}$ increment between $T_{(1)}$ and $T_{(1,p)}$
- $X^{(1,p^2)}_{T_{(1,p)}}$ ghost without increment between $T_{(1)}$ and $T_{(1,p)}$

Finite variance in the linear case (f linear in $D2u$)

- $\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[(\psi)^2\right] = O(\Delta T^2_{(1,p)})$,
- The variance of the scheme is finite for small maturities , small coefficients,
- No current proof for the full non linear case.

A first new scheme for Full Non Linear with 3 ghosts Warin [11]

- Use first order derivative weights on two successive time steps $\frac{\Delta T_{(1,p)}}{2}$.
- $(\hat{W}^{k,i})_{k=(k_1,\cdots,k_{n-1},k_n)\in\mathbb{N}^n, n>1, i=1,2}$ independent BM
- $$D^2\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[\phi(T_{(1,p)},X^{(1,p)}_{T_{(1,p)}})\right] = \mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[2(\sigma_0)^{-2}\frac{\hat{W}^{(1,p),1}_{\Delta T_{(1,p)}}}{\Delta T_{(1,p)}}\frac{(\hat{W}^{(1,p),2}_{\Delta T_{(1,p)}})}{\Delta T_{(1,p)}}\psi\right],$$
  $$\psi = \phi(T_{(1,p)},X^{(1,p)}_{T_{(1,p)}}) + \phi(T_{(1,p)},X^{(1,p^3)}_{T_{(1,p)}}) - \phi(T_{(1,p)},X^{(1,p^1)}_{T_{(1,p)}}) - \phi(T_{(1,p)},X^{(1,p^2)}_{T_{(1,p)}}).$$
- $X^{(1,p)} = X^{(1)} + \mu\Delta T_{(1,p)} + \sigma_0\left[\frac{\hat{W}^{(1,p),1}_{\Delta T_{(1,p)}}) + \hat{W}^{(1,p),2}_{\Delta T_{(1,p)}})}{\sqrt{2}}\right]$
- $X^{(1,p^3)} = X^{(1)} + \mu\Delta T_{(1,p)}$ ghost freezing position
- $X^{(1,p^1)} = X^{(1)} + \mu\Delta T_{(1,p)} + \sigma_0\frac{\hat{W}^{(1,p),1}_{\Delta T_{(1,p)}}}{\sqrt{2}}$ ghost without second $\hat{W}$ increment
- $X^{(1,p^2)} = X^{(1)} + \mu\Delta T_{(1,p)} + \sigma_0\frac{\hat{W}^{(1,p),2}_{\Delta T_{(1,p)}}}{\sqrt{2}}$ ghost without first $\hat{W}$ increment

Remark and extension

- Bounds on variance calculation indicate a potential smaller variance value of the new scheme,
- An antithetic ghost version of the second scheme with 7 ghosts can be used.
- Higher number of ghosts means higher memory requirement.
- Higher derivatives are easy to treat.

Results for full non linearity uD^2u

$$f(u, Du, D^2u) = h(t,x) + \frac{0.1}{d} u(\mathbb{1} : D^2u),$$
$$\mu = 0.2\mathbb{1}\sigma_0 = 0.5\mathbb{1}, \qquad \alpha = 0.2$$

$$h(t,x) = (\alpha + \frac{\sigma_0^2}{2}) \cos(x_1 + .. + x_d) e^{\alpha(T-t)} + 0.1 \cos(x_1 + .. + x_d)^2 e^{2\alpha(T-t)} + \mu \sin(x_1 + .. + x_d) e^{\alpha(T-t)},$$

$$u(t,x) = \cos(x_1 + .. + x_d) e^{\alpha(T-t)}.$$

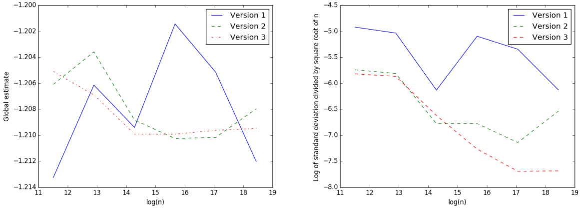


Figure: Solution $u(0, 0.5)$ obtained and error in $d = 6$ with $T = 1$, analytic solution is -1.20918.

Results for full non linearity uD^2u : derivative

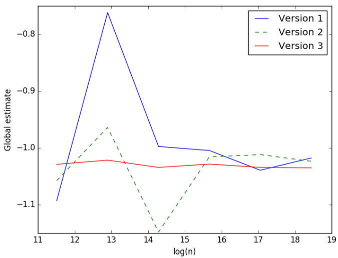


Figure: Derivative ($\mathbf{1}.Du$) obtained and error in $d = 6$ with $T = 1$.

Results for non linearity DuD^2u

$$f(u, Du, D^2u) = 0.0125(\mathbf{1}.Du)(\mathbb{1} : D^2u).$$

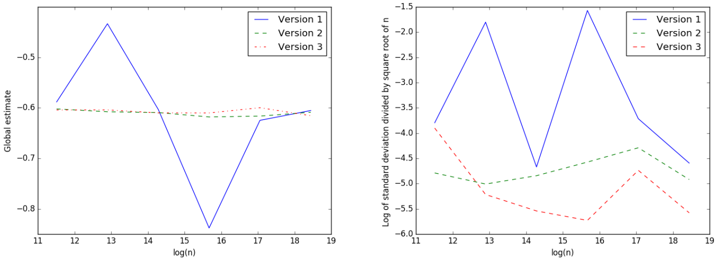


Figure: Solution $u(0, 0.5)$ and error obtained for $d = 4$ with $T = 1$.

Henry P McKean. "Application of brownian motion to the equation of kolmogorov-petrovskii-piskunov". In: *Communications on Pure and Applied Mathematics* 28.3 (1975), pp. 323–331.

Pierre Henry-Labordere. "Counterparty Risk Valuation: A Marked Branching Diffusion Approach (January 30, 2012)". In: *Available at SSRN 1995503* ().

Bruno Bouchard et al. "Numerical approximation of BSDEs using local polynomial drivers and branching processes". In: *arXiv preprint arXiv:1612.06790* (2017).

Kristian Debrabant and Espen Jakobsen. "Semi-Lagrangian schemes for linear and fully non-linear diffusion equations". In: *Mathematics of Computation* 82.283 (2013), pp. 1433–1462.

Xavier Warin. "Some non monotone schemes for time dependent Hamilton-Jacobi-Bellman equations in stochastic control". In: *arXiv preprint arXiv:1310.6121* (2013).

Arash Fahim, Nizar Touzi, and Xavier Warin. "A probabilistic numerical method for fully nonlinear parabolic PDEs". In: *The Annals of Applied Probability* (2011), pp. 1322–1364.

Pierre Henry-Labordere, Xiaolu Tan, and Nizar Touzi. "Unbiased simulation of stochastic differential equations". In: *arXiv preprint arXiv:1504.06107* (2015).

Mahamadou Doumbia, Nadia Oudjane, and Xavier Warin. *Unbiased Monte Carlo estimate of stochastic differential equations expectations*. 2016. eprint: `arXiv:1601.03139`.

Pierre Henri Labordère et al. *Truncation and renormalization techniques for solving the nonlinear PDEs by branching processes*. 2017. eprint: `arXiv:tocome`.

Xavier Warin. "Variations on branching methods for non linear PDEs". In: *arXiv preprint arXiv:1701.07660* (2017).

Xavier Warin. *Variation of branching methods for non linear PDEs.*. 2017. eprint: `arXiv:tocome`.