Multilevel Monte Carlo Simulation

Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Oxford-Man Institute of Quantitative Finance

SPA 2009, July 27 – 31, 2009

Multilevel MC Approach

Suppose we want to estimate $\mathbb{E}[P]$ where $P(\omega)$ can be simulated numerically with different levels of accuracy, and corresponding costs, giving $\widehat{P}_l, \ \ l=0,1,\ldots,L$.

$$\mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{l=1}^{L} \mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}]$$

Expected value is same – aim is to reduce variance of estimator for a fixed computational cost.

Key idea: approximate $\mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}]$ using N_l simulations with $\widehat{P}_l$ and $\widehat{P}_{l-1}$ obtained using <u>same</u> underlying sample ω).

$$\widehat{Y}_l = N_l^{-1} \sum_{i=1}^{N_l} \left(\widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)} \right)$$

Multilevel MC Approach

Using independent samples for each level, the variance of the combined estimator is

$$\mathbb{V}\left[\sum_{l=0}^{L} \widehat{Y}_l\right] = \sum_{l=0}^{L} N_l^{-1} V_l, \qquad V_l \equiv \begin{cases} \mathbb{V}[\widehat{P}_l - \widehat{P}_{l-1}], & l > 0 \\ \mathbb{V}[\widehat{P}_0], & l = 0 \end{cases}$$

and the computational cost is $\displaystyle\sum_{l=0}^{L} N_l\, C_l,$

where C_l is the cost of a single sample.

Hence, the variance is minimised for a fixed computational cost by choosing N_l to be proportional to $\sqrt{V_l/C_l}$.

Multilevel MC Approach

Since

$$\mathbb{E}\left[(\widehat{Y}-\mathbb{E}[P])^2\right] = \mathbb{V}[\widehat{Y}] + \left(\mathbb{E}[\widehat{P}_L] - \mathbb{E}[P]]\right)^2$$

can choose

- constant of proportionality for N_l so that $\mathbb{V}[\widehat{Y}] \approx \frac{1}{2}\varepsilon^2$
- finest level L so that $\left(\mathbb{E}[\widehat{P}_L - P]\right)^2 \approx \frac{1}{2}\varepsilon^2$

to get Mean Square Error equal to ε^2

Previous work

- First paper (*Operations Research, 2006 – 2008*) applied idea to SDE path simulation using Euler-Maruyama discretisation
- Second paper (*MCQMC 2006 – 2007*) used Milstein discretisation for scalar SDEs – improved strong convergence gives improved multilevel variance convergence

- Multilevel method is a generalisation of two-level control variate method of Kebaier (2005), and similar to ideas of Speight (2009)
- Also related to multilevel parametric integration by Heinrich (2001)

Multilevel Theorem

Theorem: Given multilevel estimators $\widehat{Y}_l$ based on N_l samples, each with cost C_l, and positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ with $\alpha \geq \frac{1}{2}\gamma$, such that

i) $\left| \mathbb{E}[\widehat{P}_l - P] \right| \leq c_1\, 2^{-\alpha\, l}$

ii) $\mathbb{E}[\widehat{Y}_l] = \begin{cases} \mathbb{E}[\widehat{P}_0], & l = 0 \\ \mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}], & l > 0 \end{cases}$

iii) $\mathbb{V}[\widehat{Y}_l] \leq c_2\, N_l^{-1} 2^{-\beta\, l}$

iv) $C_l \leq c_3\, 2^{\gamma\, l}$

Multilevel Theorem

then there is constant c_4 such that for any $\varepsilon < e^{-1}$ there are values L and N_l for which the multilevel estimator

$$\widehat{Y} = \sum_{l=0}^{L} \widehat{Y}_l,$$

with Mean Square Error $MSE \equiv \mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[P]\right)^2\right] < \varepsilon^2$

with a computational cost C with bound

$$C \leq \begin{cases} c_4\, \varepsilon^{-2}, & \beta > \gamma, \\ c_4\, \varepsilon^{-2} (\log \varepsilon)^2, & \beta = \gamma, \\ c_4\, \varepsilon^{-2-(\gamma-\beta)/\alpha}, & 0 < \beta < \gamma. \end{cases}$$

Multilevel path simulation

In multilevel path simulations for scalar SDEs such as

$$\mathrm{d}S = a(S,t)\,\mathrm{d}t + b(S,t)\,\mathrm{d}W, \quad 0 \le t \le T,$$

each level typically uses twice as many timesteps as the previous, so $\gamma = 1$.

Question then is: what is β?

$$V_l \;\propto\; 2^{-\beta\, l} \;\propto\; h_l^\beta$$

where h_l is timestep on level l.

Multilevel path simulation

For applications in which P is a Lipschitz function of $S(T)$, value of underlying path simulation at a fixed time, strong convergence property

$$\left(\mathbb{E}\left[(\widehat{S}_N - S(T))^2\right]\right)^{1/2} = O(h^\omega)$$

implies that

$$\mathbb{V}[\widehat{P}_l - P] = O(h_l^{2\omega})$$

and hence

$$\mathbb{V}[\widehat{P}_l - \widehat{P}_{l-1}] = O(h_l^{2\omega})$$

and therefore $\beta = 2\,\omega$.

Multilevel path simulation

	Euler		Milstein	
option	numerics	analysis	numerics	analysis
Lipschitz	$O(h)$	$O(h)$	$O(h^2)$	$O(h^2)$
Asian	$O(h)$	$O(h)$	$O(h^2)$	$O(h^2)$
lookback	$O(h)$	$O(h)$	$O(h^2)$	$o(h^{2-\delta})$
barrier	$O(h^{1/2})$	$o(h^{1/2-\delta})$	$O(h^{3/2})$	$o(h^{3/2-\delta})$
digital	$O(h^{1/2})$	$O(h^{1/2}\log h)$	$O(h^{3/2})$	$o(h^{3/2-\delta})$

Table: convergence for V_l as observed numerically and proved analytically for both the Euler and Milstein discretisations. δ can be any strictly positive constant.

Multilevel path simulation

Analysis for Euler discretisations:

- lookback and barrier: Giles, Higham & Mao (*Finance & Stochastics, 2009*)
- digital: Avikainen (*Finance & Stochastics, 2009*)

Analysis for Milstein discretisations:

- Giles, Debrabant & Rößler (TU Darmstadt)
- multilevel estimator for path-dependent options based on conditional Brownian interpolation within timesteps (or extrapolation in final timestep)

Milstein Scheme

Brownian interpolation: within each timestep, model the behaviour as simple Brownian motion (i.e. constant drift and volatility) conditional on the two end-points

$$\begin{aligned}\widehat{S}(t) &= \widehat{S}_n + \lambda(t)(\widehat{S}_{n+1} - \widehat{S}_n) \\ &\quad + b_n \Big(W(t) - W_n - \lambda(t)(W_{n+1} - W_n) \Big),\end{aligned}$$

where $\lambda(t) = \dfrac{t - t_n}{t_{n+1} - t_n}$.

There then exist analytic results for the distribution of the min/max/average over each timestep, and probability of crossing a barrier.

Milstein Scheme

Theorem: Under standard conditions,

- $$\mathbb{E}\left[\sup_{[0,T]} \left|\widehat{S}(t) - S(t)\right|^m\right] = O((h \log h)^m),$$

- $$\sup_{[0,T]} \mathbb{E}\left[\left|\widehat{S}(t) - S(t)\right|^m\right] = O(h^m),$$

- $$\mathbb{E}\left[\left(\int_0^T \widehat{S}(t) - S(t)\ \mathrm{d}t\right)^2\right] = O(h^3).$$

Milstein Scheme

The variance convergence for the Asian option comes directly from this.

Will now outline the analysis for the lookback option – the barrier is similar but more complicated.

The digital option is based on a Brownian extrapolation from one timestep before the end – the analysis is similar.

The analysis for the lookback, barrier and digital options uses the idea of “extreme” paths which are highly improbable – the variance comes mainly from non-extreme paths for which one can use asymptotic analysis.

Milstein Scheme

Computing $\widehat{P}_l - \widehat{P}_{l-1}$ requires a fine and coarse path simulation for the same underlying Brownian motion.

On the fine path, the minimum over one timestep is

$$\widehat{S}^f_{n,min} = \frac{1}{2}\left(\widehat{S}^f_n + \widehat{S}^f_{n+1} - \sqrt{\left(\widehat{S}^f_{n+1} - \widehat{S}^f_n\right)^2 - 2\,(b^f_n)^2\, h_l \log U_n}\ \right)$$

where U_m is a $(0,1]$ uniform random variable.

For the coarse path, first define $\widehat{S}^c_n$ for odd n using conditional Brownian interpolation, then use the same expression for the minimum with same U_n

Milstein Scheme

Theorem: For any $\gamma > 0$, the probability that $W(t)$, its increments ΔW_n and the corresponding SDE solution $S(t)$ and approximations $\widehat{S}^f_n$ and $\widehat{S}^c_n$ satisfy any of the following "extreme" conditions

$$\begin{aligned}
\max_n \left(\max(|S(nh)|,\ |\widehat{S}^f_n|,\ |\widehat{S}^c_n|) \right) &> h^{-\gamma} \\
\max_n \left(\max(|S(nh) - \widehat{S}^c_n|,\ |S(nh) - \widehat{S}^f_n|,\ |\widehat{S}^f_n - \widehat{S}^c_n|) \right) &> h^{1-\gamma} \\
\max_n |\Delta W_n| &> h^{1/2-\gamma}
\end{aligned}$$

is $o(h^p)$ for all $p > 0$.

Milstein Scheme

Furthermore, there exist constants c_1, c_2, c_3, c_4 such that if none of these conditions is satisfied, and $\gamma < \frac{1}{2}$, then

$$
\begin{aligned}
\max_n |\widehat{S}^f_n - \widehat{S}^f_{n-1}| &\leq c_1 \, h^{1/2-2\gamma} \\
\max_n |b^f_n - b^f_{n-1}| &\leq c_2 \, h^{1/2-2\gamma} \\
\max_n \left(|b^f_n| + |b^c_n| \right) &\leq c_3 \, h^{-\gamma} \\
\max_n |b^f_n - b^c_n| &\leq c_4 \, h^{1/2-2\gamma}
\end{aligned}
$$

where b^c_n is defined to equal b^c_{n-1} if n is odd.

Milstein Scheme

The lookback analysis splits paths into:

- "extreme" paths, which have such low probability that their contribution to the variance is negligible ($o(h^p)$ for any $p>0$)
- non-extreme paths for which it can be proved that

$$
\begin{aligned}
\left| \widehat{S}^f_{min} - \widehat{S}^c_{min} \right| &\leq \max_n \left| \widehat{S}^f_{n,min} - \widehat{S}^c_{n,min} \right| \\
&= o(h_l^{1-\delta/2})
\end{aligned}
$$

for any $\delta > 0$, and hence $V_l = o(h_l^{2-\delta})$.

SPDE application

Currently working with Christoph Reisinger on an SPDE application which arises in CDO modelling (Bush, Hambly, Haworth & Reisinger)

$$\mathrm{d}p = -\mu \, \frac{\partial p}{\partial x} \, \mathrm{d}t + \frac{1}{2} \, \frac{\partial^2 p}{\partial x^2} \, \mathrm{d}t + \sqrt{\rho} \, \frac{\partial p}{\partial x} \, \mathrm{d}W$$

with absorbing boundary $p(0,t) = 0$

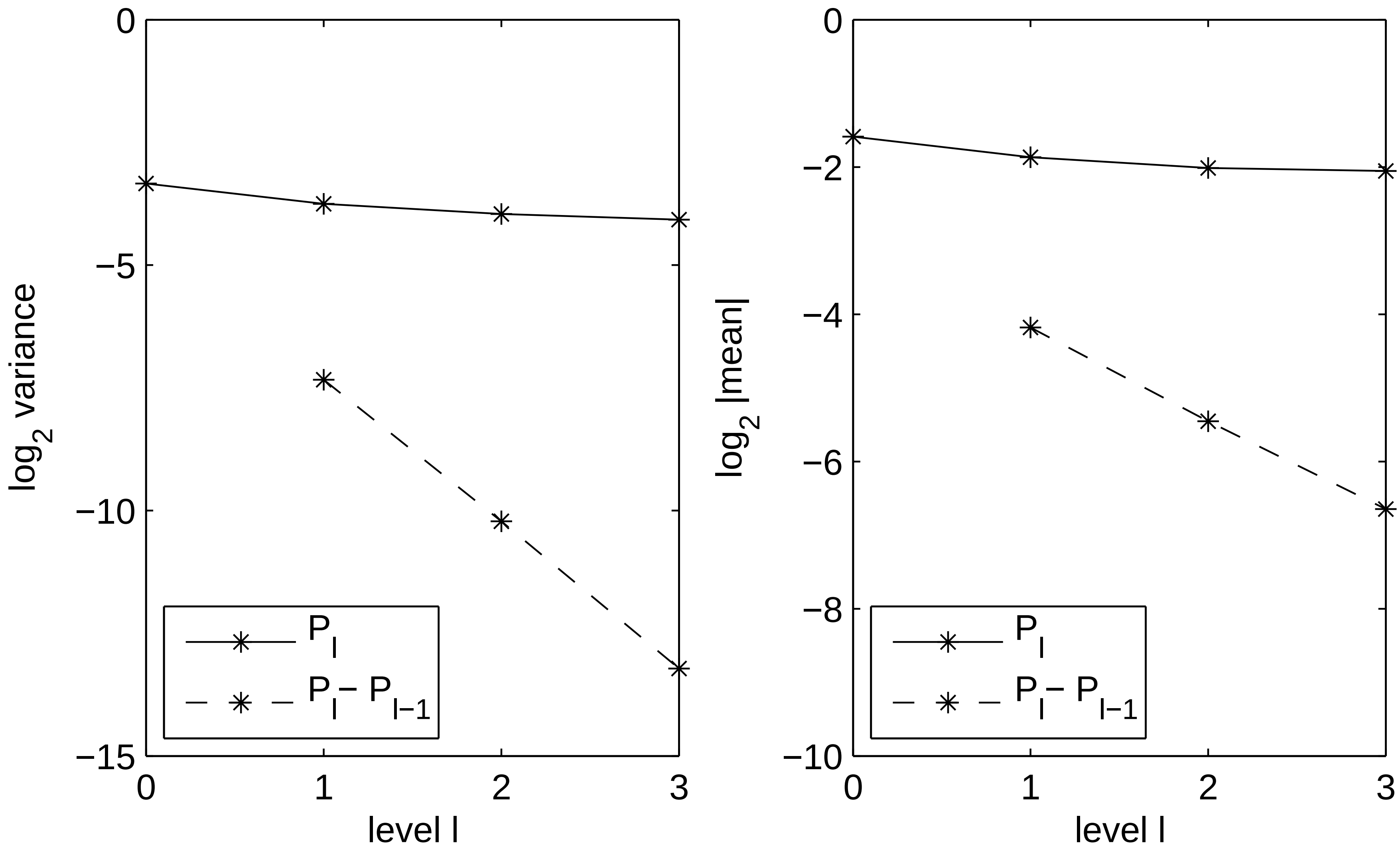
- derived in limit as number of firms $\longrightarrow \infty$
- x is distance to default
- $p(x,t)$ is probability density function
- $\mathrm{d}W$ term corresponds to systemic risk
- $\partial^2 p / \partial x^2$ comes from idiosyncratic risk

SPDE application

- numerical discretisation combines Milstein time-marching with central difference approximations
- coarsest level of approximation uses 1 timestep per quarter, and 10 spatial points
- each finer level uses four times as many timesteps, and twice as many spatial points – ratio is due to numerical stability constraints
- mean-square stability theory, with and without absorbing boundary
- computational cost $C_l \propto 8^l$
- numerical results suggest variance $V_l \propto 8^{-l}$
- can prove $V_l \propto 16^{-l}$ when no absorbing boundary

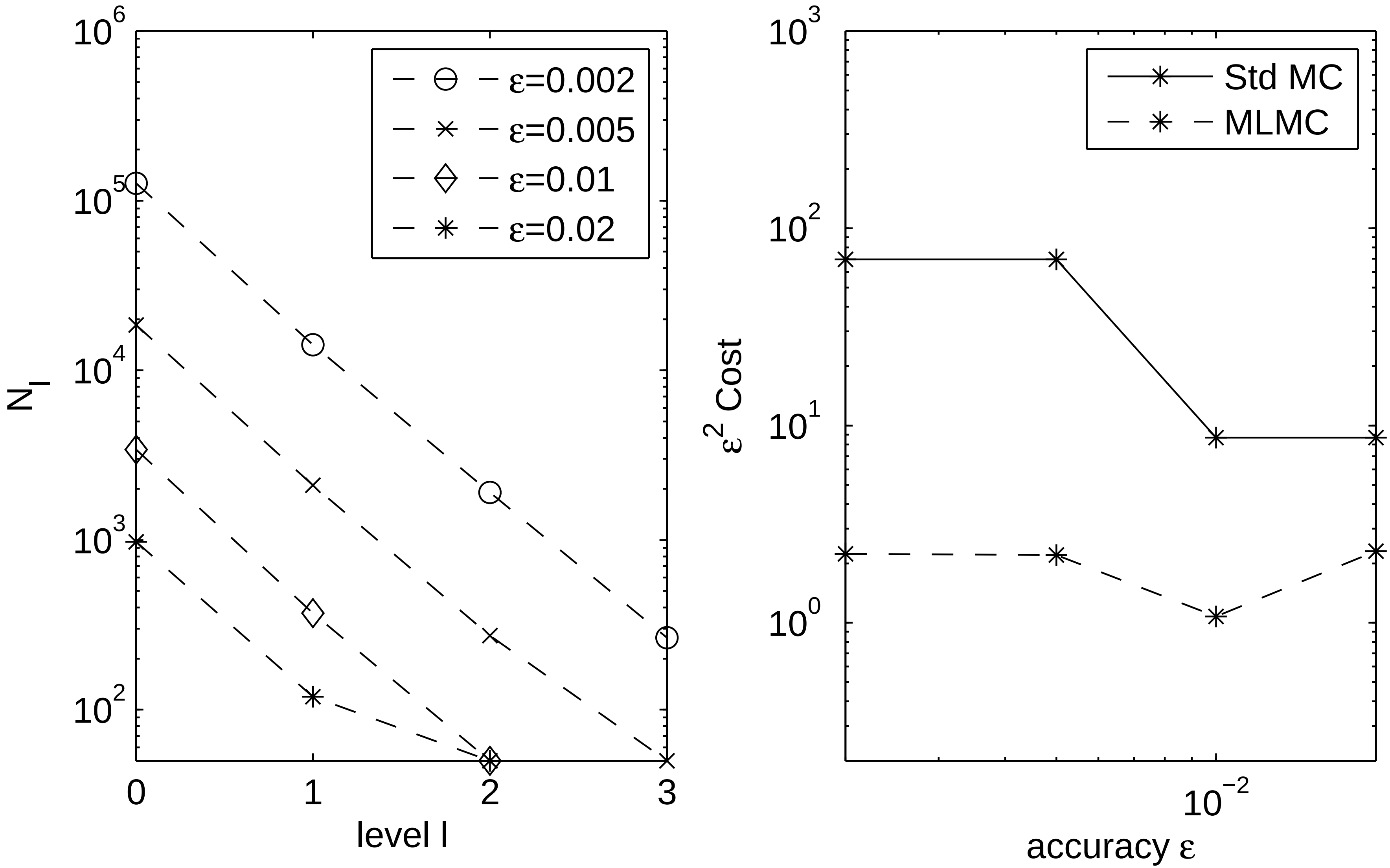
SPDE application

Fractional loss on equity tranche of a 5-year CDO:



SPDE application

Fractional loss on equity tranche of a 5-year CDO:



Future work

- “vibrato” technique for digital options:
  - current treatment uses conditional expectation one timestep before maturity, which smooths the payoff
  - the “vibrato” idea generalises this to cases without a known conditional expectation
- Greeks:
  - the multilevel approach should work well, combining pathwise sensitivities with “vibrato” treatment to cope with lack of smoothness
  - can also incorporate the adjoint approach developed with Paul Glasserman – more efficient when many Greeks are wanted for one payoff function

Future work

- variance-gamma, CGMY and other Lévy processes:
  - given exact simulation techniques, multilevel benefit is in treating path-dependent options
  - could also handle addition of a local vol surface
- American options – the next big challenge:
  - instead of Longstaff-Schwartz approach, view it as an exercise boundary optimisation problem, and use multilevel optimisation?

Conclusions

Multilevel Monte Carlo method has already achieved

- improved order of complexity
- significant benefits for model problems

but there is still a lot more research to be done, both theoretical and applied.

M.B. Giles, "Multilevel Monte Carlo path simulation", *Operations Research*, 56(3):607-617, 2008.

M.B. Giles. "Improved multilevel Monte Carlo convergence using the Milstein scheme", pp. 343-358 in *Monte Carlo and Quasi-Monte Carlo Methods 2006*, Springer, 2007.

Papers are available from:
`www.maths.ox.ac.uk/~gilesm/finance.html`