Exact discretization of SDEs

Nizar Touzi
Ecole Polytechnique, Paris

Joint work with Pierre Henry-Labordère and Xiaolu Tan

Steve's 65th Birthday

Carnegie Mellon University, June 2, 2015

Outline

Weak approximation of SDEs

Throughout this paper, objective is to approximate :

$$V_0 := \mathbb{E}\big[g(X_T)\big]$$

where X is solution of the SDE

$$dX_t \quad = \quad \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

- W is a Brownian motion
- μ and σ satisfy the Lipschitz bounded, σ^{-1} bounded
- more conditions on μ and σ will pop up

Standard method

1) discrete-time approximation

- Euler : $\pi := \{0 = t_0 < \ldots < t_n = T\}$ with $h := |\pi|$, and

$$X^{\pi}_{t_i} = X^{\pi}_{t_{i-1}} + \mu(t_i, X^{\pi}_{t_{i-1}})\Delta t + \sigma(t_i, X^{\pi}_{t_{i-1}})\Delta W_{t_i}, \quad i = 1, \ldots, n$$

strong error of order $\sqrt{h}$, weak error of order h

- Higher order discretization schemes... $\Longrightarrow$ weak error $\sim h^{\alpha}$

2) Monte Carlo approximation : Let $\left\{X^{\pi^{(i)}}_T\right\}_{1 \le i \le S}$ iid $\sim X^{\pi}_T$,

$$V_0^{h,S} \quad := \quad \frac{1}{S}\sum_{i=1}^{S} g\left(X^{\pi^{(i)}}_T\right)$$

Central limit theorem $\Longrightarrow$ statistical error $S^{-\frac{1}{2}}$

Avoiding discretization error

$\sigma = 0 \Longrightarrow$ ODE : in general, NO WAY to avoid discretization error

Beskos & Roberts : 1-dim homogeneous SDE with $\sigma > 0$

- Use Lamperti's transformation to convert the SDE to

$$dY_t = b(Y_t)dt + dW_t, \quad Y = f(X), \quad f(x) := \int_0^x \frac{d\xi}{\sigma(\xi)}$$

- Then $V_0 = \mathbb{E}\big[g(X_T)\big] = \mathbb{E}\big[g \circ f^{-1}(Y_T)\big]$, and by Girsanov :

$$V_0 = \mathbb{E}\big[Z\, g \circ f^{-1}(W_T)\big] \quad \text{with} \quad Z := e^{\int_0^T b(W_t)dW_t - \frac{1}{2}\int_0^T b(W_t)^2 dt}$$

- Rejection sampling technique to avoid discretization error for simulation of Z

More references

Exploiting further the rejection sampling technique

$\varepsilon-$strong simulation of multi-dimensional SDEs

Chen & Huang, Beskos '13, Peluchetti & Roberts '12, Pollock, Johansen & G. Roberts '14, Bayer, Friz, Riedel & Schoenmakers '13, Blanchet, Chen & Dong '14

Our algorithm in the case of constant diffusion $\sigma = I_d$ (I)

- (N_t) : Poisson process with intensity β, arrival times $(\tau_i)_{i\geq 1}$
- Set $\tau_0 := 0$, $T_i := \tau_i \wedge T$, and

$$\Delta T_i := T_i - T_{i-1}, \qquad \Delta W_{T_i} := W_{T_i} - W_{T_{i-1}}$$

- Consider the "Euler discretization along the arrival times τ_i"

$$\hat{X}_{T_i} = \hat{X}_{T_{i-1}} + \mu(T_{i-1}, \hat{X}_{T_{i-1}})\Delta T_i + \Delta W_{T_i},$$

$$\text{for } i = 1, \ldots, N_T + 1$$

Our algorithm in the case of constant diffusion $\sigma = I_d$ (II)

Define the **exactly simulatable** r.v.

$$\hat{\xi} \;:=\; \beta^{-N_T} e^{\beta T}\left[g(\hat{X}_T) - g(\hat{X}_{T_{N_T}})\mathbb{1}_{\{N_T>0\}}\right]\prod_{k=1}^{N_T}\hat{\mathcal{W}}_k^1$$

where

$$\hat{\mathcal{W}}_k^1 \;:=\; \left(\mu(T_k,\hat{X}_{T_k}) - \mu(T_{k-1},\hat{X}_{T_{k-1}})\right)\cdot\frac{\Delta W_{T_{k+1}}}{\Delta T_{k+1}}.$$

Theorem

Assume g Lipschitz. Then $\hat{\xi} \in \mathbb{L}^2$ and $\mathbb{E}[g(X_T)] = \mathbb{E}[\,\hat{\xi}\,]$

Varying drift versus varying volatility in 1-dim

European option valuation in the local volatility model

$$dX_t = \frac{0.8}{1+X_t^2} dW_t$$

Lamperti transformation leads to SDE with varying drift, unit vol

$$dY_t = \frac{0.8X_t}{(1+X_t^2)^2} dt + dW_t$$

• Comparison with Euler discretization with time steps 1/10, 1/50, 1/100, 1/400 $\Longrightarrow$ agree with our exact discretization for 1/400

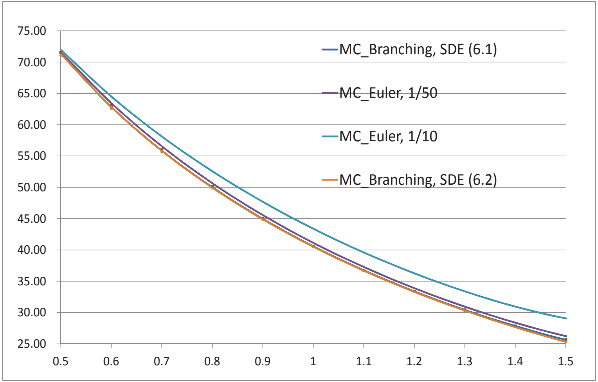

Figure: $V_0(K)$ quoted in implied volatility $\times 100$ as a function of K. The dots correspond to the standard deviation error.

N	$\beta = 0.1$	$\beta = 0.2$	Euler
12	0.32	0.34	0.30
14	0.16	0.17	0.15
16	0.08	0.09	0.08
18	0.05	0.04	0.04
20	0.02	0.02	0.02
22	0.01	0.02	0.01
24	0.01	0.01	0.00

Table: Standard deviation for an at-the-money call option with $K = 1$, $T =$ one year as a function of the Monte-Carlo paths 2^N.

A multi-dimensional example

Basket option $\mathbb{E}[\big(\frac{1}{n}\sum_{i=1}^n X_T^i - K\big)^+]$ in the model :

$$\frac{dX_t^i}{X_t^i} = \frac{1}{2}\, dW_t^i + 0.1\, (\sqrt{X_t^i} - 1)dt \quad d\langle W^i, W^j\rangle_t = 0.5dt, \quad i \neq j$$

Our method is compared to a (log)-Euler discretization scheme with a time step 1/10, 1/50, 1/100 :

$$X_{t+\Delta}^{\Delta} \;=\; X_t^{\Delta} \exp\left(\frac{1}{2}\Delta W_t + \left(0.1\, \left(\sqrt{X_t^i} - 1\right) - \frac{1}{8}\right)\Delta\right).$$

- $\Delta = 1/100$ converges exactly to our exact scheme

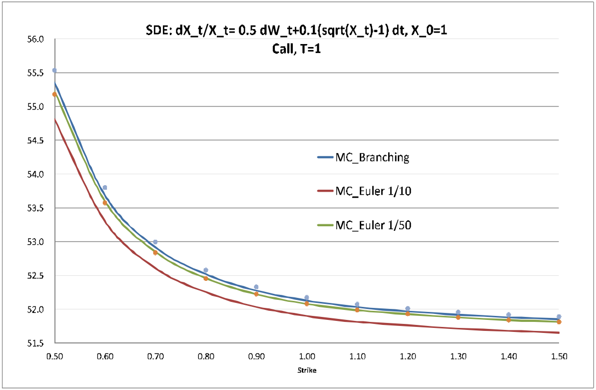

Figure: $d = 1$. $V_0(K)$ quoted in implied volatility $\times 100$ as a function of K. The dots correspond to the standard deviation error.

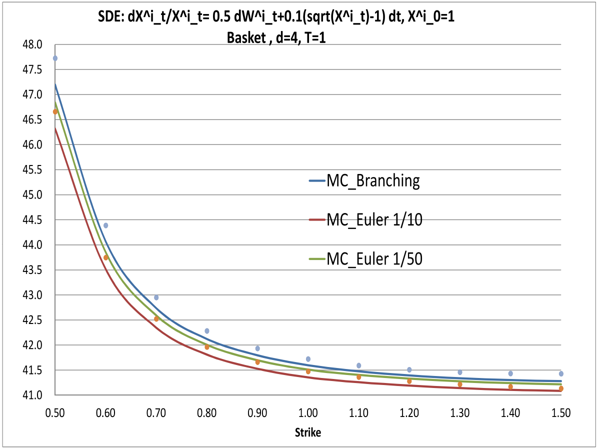

Figure: $d = 4$. $V_0(K)$ quoted in implied volatility $\times 100$ as a function of K. The dots correspond to the standard deviation error.

Objective of this talk

1. Why does it work ?
2. The case of non-constant diffusion coefficient
3. From regime switching diffusions to branching diffusions

$\Longrightarrow$ **forward** Monte Carlo approximation of nonlinear partial differential equations

Outline

A regime switching diffusion

Let $(\hat{\mu}, \hat{\sigma}) : (s, y, t, x) \in [0, T] \times \mathbb{R}^d \times [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{S}^d$ be Lipschitz in x, continuous in t, and define

$$\hat{X}_0 := X_0, \;\; d\hat{X}_t = \hat{\mu}(\Theta_t, t, \hat{X}_t)dt + \hat{\sigma}(\Theta_t, t, \hat{X}_t)dW_t$$

with $\Theta_t := (T_{N_t}, \hat{X}_{T_{N_t}})$. In other words,

$$\hat{X}_{T_{k+1}} = \hat{X}_{T_k} + \int_{T_k}^{T_{k+1}} \hat{\mu}\big(T_k, X_{T_k}, s, \hat{X}_s\big)ds + \int_{T_k}^{T_{k+1}} \hat{\sigma}\big(T_k, X_{T_k}, s, \hat{X}_s\big)dW_s$$

i.e. the coefficients of the diffusion change at each arrival time T_k

First main idea

Define $u(t,x) := \mathbb{E}_{t,x}\big[g(X_T)\big]$, $t \leq T$, $x \in \mathbb{R}$

Proposition

Let $\beta > 0$, $\theta \in [0,T) \times \mathbb{R}^d$, $(t,x) \in [0,T) \times \mathbb{R}^d$. *Then*

$$u(t,x) = e^{\beta(T-t)}\mathbb{E}_{t,x,\theta}\Big[\mathbb{1}_{\{N_T=0\}}\, g\big(\hat{X}_T\big) + \mathbb{1}_{\{N_T>0\}}\frac{1}{\beta}\Delta f \bullet (Du, D^2u)\big(T_1, \hat{X}_{T_1}\big)\Big]$$

where $\Delta f := (\mu, a) - (\hat{\mu}, \hat{a})(\theta, .)$, $(x,A) \bullet (y,B) := x \cdot y + Tr[AB]$

Here $a := \frac{1}{2}\sigma^2$, $\hat{a} := \frac{1}{2}\hat{\sigma}^2$

Sketch of proof of the lemma

The function $\tilde{u} := e^{-\beta(T-t)}\mathbb{E}_{t,x}\big[g(X_T)\big]$ solves

$$-\partial_t \tilde{u} - \mu \cdot D\tilde{u} - a : D^2\tilde{u} + \beta\tilde{u} = 0 \quad \text{and} \quad \tilde{u}(T,.) = g$$

Equivalently, with $\phi := (\mu - \hat{\mu}) \cdot D\tilde{u} + (a - \hat{a}) : D^2\tilde{u}$,

$$-\partial_t \tilde{u} - \hat{\mu} \cdot D\tilde{u} - \hat{a} : D^2\tilde{u} + \beta\tilde{u} = \phi \quad \text{and} \quad \tilde{u}(T,.) = g$$

By the Feynman-Kac representation :

$$\begin{aligned} u(0, X_0) &= e^{\beta T}\mathbb{E}\Big[e^{-\beta T} g(\hat{X}_T) + \int_0^T e^{-\beta t}\phi(t, \hat{X}_t)dt\Big] \\ &= e^{\beta T}\mathbb{E}\Big[g(\hat{X}_T)\mathbb{1}_{\{\tau \geq T\}} + \frac{1}{\beta}\phi(\tau, \hat{X}_\tau)\mathbb{1}_{\{\tau < T\}}\Big] \end{aligned}$$

where τ is an independent Expo(β)

Second main idea : Monte Carlo automatic differentiation

Assumption

For all $\theta \in [0,T) \times \mathbb{R}^d$, and $(t,x) \in [0,T) \times \mathbb{R}^d$, there is a pair of random functions $(\hat{\mathcal{W}}^1_\theta(\cdot), \hat{\mathcal{W}}^2_\theta(\cdot))$, called Malliavin weights, depending only on $(t, x, T^t_1, (W_s - W_t)_{s \le T_1})$ s.t.

$$D^i\, \mathbb{E}_{t,x,\theta}\big[\phi(T_1, \hat{X}_{T_1})\big] = \mathbb{E}_{t,x,\theta}\Big[\phi(T_1, \hat{X}_{T_1})\ \hat{\mathcal{W}}^i\Big]$$

$i = 1, 2,$ *for all bounded* $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$

This assumption corresponds to

- likelihood ratio method for Greeks, Broadie & Glasserman
- El Worthy formula, see Fouriné, Lasry, Lions, Lebuchoux & NT

Back to the constant diffusion case $\sigma = I_d$

Recall that our algorithm in this case uses the Euler scheme sampled at the arrival times of the Poisson process $(N_t)_{t\geq 0}$. Then

$$\begin{aligned}
\partial_x \mathbb{E}_{t,x,\theta}\big[\phi\big(T_1, \hat{X}_{T_1}\big)\big] &= \partial_x \mathbb{E}\big[\phi\big(T_1, x + \mu(\Theta_0)T_1 + \Delta W_{T_1}\big)\big] \\
&= \partial_x \mathbb{E}\int \phi(T_1, y)\frac{e^{\frac{-1}{2T_1}|y-x+\mu(\Theta_0)T_1|^2}}{(2\pi T_1)^{-d/2}}dy \\
&= \mathbb{E}\int \phi(T_1, y)\frac{y - x + \mu(\Theta_0)T_1}{T_1}\frac{e^{\frac{-1}{2T_1}|y-x+\mu(\Theta_0)T_1|^2}}{(2\pi T_1)^{-d/2}}dy \\
&= \mathbb{E}_{t,x,\theta}\left[\phi\Big(T_1, \hat{X}_{T_1}\Big)\frac{\Delta W_{T_1}}{T_1}\right]
\end{aligned}$$

and, similarly,

$$\partial^2_{xx} \mathbb{E}_{t,x,\theta}\big[\phi\big(T_1, \hat{X}_{T_1}\big)\big] = \mathbb{E}_{t,x,\theta}\left[\phi\big(T_1, \hat{X}_{T_1}\big)\frac{(\Delta W_{T_1})^2 - T_1}{(T_1)^2}\right]$$

Combining automatic differentiation with first main idea

Recall from the Proposition that $u(t,x) := \mathbb{E}_{t,x}\big[g(X_T)\big]$ satisfies

$$
\begin{aligned}
u(t,x) &= \mathbb{E}_{t,x,\theta}\Big[e^{\beta(T_1-t)}\Big(\mathbb{1}_{\{N_T=0\}}\, g(\hat{X}_T) \\
&\qquad + \mathbb{1}_{\{N_T>0\}} \frac{\Delta f_{T_1}}{\beta} \bullet (Du, D^2u)(T_1, \hat{X}_{T_1})\Big)\Big] \\
&= \mathbb{E}_{t,x,\theta}\Big[e^{\beta(T_1-t)}\Big(\mathbb{1}_{\{N_T=0\}}\, g(\hat{X}_T) \\
&\qquad + \mathbb{1}_{\{N_T=1\}} \frac{\Delta f_{T_1}}{\beta} \bullet (\hat{\mathcal{W}}^1, \hat{\mathcal{W}}^2) g(\hat{X}_T) \\
&\qquad + \mathbb{1}_{\{N_T>1\}} \frac{\Delta f_{T_2}}{\beta^2} \bullet (Du, D^2u)(T_2, \hat{X}_{T_2})\Big)\Big]
\end{aligned}
$$

by the assumption. And so on...

Back to unit diffusion : square integrability lost... in general

Iterating as above, and passing to limits, we would arrive at

$$\mathbb{E}[\xi] \quad \text{where} \quad \xi := \beta^{-N_T} e^{\beta T} g(\hat{X}_T) \prod_{k=1}^{N_T} \hat{\mathcal{W}}_k^1$$

where, in the case of unit diffusion :

$$\hat{\mathcal{W}}_k^1 \quad := \quad \left[\mu(T_k, \hat{X}_{T_k}) - \hat{\mu}(T_{k-1}, \hat{X}_{T_{k-1}})\right] \cdot \frac{\Delta W_{T_{k+1}}}{\Delta T_{k+1}}$$

However $\frac{\Delta W_{T_1}}{\Delta T_1} \sim (\Delta T_1)^{-1/2}$ and $(\Delta T_1 | N_T = 1)$ is Unif$[0, T]$, so

in general, $\xi \in \mathbb{L}^1$, but $\xi \notin \mathbb{L}^2$!

Recovering square integrability in the unit diffusion case

Choose $\hat{\mu}(s,x,t,y) := \mu(s,x,s,x)$ (Euler !), leads to

$$\xi := \beta^{-N_T} e^{\beta T} g(\hat{X}_T) \prod_{k=1}^{N_T} \hat{\mathcal{W}}_k^1$$

where, by the Lipschitz property of μ :

$$\hat{\mathcal{W}}_k^1 = [\mu(T_k, \hat{X}_{T_k}) - \mu(T_{k-1}, \hat{X}_{T_{k-1}})] \cdot \frac{\Delta W_{T_{k+1}}}{\Delta T_{k+1}} \sim (\Delta T_k)^{1/2} (\Delta T_{k+1})^{-1/2}$$

It remains to deal with the last term $\hat{\mathcal{W}}_{N_T}^1$. For this, we notice that

$$\mathbb{E}[\xi] = \mathbb{E}[\hat{\xi}], \text{ where } \hat{\xi} := \beta^{-N_T} e^{\beta T} [g(\hat{X}_T) - g(\hat{X}_{T_{N_T}}) \mathbb{1}_{\{N_T > 0\}}] \prod_{k=1}^{N_T} \hat{\mathcal{W}}_k^1$$

so that $\hat{\xi} \in \mathbb{L}^2$ by the Lipschitz assumption on g

Driftless one-dimensional diffusion

We now consider the one-dimansional SDE

$$dX_t \;=\; \sigma(t, X_t)dW_t$$

Iterating the Proposition and the Assumption, we arrive to the r.v.

$$\hat{\xi} \;:=\; \beta^{-N_T} e^{\beta T} g\big(\hat{X}_T\big) \prod_{k=1}^{N_T} \hat{\mathcal{W}}_k^2$$

where, denoting $a := \frac{1}{2}\sigma^2$ and $\hat{a} := \frac{1}{2}\hat{\sigma}^2$:

$$\hat{\mathcal{W}}_k^1 \;:=\; \big[a(T_k, \hat{X}_{T_k}) - \hat{a}(T_{k-1}, \hat{X}_{T_{k-1}})\big] \cdot \frac{\big(\Delta W_{T_{k+1}}\big)^2 - \Delta T_{k+1}}{(\Delta T_{k+1})^2}$$

Situation is worse : $\frac{(\Delta W_{T_1})^2 - \Delta T_1}{(\Delta T_1)^2} \sim (\Delta T_1)^{-1}$ and $\big(\Delta T_1 | N_T = 1\big)$ is Unif$[0, T]$!

in general, $\hat{\xi}$ is not integrable ! unless good choice of $\hat{a}$

Choice of the regime switching diffusion

Let

$$\hat{\mu}(\cdot) \equiv 0 \quad \text{and} \quad \hat{\sigma}(s,y,t,x) := \sigma(s,y) + \partial_x \sigma(s,y)(x-y).$$

Then $\hat{X}$ is defined by

$$d\hat{X}_t = \left(c_1^k + c_2^k \hat{X}_t\right) dW_t \quad \text{on each} \quad [T_k, T_{k+1}]$$

where

$$c_1^k := \sigma(T_k, \hat{X}_{T_k}) - \partial_x \sigma(T_k, \hat{X}_{T_k}) \hat{X}_{T_k}, \quad c_2^k := \partial_x \sigma(T_k, \hat{X}_{T_k})$$

$\Longrightarrow$ Explicit solution...
$\Longrightarrow$ Explicit and simulatable Malliavin weight...

Exact simulation of local volatility SDE : first try

$$\xi \;:=\; \beta^{-N_T} e^{\beta T}\left[g(\hat{X}_T) - g(\hat{X}_{T_{N_T}})\mathbb{1}_{\{N_T>0\}}\right] \prod_{k=1}^{N_T} \hat{\mathcal{W}}_k^2$$

where the weight is

$$\hat{\mathcal{W}}_k^2 = \frac{a(\Theta_k) - \hat{a}(\Theta_{k-1}, \Theta_k)}{2a(\Theta_k)}\left(-\partial_x\sigma(\Theta_k)\frac{\Delta W_{T_{k+1}}}{\Delta T_{k+1}} + \frac{\Delta W_{T_{k+1}}^2 - \Delta T_{k+1}}{\Delta T_{k+1}^2}\right)$$

Theorem

Assume in addition $\partial_x\sigma$ Lip in x. Then $\hat{\xi} \in \mathbb{L}^1$ and $V_0 = \mathbb{E}[\,\hat{\xi}\,]$

But square integrability fails, in general !

Exact simulation of local volatility SDE : restoring square integrability

Use the technique of antithetic variables :

- define $\hat{X}_T^-$ exactly as $\hat{X}_T$, except that
- the sign of the last increment of Brownian motion ΔW_T
- introduce the corresponding r.v. $\check{\xi}$

Finally define

$$\bar{\xi} := \frac{1}{2}(\hat{\xi} + \check{\xi})$$

Theorem

Suppose in addition $g \in C_b^2$. Then $\bar{\xi} \in \mathbb{L}^2$ and $V_0 = \mathbb{E}[\,\bar{\xi}\,]$

Choice of the intensity β of the Poisson process

- In the constant diffusion case, we compute directly that

$$\mathbb{E}[\hat{\xi}^2] \leq F(\beta) := Ce^{-\beta T + L' T/\beta}$$

with explicit L'

- The computation effort is proportional to N_T

- A reasonable criterion for the choice of $\beta > 0$ is then :

$$\min_{\beta>0} \frac{F(\beta)}{\mathbb{E}[N_T]} \implies \beta^* := \sqrt{L' + T^2/4} + \frac{T}{2}$$

Limitations

• Multidimensional driftess SDE : reduces to

$$d\hat{X}_t = (A + \langle B, \hat{X}_t \rangle) dW_t \quad \text{where} \quad A \in \mathbb{S}_d,\ B \in \mathcal{L}(\mathbb{R}^d, \mathbb{S}_d)$$

Exact simulation is not available !

• 1-dim SDE with varying drift and volatility : reduces to

$$d\hat{X}_t = (b_0 + b_1 \hat{X}_t) dt + (\sigma_0 + \sigma_1 \hat{X}_t) dW_t$$

Exact simulation is not available !
Moreover, volatility may vanish... Malliavin integration by parts fails
But we can still use Lamperti's transformation

Outline

A first class of Nonlinear Path-Dependent PDEs

Let $p_k \geq 0$ with $\sum_{k=0}^{n} p_k = 1$, and consider the equation :

$$\partial_t v + \frac{1}{2}\partial^2_{\omega\omega} v + \beta\Big(\sum_{k=0}^{n} p_k v^k - v\Big) = 0, \quad v_T = \xi$$

Define the branching Brownian motion :

- Start from one particle driven by a Brownian motion
- T_1 independent exponential distribution with parameter β
- if $T_1 < T$, the first particle dies out and is replaced by k independent particles with probability p_k
- V_T : Number of living particles at time T
- $Z^i_\cdot$: path of particle i

Then (Watanabe, McKean)

$$v(0, X_0) \quad = \quad \mathbb{E}\Big[\prod_{i=1}^{V_T} \xi(Z^i_\cdot)\Big]$$

Path-Dependent KPP equation

Let $(a_k)_{0\le i\le n}$, and consider the equation :

$$\partial_t v + \frac{1}{2}\partial^2_{\omega\omega} v + \beta\Big(\sum_{k=0}^{n} a_k v^k - v\Big) = 0, \quad v_T = \xi$$

Define the branching Brownian motion with probabilities $(p_k)_{0\le k\le n}$. Then

$$v(0, X_0) = \mathbb{E}\Big[\prod_{i=1}^{V_T}\Big(\frac{a_i}{p_i}\Big)^{\ell_i}\xi(Z^i_\cdot)\Big], \quad \ell_i = \#\text{ arrivals for } i$$

- Possible extension to include random drift and random diffusion
- For an analytic nonlinearity $R(v) = \sum_{i=0}^{\infty} a_k v^k$, approximation by substitution $R_n(v) := \sum_{i=0}^{n} a_k v^k$ to $R(v)$

Monte Carlo approximation of nonlinear PDEs

• Purely forward Monte Carlo scheme for KPP equation. Compare with Longstaff-Schwartz backward repeated regression algorithm

• Work in progress : semilinear PDEs

$$\partial_t v + \frac{1}{2}\partial^2_{\omega\omega} v + \beta\Big(\sum_{k=0}^{n} a_k v^{i_k} (\partial_\omega v)^{j_k} - v\Big) = 0, \quad v_T = \xi$$

... key ingredient : automatic differentiation

• Fully nonlinear PDEs... (e.g. HJB equations)

Figure: Happy Birthday Steve