Variations on branching methods for non linear PDEs

P. H. Labordère & N. Touzi & X. Tan & Xavier Warin

January 11, 2017

Context

Try to solve Semi Linear

$$-\partial_t u - \mathcal{L}u = f(u), u_T = g, \qquad t < T,\ x \in \mathbb{R}^d,$$

Or Full non linear

$$-\partial_t u - \mathcal{L}u = f(u, Du, D^2u), u_T = g, \qquad t < T,\ x \in \mathbb{R}^d,$$

with

- f polynomial in u, Du, D^2u,
- g Lipschitz, coefficients bounded continuous.
- Deterministic methods (Finite Difference...) suffers highly from curse of dimentionality,
- Probabilitic methods such to solve BSDE (Semi Linear), SOBSDE (Full linear) can be based on regression, so suffer (to a less extend) from curse of dimensionality.

Branching dates and particle trajectories (burgers)

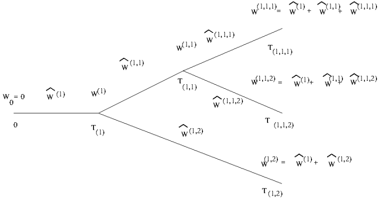


Figure: Galton-Watson tree : brownian for Burgers

- $(\tau^k)_{k=(k_1,..,k_n),k_i \in \{1,2\},n>1}$ iid rv with density ρ,
- Sequence of branching dates $T_{k=(k_1,..,k_n)} = T_{k-=(k_1,..,k_{n-1})} + \tau^{(k_1,..,k_n)} \wedge T$,
- $\Delta T_k = T_k - T_{k-}$ the time increments
- $(\hat{W}^k)_{k=(k_1,\cdots,k_{n-1},k_n)\in\{1,2\}^n,n>1}$ independent Brownian motion
- Each particle $k = (k_1, \cdots, k_{n-1}, k_n) \in \{1,2\}^n, n > 1$ equiped with a brownian with increments

$$W^k_{T_k} \;:=\; W^{k-}_{T_{k-}} + \hat{W}^k_{\Delta T_k}$$

- Particule k position : $X^k_t := x + \mu t + \sigma_0 W^k_t$ for $t \in [T_{k-}, T_k]$

Original branching methods for semi linear (Henry-Labordere et al. [1]) (exemple Burgers) in 1D

- PDE : $\partial_t u + \frac{1}{2}\sigma_0^2 D^2 u + \mu D u + b u D u = 0$,
- Using Feyman-Kac:

$$u(0,x) = \mathbb{E}_{0,x}\Big[\bar{F}(T)\frac{g(X_T)}{\bar{F}(T)} + \int_0^T \frac{buDu(t,X_t)}{\rho(t)}\rho(t)dt\Big]$$
$$= \mathbb{E}_{0,x}\big[\phi\big(T_{(1)}, X^{(1)}_{T_{(1)}}\big)\big]$$

with $\bar{F}(t) := \int_t^\infty \rho(s)ds$ complementary CDF of $\tau_{(1)}$,

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

Original algorithm for semi linear

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

Original algorithm for semi linear

- On $\{\mathbf{1}_{\{T_{(1)}<T\}}\}$,

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

$$D\mathbb{E}_{T_{(1)}, X_{T_{(1)}}} \left[\phi\big(T_{(1,2)}, X^{(1,p)}_{T_{(1,2)}}\big) \right]$$

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

Original algorithm for semi linear

- On $\{\mathbf{1}_{\{T_{(1)}<T\}}\}$,

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

$$D\mathbb{E}_{T_{(1)}, X_{T_{(1)}}}\left[\phi\left(T_{(1,2)}, X^{(1,p)}_{T_{(1,2)}}\right)\right]$$

- Use automatic differentiation :

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

Backward recursion for semi linear

- $$\widehat{\psi}_k := \frac{g(X_T^k) - g(X_{T_{k-}}^k)\,\mathbf{1}_{\{\theta_k \neq 0\}}}{\overline{F}(\Delta T_k)} \text{ if } T_k = T,$$

  Control variate

- $$\widehat{\psi}_k := \frac{b}{\rho(\Delta T_k)} \prod_{\tilde{k} \in \{(k,1),(k,2)\}} \widehat{\psi}_{\tilde{k}} \mathcal{W}_{\tilde{k}}, \quad \text{if } T_k \neq T$$

  where

  $$\mathcal{W}_k = \mathbf{1}_{\{\theta_k = 0\}} + \mathbf{1}_{\{\theta_k \neq 0\}} \frac{(\sigma_0)^{-1} \hat{W}_{\Delta T_k}^k}{\Delta T_k}.$$

  Weight for u term Du term

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

Finite variance requirement

- $\dfrac{1}{x\rho(x)^2} = O(x^\alpha)$ as $x \longrightarrow 0$ with $\alpha \geq 0$. Use Gamma laws parameters $\kappa \leq 0.5$,
- Implies a lot of branching and a higher computational cost.
- only finite for small coefficients and small maturities.

Mofication to handle longer maturities :

- Use exponential law for u terms, use gamma law for Du term.
- Compute conditional expectation with 2 or 4 particles.

Results on burgers type equation dimension 4

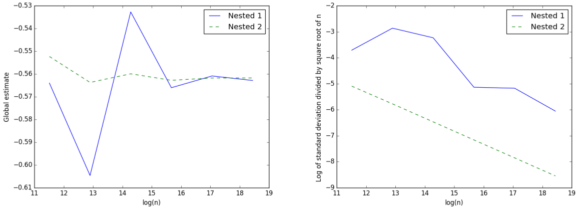


Figure: Estimation and error in $d = 4$. Maturity $T = 1.5$

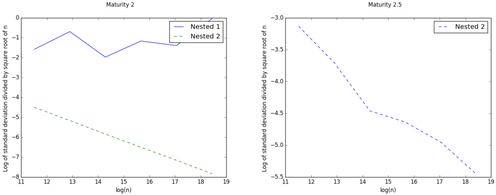
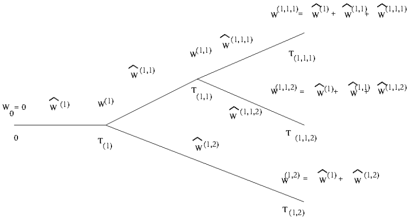


Figure: Error in $d = 4$, maturity $T = 2.$, $T = 2.5$

Re-normalization Labordère et al. [2]

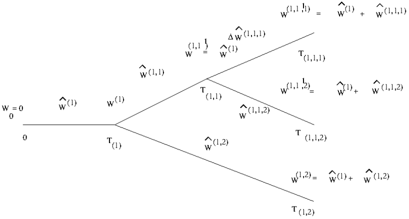
- For gradient term :
  $$\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\Big[\frac{\hat{W}^{(1,p)}_{\Delta T_{(1,p)}}}{\sigma_0 \Delta T_{(1,p)}}\Big(\phi\big(T_{(1,p)}, X^{(1,p)}_{T_{(1,p)}}\big) - \phi\big(T_{(1,p)}, X^{(1,p^1)}_{T_{(1,p)}}\big)\Big)\Big],$$
  $p = 1, 2$
- $X^{(1,p^1)}$ has the same past as $X^{(1,p)}$ at date $T_{(1)}$, same increment between $T_{(1,p)}$ and T, no brownian increment between $T_{(1)}$ and $T_{(1,p)}$
- Acts as a control variate.
- $\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\Big[\Big(\phi\big(T_{(1,p)}, X^{(1,p)}_{T_{(1,p)}}\big) - \phi\big(T_{(1,p)}, X^{(1,p^1)}_{T_{(1,p)}}\big)\Big)^2\Big] = O(\Delta T_{(1,p)})$.
- 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



(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. [2]

$$
\begin{aligned}
\widehat{\psi}_k &:= \frac{g(X_T^k)}{\overline{F}(\Delta T_k)} \text{ if } T_k = T \\
\widehat{\psi}_k &:= \frac{b}{\rho(\Delta T_k)} \prod_{\tilde{k}=\{(k,1),(k,2)\}} (\widehat{\psi}_{\tilde{k}} - \widehat{\psi}_{\tilde{k}^1} \mathbf{1}_{\{\theta(\tilde{k}) \neq 0\}}) \mathcal{W}_{\tilde{k}}, \quad \text{if } T_k < T \\
u(0,x) &= \mathbb{E}_{0,x}\left[\widehat{\psi}_{(1)}\right].
\end{aligned}
$$

Re-normalization with antithetic ghosts Warin [3]

- $$\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\Big[(\sigma_0^\top)^{-1}\frac{\hat{W}^{(1,p)}_{\Delta T_{(1,p)}}}{\Delta T_{(1,p)}}\frac{1}{2}\big(\phi\big(T_{(1,p)},X^{(1,p)}_{T_{(1,p)}}\big)-\phi\big(T_{(1,p)},X^{(1,p^1)}_{T_{(1,p)}}\big)\big)\Big].$$
- $X^{(1,p^1)}$ has the same past as $X^{(1,p)}$ at date $T_{(1)}$, same increment 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.

Numerical original ghost versus antithetic ghosts for u calculation.

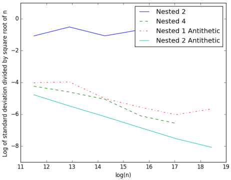


Figure: Error in $d = 6$ $T = 3$ for Burgers

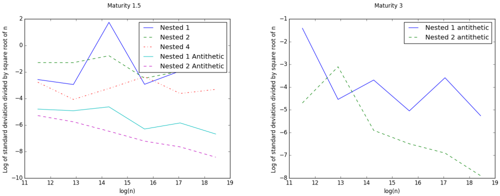


Figure: Error in $d = 6$ for $(Du)^2$ non linearity.

Numerical original ghost versus antithetic ghosts for Du calculation.

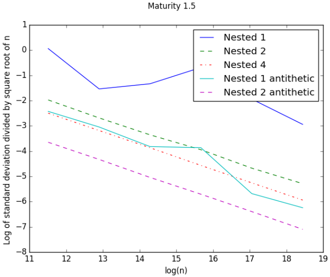


Figure: Error in $d = 6$ for the term $b.Du$ on Burgers test case for $T = 1.5$.

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. [2]

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

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

- $X^{(1,p)}_{T_{(1,p)}}$ 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

- $\mathbb{E}_{T_{(1)},X_{T_{(1)}}}\left[(\psi)^2\right] = O(\Delta T_{(1,p)}^2)$,
- 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 [3]

- 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
-

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

$$\psi = \phi(T_{(1,p)},X^{(1,p)}_{T_{(1,p)}}) + \phi(T_{(1,p)},X^{(1,p^3)}_{T_{(1,p)}}) - \phi(T_{(1,p)},X^{(1,p^1)}_{T_{(1,p)}}) - \phi(T_{(1,p)},X^{(1,p^2)}_{T_{(1,p)}}).$$

- $X^{(1,p)} = X^{(1)} + \mu\Delta T_{(1,p)} + \sigma_0\left[\frac{\hat{W}^{(1,p),1}_{\Delta T_{(1,p)}}) + \hat{W}^{(1,p),2}_{\Delta T_{(1,p)}}}{\sqrt{2}}\right]$
- $X^{(1,p^3)} = X^{(1)} + \mu\Delta T_{(1,p)}$ 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

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

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

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

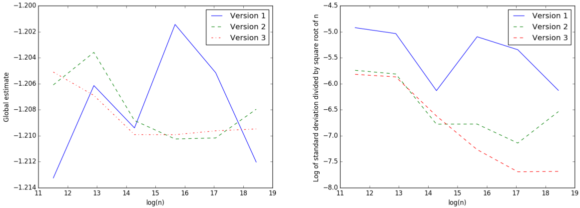


Figure: Solution $u(0, 0.5)$ obtained and error in $d = 6$ with $T = 1$, analytical solution is -1.20918.

Results for full non linearity uD^2u : derivative

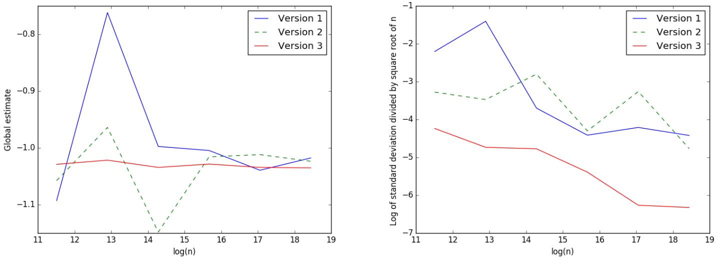


Figure: Derivative $(\mathbf{1}.Du)$ obtained and error in $d = 6$ with $T = 1$.

Results for non linearity DuD^2u

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

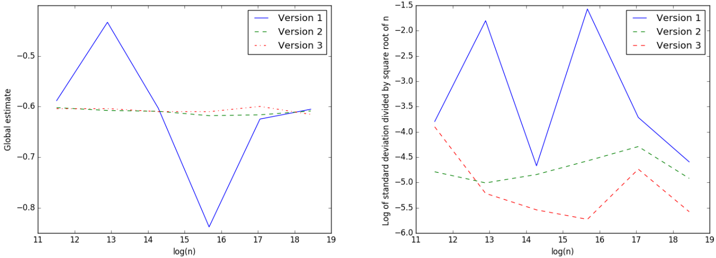


Figure: Solution $u(0, 0.5)$ and error obtained for $d = 4$ with $T = 1$.

Pierre Henry-Labordere et al. "Branching diffusion representation of semilinear PDEs and Monte Carlo approximation". In: *arXiv preprint arXiv:1603.01727* (2016).

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

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