"Vibrato" Monte Carlo evaluation of Greeks

(Smoking Adjoints: part 3)

Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Oxford-Man Institute of Quantitative Finance

MCQMC 2008, July 7-11

"Smoking Adjoints"

Paper with Paul Glasserman in *Risk* in 2006 showed how adjoints can be used in computing pathwise sensitivities – gives lots of first order sensitivities for negligible cost

This attracted a lot of interest, and questions:

- what is involved in practice in creating an adjoint code, and can it be simplified?
  (see HERCMA paper, available from website)
- do we really have to differentiate the payoff?
- what about discontinuous payoffs?
- what about American options?
  (not addressed yet!)

Outline

- different approaches to computing Greeks
  - finite differences
  - likelihood ratio method
  - pathwise sensitivity
- use of conditional expectation for a digital option
- “vibrato” extension for scalar SDE
- generalisation to multidimensional SDEs

Generic Problem

Stochastic differential equation with general drift and volatility terms:

$$\mathrm{d}S_t = a(S_t, t)\,\mathrm{d}t + b(S_t, t)\,\mathrm{d}W_t$$

For a simple European option we want to compute the expected discounted payoff value dependent on the terminal state:

$$V = \mathbb{E}[f(S_T)]$$

Note: the drift and volatility functions are almost always differentiable, but the payoff $f(S)$ is often not.

Generic Problem

Euler discretisation with timestep h:

$$\widehat{S}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\, h + b(\widehat{S}_n, t_n)\, \Delta W_n$$

Simplest Monte Carlo estimator for expected payoff is an average of M independent path simulations:

$$M^{-1} \sum_{i=1}^{M} f(\widehat{S}_N^{(i)})$$

Greeks: for hedging and risk management we also want to estimate derivatives of expected payoff V

Simple Problem

For Geometric Brownian motion

$$\mathrm{d}S_t = r\, S_t\, \mathrm{d}t + \sigma\, S_t\, \mathrm{d}W_t$$

the SDE can be solved analytically to give

$$S_T = S_0 \exp\left((r - \tfrac{1}{2}\sigma^2)\, T + \sigma\, W_T \right)$$

In this case, we can directly sample W_T to get

$$V \equiv \mathbb{E}\left[f(S_T) \right] \approx M^{-1} \sum_{i=1}^{M} f(S_T^{(i)})$$

– will use this to illustrate approaches to calculating sensitivities

Finite Differences

Simplest approach is to use a finite difference approximation,

$$\frac{\partial V}{\partial \theta} \approx \frac{V(\theta+\Delta\theta) - V(\theta-\Delta\theta)}{2\,\Delta\theta}$$

$$\frac{\partial^2 V}{\partial \theta^2} \approx \frac{V(\theta+\Delta\theta) - 2\,V(\theta) + V(\theta-\Delta\theta)}{(\Delta\theta)^2}$$

– very simple, but expensive and inaccurate if $\Delta\theta$ is too big, or too small in the case of discontinuous payoffs

Likelihood Ratio Method

For simple cases where we know the terminal probability distribution

$$V \equiv \mathbb{E}\left[f(S_T)\right] = \int f(S)\, p_S(\theta; S)\, \mathrm{d}S$$

we can differentiate this to get

$$\frac{\partial V}{\partial \theta} = \int f\, \frac{\partial p_S}{\partial \theta}\, \mathrm{d}S = \int f\, \frac{\partial (\log p_S)}{\partial \theta}\, p_S\, \mathrm{d}S = \mathbb{E}\left[f\, \frac{\partial (\log p_S)}{\partial \theta} \right]$$

This is the Likelihood Ratio Method (Broadie & Glasserman, 1996) – its great strength is that it can handle discontinuous payoffs

Likelihood Ratio Method

The LRM weakness is in its generalisation to full path simulations for which we get the multi-dimensional integral

$$\widehat{V} = \mathbb{E}[f(\widehat{S})] = \int f(\widehat{S})\, p(\widehat{S})\, \mathrm{d}\widehat{S},$$

where $$\mathrm{d}\widehat{S} \equiv \mathrm{d}\widehat{S}_1\ \mathrm{d}\widehat{S}_2\ \mathrm{d}\widehat{S}_3\ \ldots\ \mathrm{d}\widehat{S}_N$$

and the joint probability density function $p(\widehat{S})$ is the product of the p.d.f.s for each timestep

$$\begin{aligned} p(\widehat{S}) &= \prod_n p_n(\widehat{S}_{n+1}|\widehat{S}_n) \\ \log p(\widehat{S}) &= \sum_n \log p_n(\widehat{S}_{n+1}|\widehat{S}_n) \end{aligned}$$

Likelihood Ratio Method

When computing Vega from an Euler discretisation of Geometric Brownian motion this leads to

$$\frac{\partial \widehat{V}}{\partial \sigma} = \mathbb{E}\left[\left(\sum_n \frac{Z_n^2 - 1}{\sigma} \right) f(\widehat{S}_N) \right]$$

where Z_n is the unit Normal used in the n^{th} timestep

$$\widehat{S}_{n+1} = \widehat{S}_n (1 + r\, h) + \sigma\, \widehat{S}_n \sqrt{h}\, Z_n$$

Since $\mathbb{V}[Z_n^2 - 1] = 2$ it follows that the variance of the estimator is $O(h^{-1})$

This blow-up as $h \rightarrow 0$ is the weakness of the LRM.

Pathwise sensitivities

Alternatively, for simple Geometric Brownian Motion

$$V \equiv \mathbb{E}\left[f(S_T)\right] = \int f(S_T(\theta; W))\; p_W(W)\, \mathrm{d}W$$

and differentiating this gives

$$\frac{\partial V}{\partial \theta} = \int \frac{\partial f}{\partial S}\, \frac{\partial S_T}{\partial \theta}\; p_W\; \mathrm{d}W = \mathbb{E}\left[\frac{\partial f}{\partial S}\, \frac{\partial S_T}{\partial \theta}\right]$$

with $\partial S_T/\partial\theta$ being evaluated at fixed W.

This is the pathwise sensitivity approach – it can't handle discontinuous payoffs, but generalises well to full path simulations

Pathwise sensitivities

The generalisation involves differentiating the Euler path discretisation,

$$\widehat{S}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\, h + b(\widehat{S}_n, t_n)\, \Delta W_n$$

holding fixed the Brownian increments, to get

$$\frac{\partial \widehat{S}_{n+1}}{\partial \theta} = \left(1 + \frac{\partial a}{\partial S}\, h + \frac{\partial b}{\partial S}\, \Delta W_n \right) \frac{\partial \widehat{S}_n}{\partial \theta} + \frac{\partial a}{\partial \theta}\, h + \frac{\partial b}{\partial \theta}\, \Delta W_n$$

leading to

$$\frac{\partial \widehat{V}}{\partial \theta} = \mathbb{E}\left[\frac{\partial f}{\partial S}(\widehat{S}_N)\, \frac{\partial \widehat{S}_N}{\partial \theta} \right].$$

Pathwise sensitivities

In the case of Vega for an Euler discretisation of GBM

$$\widehat{S}_{n+1} = \widehat{S}_n + r\,\widehat{S}_n\,h + \sigma\,\widehat{S}_n\,\Delta W_n$$

we get

$$\frac{\partial \widehat{S}_{n+1}}{\partial \sigma} = \Big(1 + r\,h + \sigma\,\Delta W_n\Big)\,\frac{\partial \widehat{S}_n}{\partial \sigma} + \widehat{S}_n\,\Delta W_n$$

and the variance

$$\mathbb{V}\left[\frac{\partial f}{\partial S}(\widehat{S}_N)\;\frac{\partial \widehat{S}_N}{\partial \sigma}\right]$$

is $O(1)$ if $f(S)$ is Lipschitz.

Vibrato Monte Carlo

What is best if payoff is discontinuous?

- LRM
  - estimator variance $O(h^{-1})$
- Malliavin calculus
  - estimator variance $O(1)$
  - recent paper by Glasserman & Chen shows it can be viewed as a pathwise/LRM hybrid
  - might be good choice when few Greeks needed
- new “vibrato” Monte Carlo idea
  - also a pathwise/LRM hybrid
  - estimator variance $O(h^{-1/2})$
  - efficient adjoint implementation

Vibrato Monte Carlo

- new idea is based on use of conditional expectation for a simple digital option in Paul Glasserman's book
- output of each SDE path calculation becomes a narrow (multivariate) Normal distribution
- combine pathwise sensitivity for the differentiable SDE, with LRM for the discontinuous payoff
- avoiding the differentiation of the payoff also simplifies the implementation in real-world setting

Vibrato Monte Carlo

Final timestep of Euler path discretisation is

$$\widehat{S}_N = \widehat{S}_{N-1} + a(\widehat{S}_{N-1}, t_{N-1})\, h + b(\widehat{S}_{N-1}, t_{N-1})\, \Delta W_{N-1}$$

Instead of using random number generator to get a value for ΔW_{N-1}, consider the whole distribution of possible values, so $\widehat{S}_N$ has a Normal distribution with mean

$$\mu_W = \widehat{S}_{N-1} + a(\widehat{S}_{N-1}, t_{N-1})\, h$$

and standard deviation

$$\sigma_W = b(\widehat{S}_{N-1}, t_{N-1})\, \sqrt{h}$$

where $W \equiv (\Delta W_0, \Delta W_1, \ldots \Delta W_{N-2})$.

Vibrato Monte Carlo

For a particular path given by a particular vector W, the expected payoff is

$$\mathbb{E}_Z[f(\mu_W + \sigma_W\, Z)]$$

where Z is a unit Normal random variable.

Averaging over all W then gives the same overall expectation as before.

Note also that, for given W, $\widehat{S}_N$ has a Normal distribution

$$p_S(\widehat{S}) = \frac{1}{\sqrt{2\pi}\,\sigma_W}\ \exp\left(-\frac{(\widehat{S}-\mu_W)^2}{2\,\sigma_W{}^2}\right)$$

Vibrato Monte Carlo

In the case of a simple digital call with strike K, the analytic solution is

$$\mathbb{E}_Z[f(\mu_W + \sigma_W Z)] = \exp(-rT)\, \Phi\left(\frac{\mu_W - K}{\sigma_W}\right).$$

- for each W, the payoff is now smooth, differentiable
- derivative is $O(h^{-1/2})$ near strike, near zero elsewhere
  $\Longrightarrow$ variance is $O(h^{-1/2})$
- analytic evaluation of conditional expectation not possible in general for multivariate cases
  $\Longrightarrow$ use Monte Carlo estimation!

Vibrato Monte Carlo

Main novelty comes in calculating the sensitivity.

For a particular W, we have a Normal probability distribution for $\widehat{S}_N$ and can apply the Likelihood Ratio method to get

$$\frac{\partial}{\partial\theta}\,\mathbb{E}_Z\left[f(\widehat{S}_N)\right] \;=\; \mathbb{E}_Z\left[f(\widehat{S}_N)\;\frac{\partial(\log p_S)}{\partial\theta}\right],$$

where

$$\begin{aligned}\frac{\partial(\log p_S)}{\partial\theta} &= \frac{\partial(\log p_S)}{\partial\mu_W}\;\frac{\partial\mu_W}{\partial\theta}+\frac{\partial(\log p_S)}{\partial\sigma_W}\;\frac{\partial\sigma_W}{\partial\theta}\\ &= \frac{Z}{\sigma_W}\;\frac{\partial\mu_W}{\partial\theta}+\frac{Z^2-1}{\sigma_W}\;\frac{\partial\sigma_W}{\partial\theta}.\end{aligned}$$

Averaging over all W then gives the expected sensitivity.

Vibrato Monte Carlo

To improve the variance, we note that

$$\begin{aligned}\mathbb{E}_Z\left[f(\mu_W\!+\!\sigma_W Z)\ Z\ \right] &= \mathbb{E}_Z\left[-f(\mu_W\!-\!\sigma_W Z)\ Z\ \right] \\ &= \tfrac{1}{2}\,\mathbb{E}_Z\left[\Big(f(\mu_W\!+\!\sigma_W Z) - f(\mu_W\!-\!\sigma_W Z)\Big)\ Z\ \right]\end{aligned}$$

and similarly

$$\begin{aligned}&\mathbb{E}_Z\left[f(\mu_W\!+\!\sigma_W Z)\ (Z^2\!-\!1)\ \right] \\ &= \tfrac{1}{2}\,\mathbb{E}_Z\left[\Big(f(\mu_W\!+\!\sigma_W Z) - 2f(\mu_W) + f(\mu_W\!-\!\sigma_W Z)\Big)\ (Z^2\!-\!1)\ \right]\end{aligned}$$

This gives an estimator with $O(1)$ variance when $f(S)$ is Lipschitz, and $O(h^{-1/2})$ variance when it is discontinuous.

Vibrato Monte Carlo

Test case: Geometric Brownian motion

$$\mathrm{d}S_t = r\, S_t\, \mathrm{d}t + \sigma\, S_t\, \mathrm{d}W_t$$

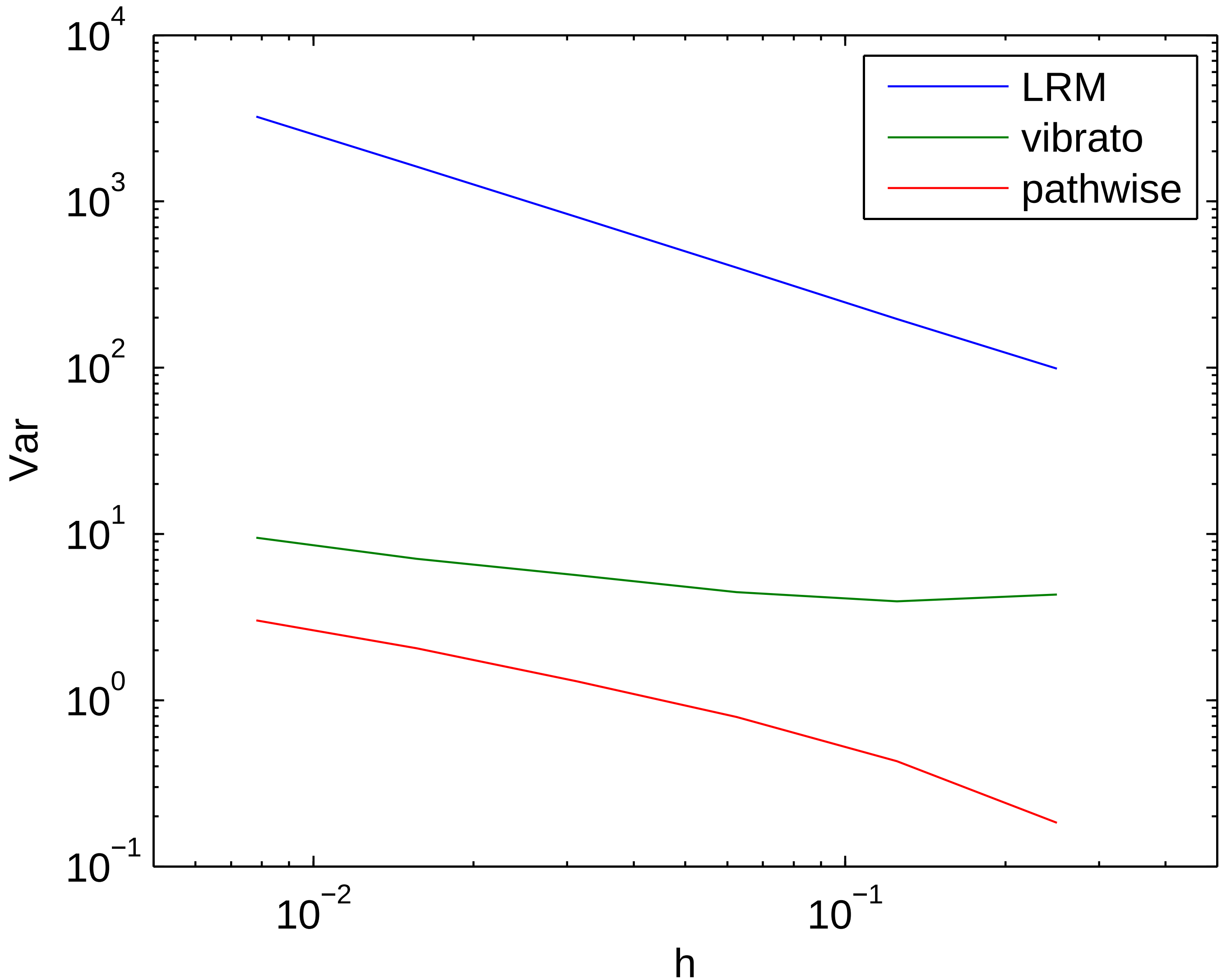
with simple digital call option.

Parameters: $r = 0.05,\ \sigma = 0.2,\ T = 1,\ S_0 = 100,\ K = 100$

Numerical results compare:

- LRM
- vibrato with one Z per W
- pathwise with conditional expectation

Vibrato Monte Carlo



Vibrato Monte Carlo

These results used just one Z per path. If M_Z are used, the variance is

$$\mathbb{V}_W\left[\mathbb{E}_Z[g(W,Z)]\right] + M_Z^{-1}\mathbb{E}_W\left[\mathbb{V}_Z[g(W,Z)]\right]$$

where $g(W,Z)$ is the estimator.

The limit $M_z \to \infty$ gives the variance for the estimator based on the analytic conditional expectation.

The optimal M_Z can be determined if one knows/estimates $\mathbb{V}_W\left[\mathbb{E}_Z[g(W,Z)]\right]$ and $\mathbb{E}_W\left[\mathbb{V}_Z[g(W,Z)]\right]$, and the relative cost of the path simulation and the payoff evaluation.

Multivariate extension

In general we have

$$\widehat{S}(W,Z) = \mu_W + C_W\, Z$$

where $\Sigma_W = C_W\, C_W{}^T$ is the covariance matrix, and Z is a vector of uncorrelated Normals. The joint p.d.f. is

$$\log p_S = -\tfrac{1}{2}\log|\Sigma_W| - \tfrac{1}{2}(\widehat{S}-\mu_W)^T \Sigma_W{}^{-1}(\widehat{S}-\mu_W) - \tfrac{1}{2}d\, \log(2\pi)$$

and so

$$\frac{\partial \log p_S}{\partial \mu_W} = C_W{}^{-T} Z,$$

$$\frac{\partial \log p_S}{\partial \Sigma_W} = \tfrac{1}{2}\, C_W{}^{-T} \left(Z Z^T - I \right) C_W{}^{-1}$$

Multivariate extension

This leads to

$$\frac{\partial}{\partial\theta}\,\mathbb{E}_Z\left[f(\widehat{S})\right] \;=\; \mathbb{E}_Z\left[f(\widehat{S})\;\frac{\partial(\log p_S)}{\partial\theta}\right]$$

where

$$\frac{\partial(\log p_S)}{\partial\theta} = \left(\frac{\partial\log p_S}{\partial\mu_W}\right)^T\frac{\partial\mu_W}{\partial\theta} + \text{tr}\left(\frac{\partial\log p_S}{\partial\Sigma_W}\;\frac{\partial\Sigma_W}{\partial\theta}\right)$$

and $\dfrac{\partial\mu_W}{\partial\theta},\ \dfrac{\partial\Sigma_W}{\partial\theta}$ come from pathwise sensitivity analysis.

A more efficient estimator can be obtained by similar reasoning to the scalar case.

Vibrato Monte Carlo

Test case: Geometric Brownian motion

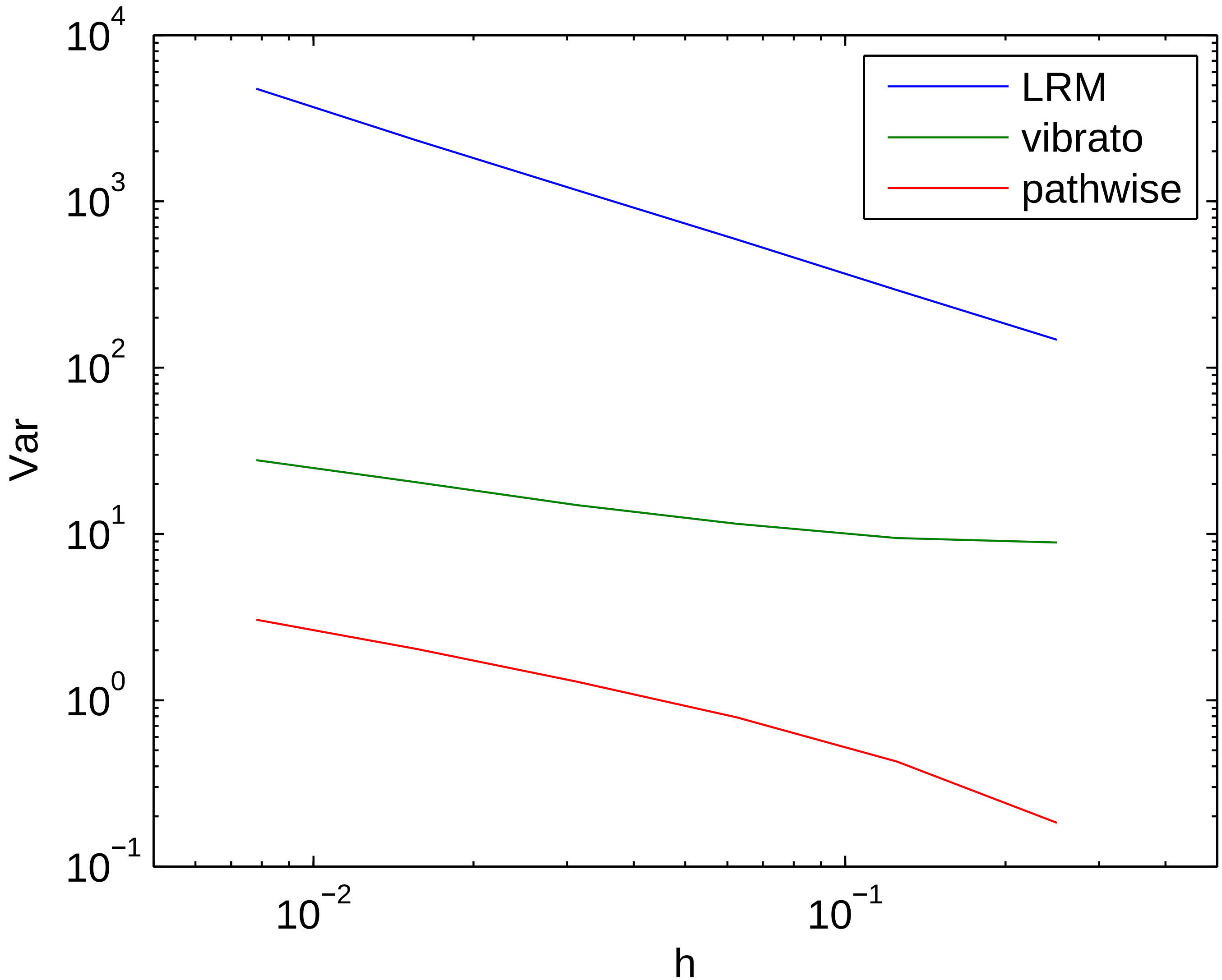
$$\begin{aligned} \mathrm{d}S_t^{(1)} &= r\, S_t^{(1)}\, \mathrm{d}t + \sigma^{(1)}\, S_t^{(1)}\, \mathrm{d}W_t^{(1)} \\ \mathrm{d}S_t^{(1)} &= r\, S_t^{(2)}\, \mathrm{d}t + \sigma^{(2)}\, S_t^{(2)}\, \mathrm{d}W_t^{(2)} \end{aligned}$$

with a simple digital call option based solely on $S_T^{(1)}$.

Parameters: $r = 0.05,\ \sigma^{(1)} = 0.2,\ \sigma^{(2)} = 0.3,\ T = 1,\ S_0^{(1)} = S_0^{(2)} = 100,\ K = 100,\ \rho = 0.5$

Numerical results again compare LRM, vibrato with one Z per W, and pathwise with conditional expectation.

Vibrato Monte Carlo



Multivariate extension

Can also treat payoffs dependent on $S(\tau)$ at intermediate times, by taking

$$t_n < \tau < t_{n+1}$$

and using simple Brownian motion interpolation between $\widehat{S}_n$ and $\widehat{S}_{n+1}$ to get a Normal distribution for $\widehat{S}(\tau)$, with

mean: $$\widehat{S}_n + \frac{\tau - t_n}{t_{n+1} - t_n} \left(\widehat{S}_{n+1} - \widehat{S}_n \right)$$

variance: $$\frac{(\tau - t_n)(t_{n+1} - \tau)}{t_{n+1} - t_n} \, b^2(\widehat{S}_n, t_n)$$

Conclusions

"Vibrato" idea for computing Greeks offers

- $O(1)$ variance for Lipschitz payoffs, and easy implementation – no derivatives required
- $O(h^{-1/2})$ variance for discontinuous payoffs
- adjoint implementation for multiple Greeks

Future work:

- similar idea for digital options in multilevel Monte Carlo path simulation – introduces Radon-Nikodym derivative from change in measure

Acknowledgements

- Paul Glasserman for collaboration on adjoint technique and discussions on vibrato Monte Carlo
- Funding from Microsoft, EPSRC and Man Investments

Further information

- `www.maths.ox.ac.uk/∼gilesm/`
- Email: `mike.giles@maths.ox.ac.uk`