Multilevel Monte Carlo Path Simulation

Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Oxford-Man Institute of Quantitative Finance

DqF Workshop on SDEs: Models and Numerics

KTH, October 20-22, 2008

Generic Problem

Stochastic differential equation with general drift and volatility terms:

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

In many finance applications, we want to compute the expected value of an option dependent on the terminal state

$$P = f(S(T))$$

with a uniform Lipschitz bound,

$$|f(U) - f(V)| \ \leq c\,\|U - V\|\,, \quad \forall\, U, V.$$

Standard MC Approach

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 estimator for expected payoff is an average of N independent path simulations:

$$\widehat{Y} = N^{-1} \sum_{i=1}^{N} f(\widehat{S}^{(i)}_{T/h})$$

- weak convergence – $O(h)$ error in expected payoff
- strong convergence – $O(h^{1/2})$ error in individual path

Standard MC Approach

Mean Square Error is $O\left(N^{-1} + h^2\right)$

- first term comes from variance of estimator
- second term comes from bias due to weak convergence

To make this $O(\varepsilon^2)$ requires

$$N = O(\varepsilon^{-2}), \quad h = O(\varepsilon) \quad \Longrightarrow \quad \text{cost} = O(N\,h^{-1}) = O(\varepsilon^{-3})$$

Aim is to improve this cost to $O\left(\varepsilon^{-2}(\log \varepsilon)^2\right)$, by combining simulations with different numbers of timesteps – same accuracy as finest calculations, but at a much lower computational cost.

Multilevel MC Approach

Consider multiple sets of simulations with different timesteps $h_l = 2^{-l}\,T,\ l = 0, 1, \ldots, L$, and payoff $\widehat{P}_l$

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

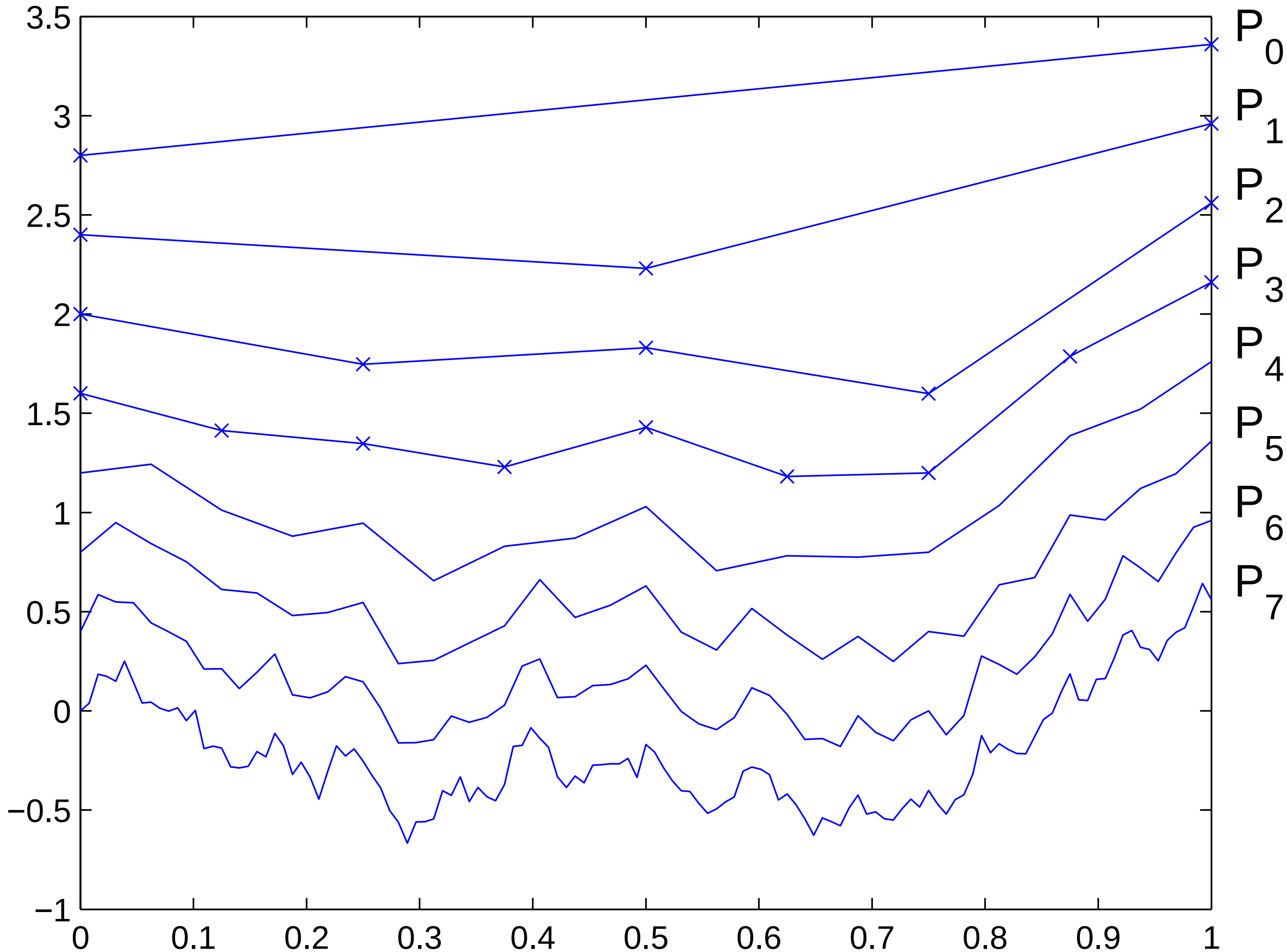
Expected value is same – aim is to reduce variance of estimator for a fixed computational cost.

Key point: 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> Brownian path.

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

Multilevel MC Approach

Discrete Brownian path at different levels



Multilevel MC Approach

- each level adds more detail to Brownian path
- $\mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}]$ reflects impact of that extra detail on the payoff
- different timescales handled by different levels
  – similar to different wavelengths being handled by different grids in multigrid

Multilevel MC Approach

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

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

and the computational cost is proportional to $\displaystyle\sum_{l=0}^{L} N_l\, h_l^{-1}$.

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

The constant of proportionality can be chosen so that the combined variance is $O(\varepsilon^2)$.

Multilevel MC Approach

For the Euler discretisation and the Lipschitz payoff function

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

and the optimal N_l is asymptotically proportional to h_l.

To make the combined variance $O(\varepsilon^2)$ requires

$$N_l = O(\varepsilon^{-2} L\, h_l).$$

To make the bias $O(\varepsilon)$ requires

$$L = \log_2 \varepsilon^{-1} + O(1) \quad \Longrightarrow \quad h_L = O(\varepsilon).$$

Hence, we obtain an $O(\varepsilon^2)$ MSE for a computational cost which is $O(\varepsilon^{-2} L^2) = O(\varepsilon^{-2} (\log \varepsilon)^2)$.

Results

Geometric Brownian motion:

$$\mathrm{d}S = r\, S\, \mathrm{d}t + \sigma\, S\, \mathrm{d}W, \qquad 0 < t < T,$$

$T\!=\!1, \;\; S(0)\!=\!1, \;\; r\!=\!0.05, \;\; \sigma\!=\!0.2$

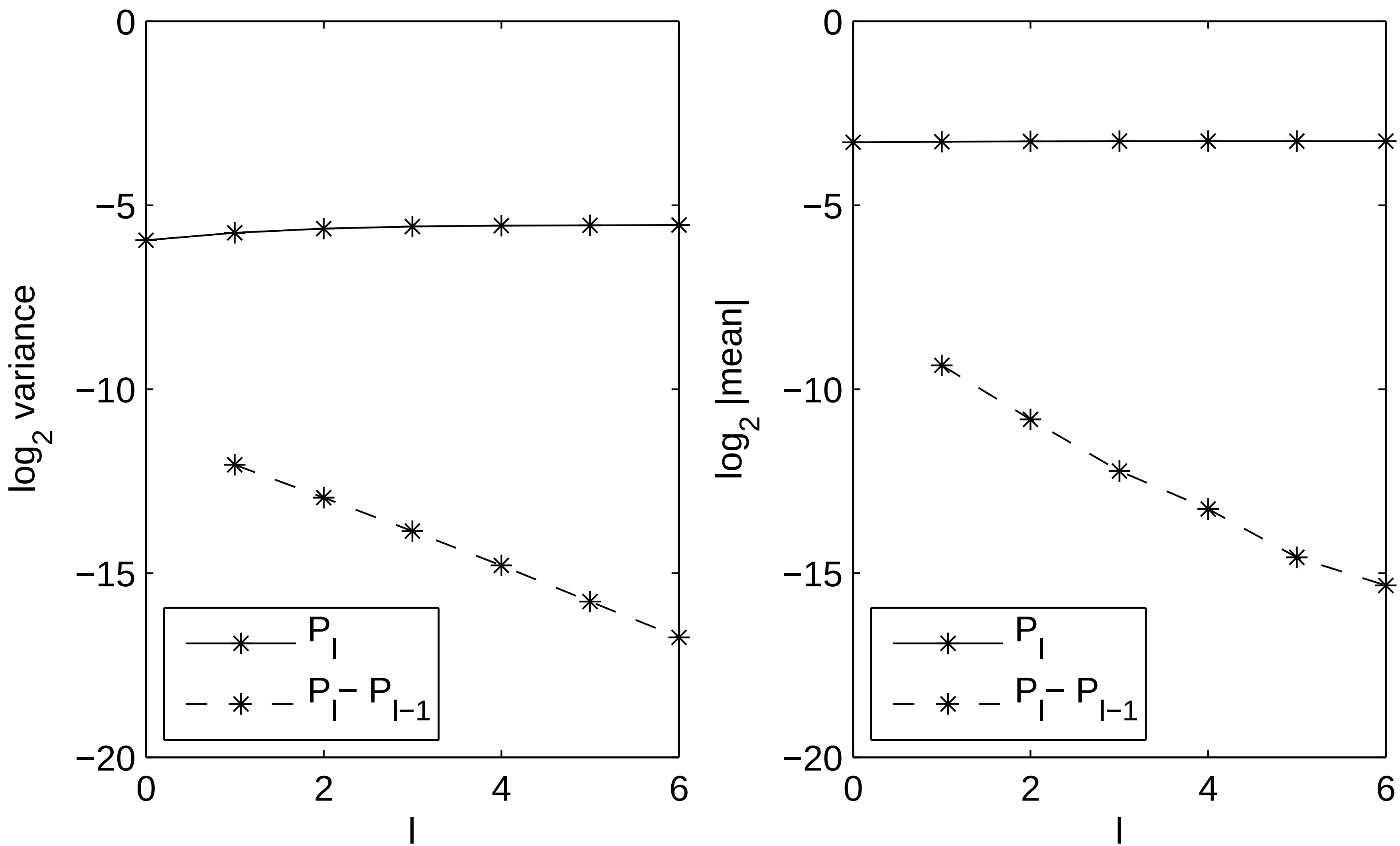
European call option with discounted payoff

$$\exp(-rT)\;\max(S(T)\!-\!K, 0)$$

with $K\!=\!1$.

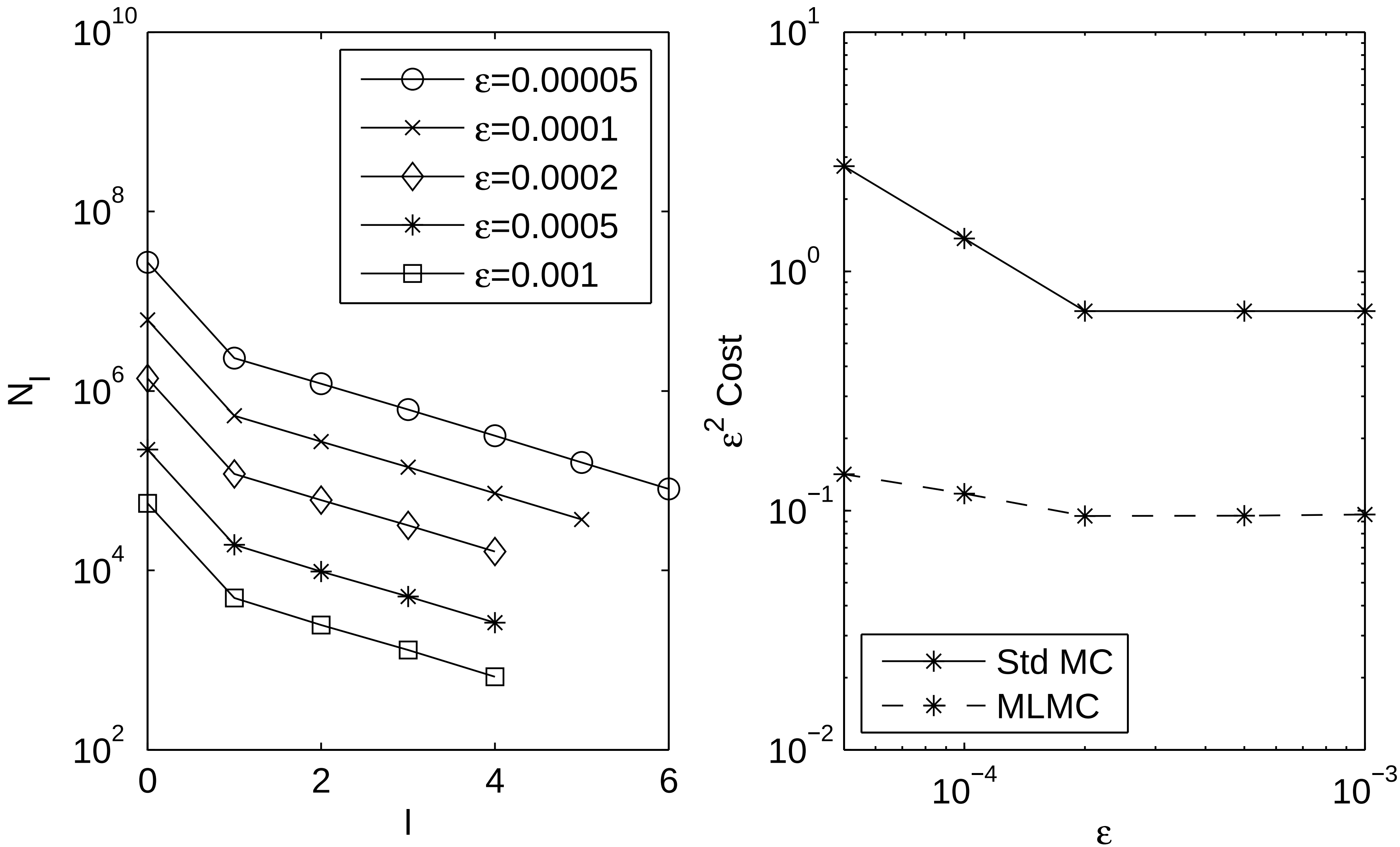
MLMC Results

GBM: European call, $\exp(-rT)\ \max(S(T)\!-\!K, 0)$



MLMC Results

GBM: European call, $\exp(-rT)\,\max(S(T)-K,0)$



Multilevel MC Approach

Theorem: *Let P be a functional of the solution of a stochastic o.d.e., and $\widehat{P}_l$ the discrete approximation using a timestep $h_l = M^{-l}\,T$.*

If *there exist independent estimators $\widehat{Y}_l$ based on N_l Monte Carlo samples, and positive constants $\alpha \geq \frac{1}{2}, \beta, c_1, c_2, c_3$ such that*

i) $\mathbb{E}[\widehat{P}_l - P] \leq c_1\, h_l^\alpha$

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

iv) C_l*, the computational complexity of* $\widehat{Y}_l$*, is bounded by*

$$C_l \leq c_3\, N_l\, h_l^{-1}$$

Multilevel MC Approach

then *there exists a positive constant* c_4 *such that for any* $\varepsilon < e^{-1}$ *there are values* L *and* N_l *for which the multi-level estimator*

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

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

with a computational complexity C *with bound*

$$C \leq \begin{cases} c_4\,\varepsilon^{-2}, & \beta > 1, \\ c_4\,\varepsilon^{-2}(\log\varepsilon)^2, & \beta = 1, \\ c_4\,\varepsilon^{-2-(1-\beta)/\alpha}, & 0 < \beta < 1. \end{cases}$$

Milstein Scheme

The theorem suggests use of Milstein approximation
– better strong convergence, same weak convergence

Generic scalar SDE:

$$\mathrm{d}S(t) = a(S,t)\;\mathrm{d}t + b(S,t)\;\mathrm{d}W(t), \qquad 0<t<T.$$

Milstein scheme:

$$\widehat{S}_{n+1} = \widehat{S}_n + a\,h + b\,\Delta W_n + \tfrac{1}{2}\,b'\,b\left((\Delta W_n)^2 - h\right).$$

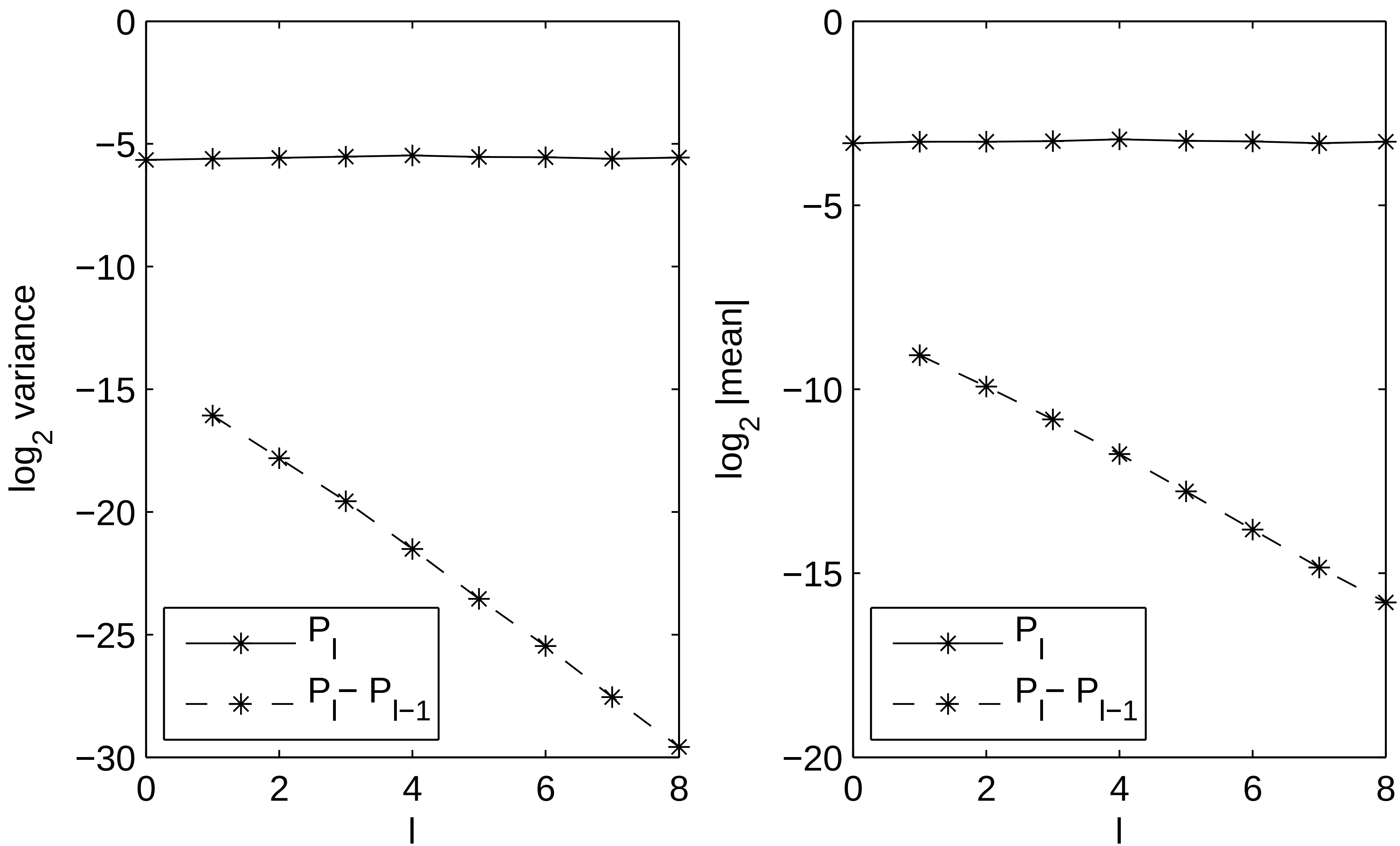
Milstein Scheme

In scalar case:

- $O(h)$ strong convergence
- $O(\varepsilon^{-2})$ complexity for Lipschitz payoffs – trivial
- $O(\varepsilon^{-2})$ complexity for more complex cases using carefully constructed estimators based on Brownian interpolation or extrapolation
  - digital, with discontinuous payoff
  - Asian, based on average
  - lookback and barrier, based on min/max

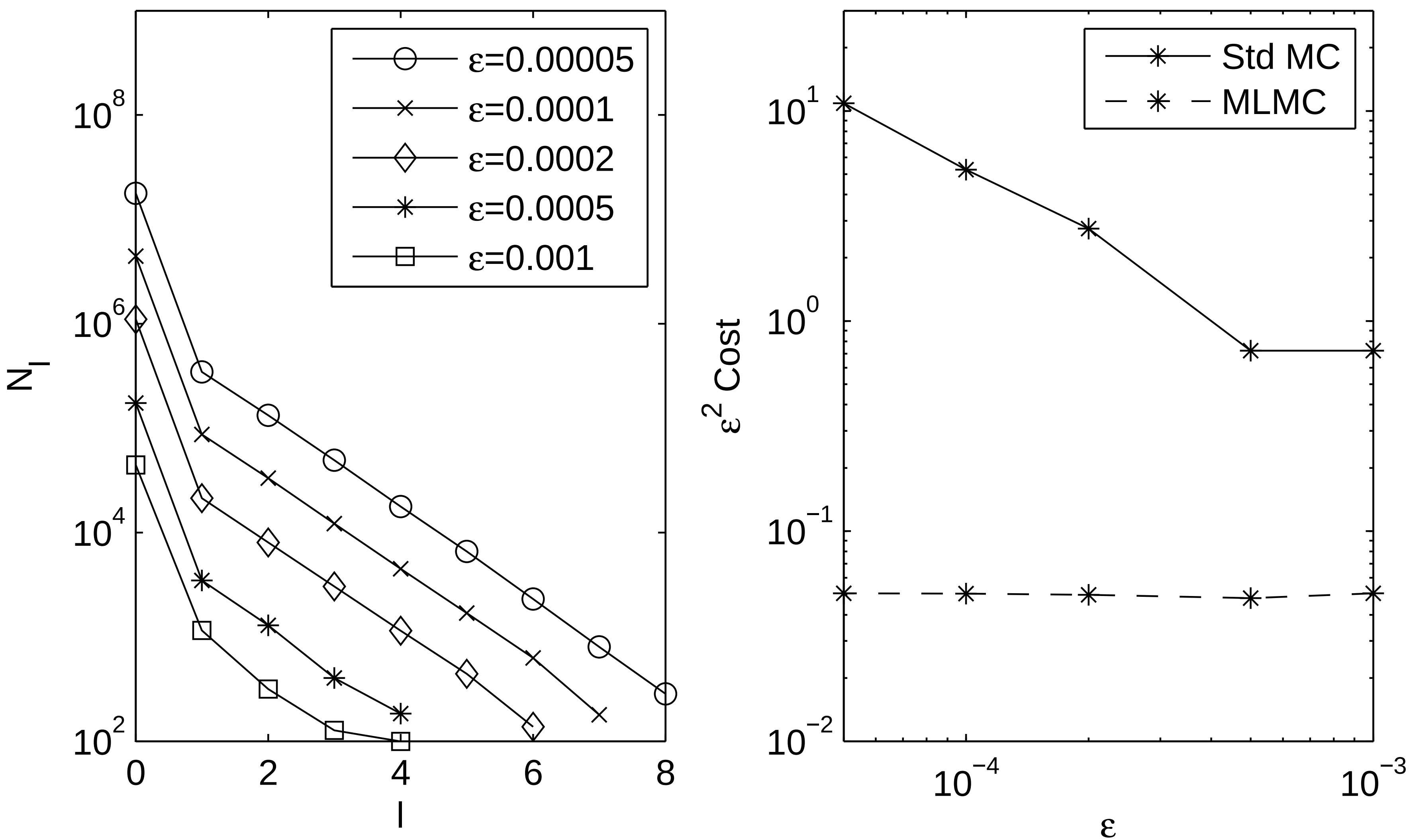
MLMC Results

GBM: European call, $\exp(-rT)\ \max(S(T)-K,0)$



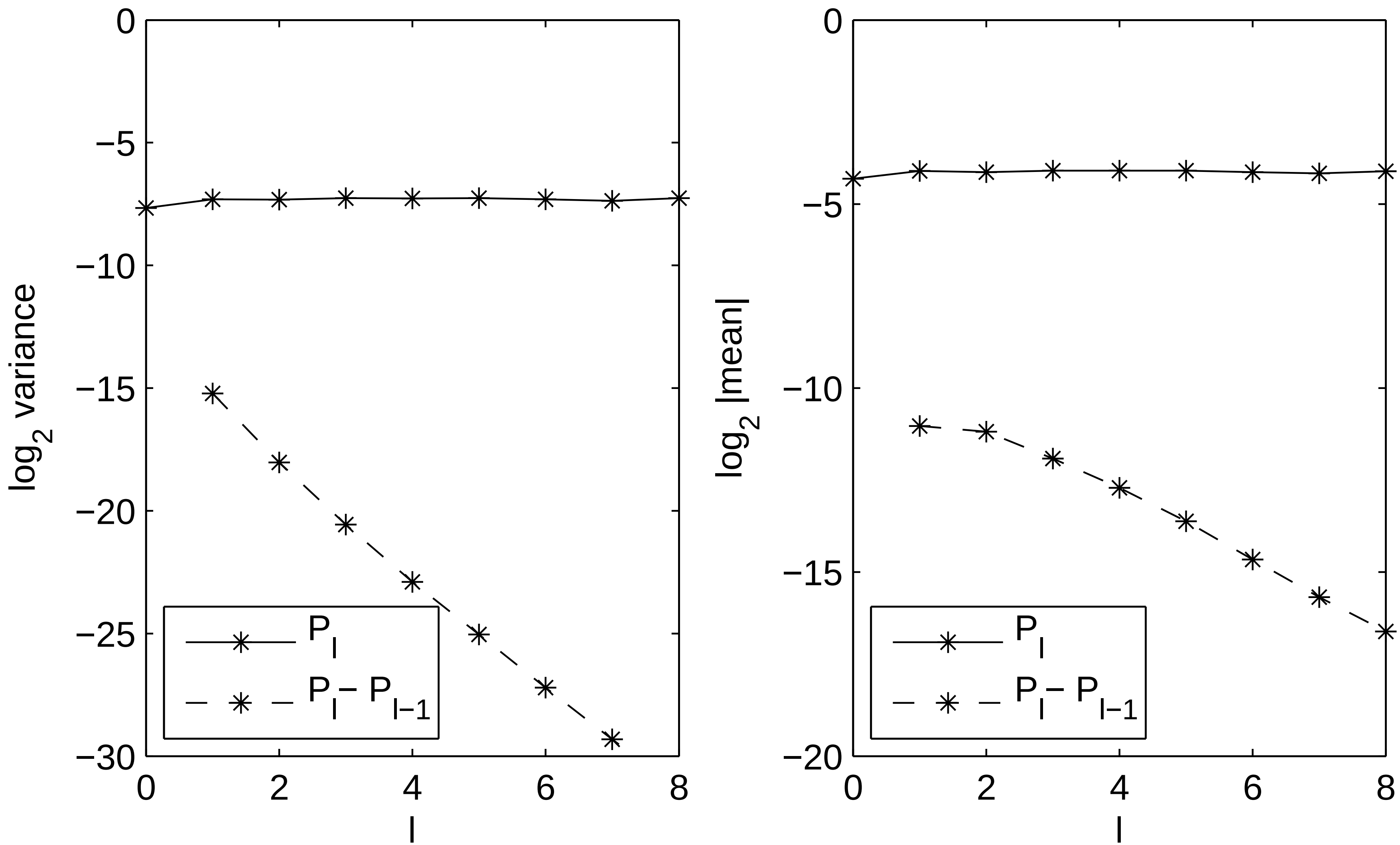
MLMC Results

GBM: European call, $\exp(-rT)\,\max(S(T)-K,0)$



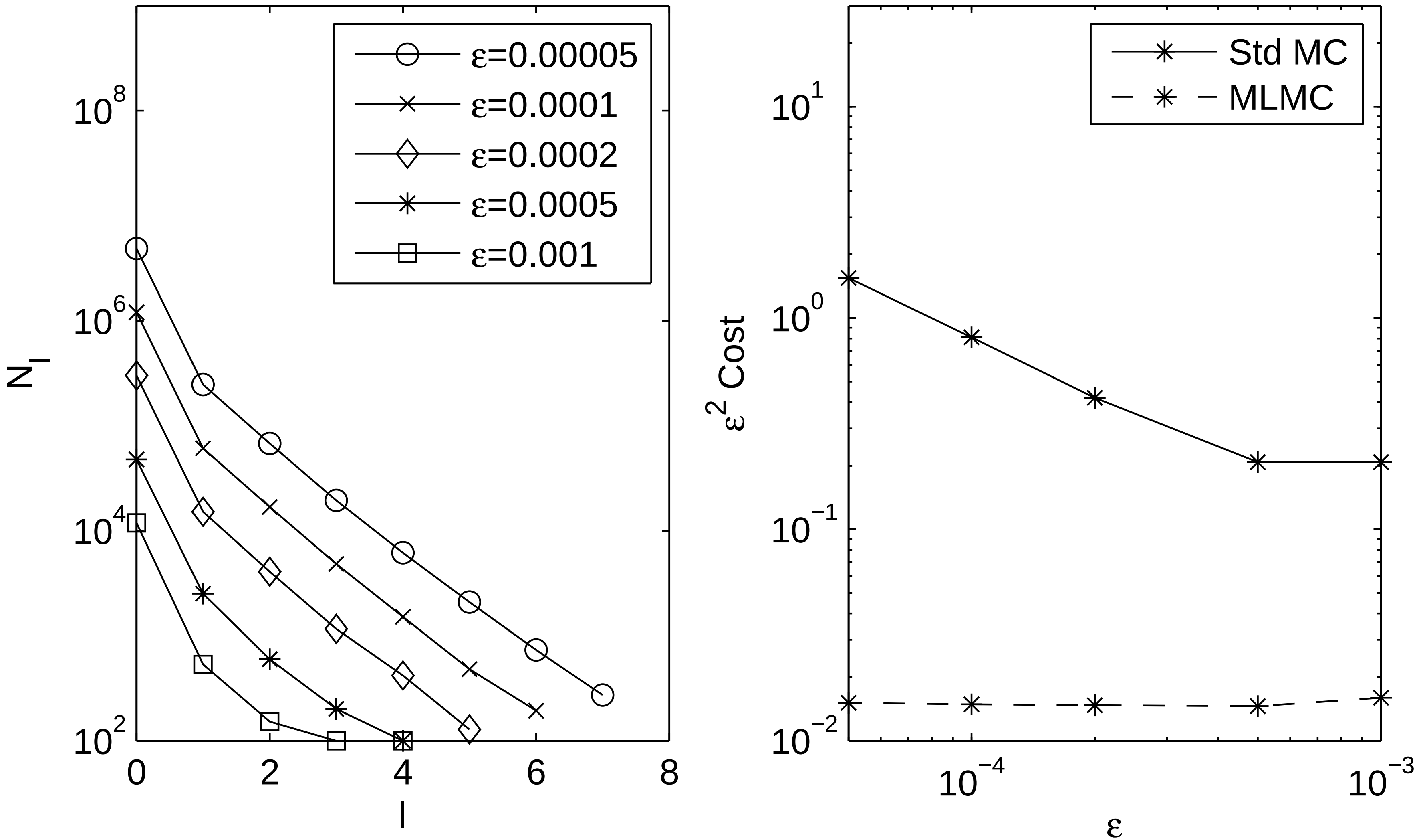
MLMC Results

GBM: Asian option, $\exp(-rT)\ \max(T^{-1}\int_0^T S(t)\,\mathrm{d}t - 1, 0)$



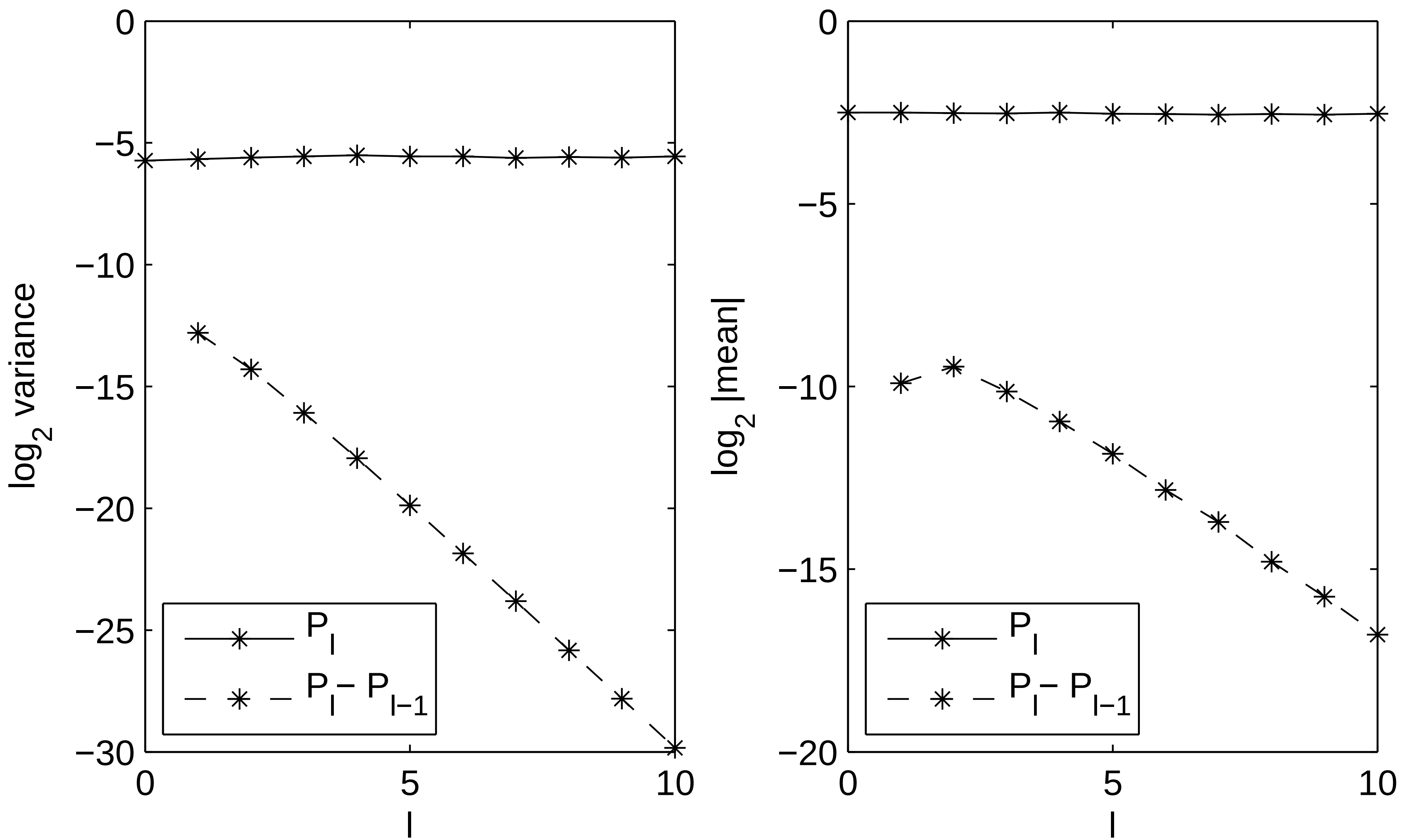
MLMC Results

GBM: Asian option, $\exp(-rT)\ \max(T^{-1}\int_0^T S(t)\,\mathrm{d}t - 1, 0)$



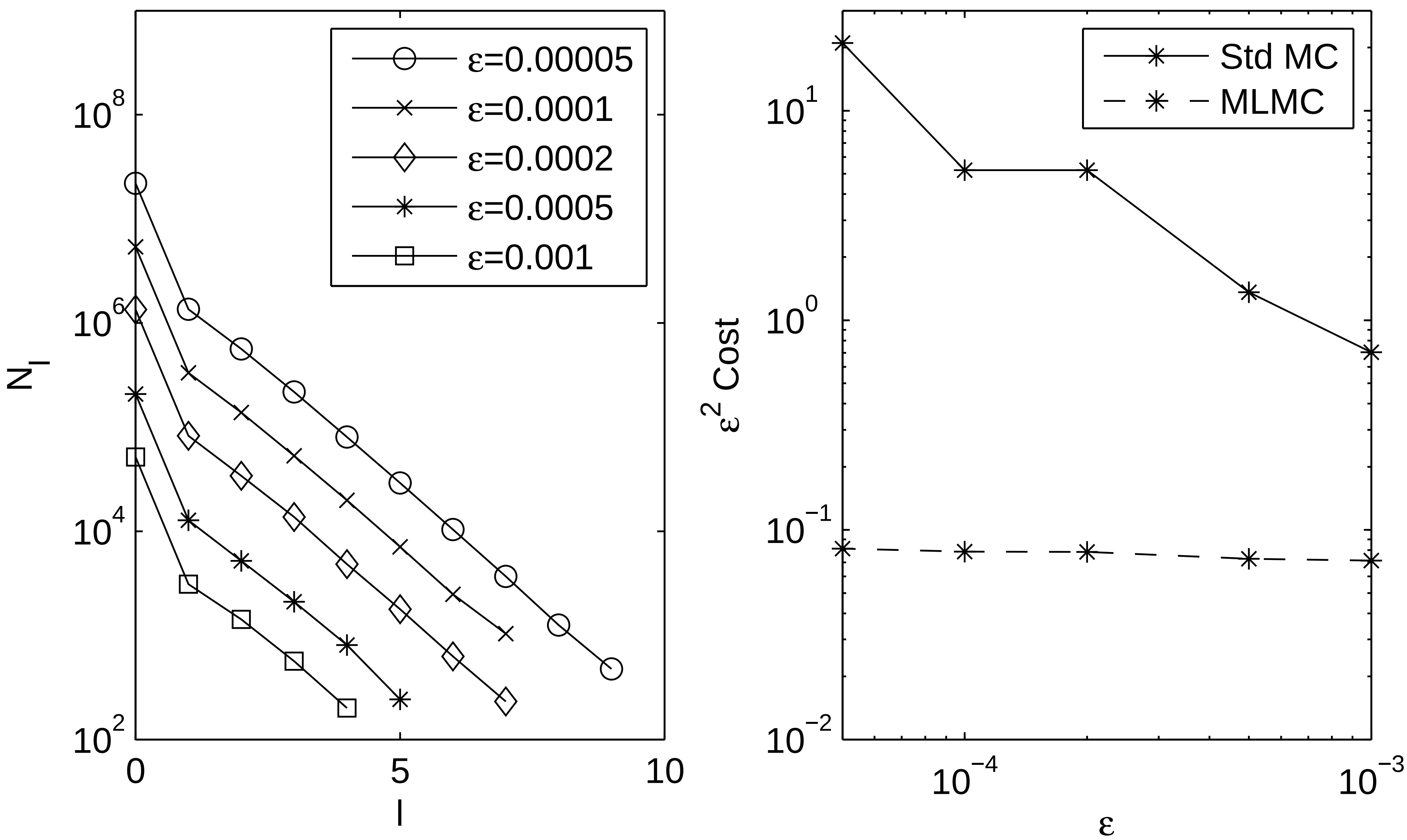
MLMC Results

GBM: lookback option, $\exp(-rT)\,(S(T) - \min_{0<t<T} S(t))$



MLMC Results

GBM: lookback option, $\exp(-rT)\,(S(T) - \min_{0<t<T} S(t))$



Extensions

1) Milstein scheme for vector SDEs

- significantly more difficult because it involves Lévy areas defined as

$$A_{jk,n} = \int_{t_n}^{t_{n+1}} (W_j(t) - W_j(t_n))\, \mathrm{d}W_k - (W_k(t) - W_k(t_n))\, \mathrm{d}W_j .$$

- $O(h)$ strong convergence if Lévy areas are simulated correctly – expensive
- $O(h^{1/2})$ strong convergence in general if Lévy areas are omitted, except if a certain commutativity condition is satisfied (useful for a number of real cases)
- Lipschitz payoffs can be handled well using antithetic variables

Extensions

2) Quasi-Monte Carlo

- standard Monte Carlo has a random sampling error proportional to $N^{-1/2}$
- Quasi-Monte Carlo uses a deterministic choice of sample "points" to achieve an error which is nearly $O(N^{-1})$ in the best cases
- Not much applicable theory because financial payoffs don't have required smoothness
- In practice, get great results using rank-1 lattice rules developed by Ian Sloan's group at UNSW
- Haven't yet tried Sobol sequences

Extensions

3) Numerical Analysis

- paper with Des Higham and Xeurong Mao (Strathclyde) on analysis of Euler discretisation with complex options
- Klaus Ritter (Darmstadt) has generalised analysis of Euler discretisation to path dependent options with Lipschitz property
- more work needed to analyse Milstein approximation

Extensions

4) "Greeks"

- this is the name given to derivatives such as $\dfrac{\partial}{\partial S_0}\,\mathbb{E}[P]$
- under certain circumstance, this is equal to $\mathbb{E}\left[\dfrac{\partial P}{\partial S_0}\right]$

  – this leads to the pathwise differentiation approach
- the multilevel approach should again work well but not tried yet
- can also incorporate the adjoint approach developed with Paul Glasserman – more efficient when many Greeks are wanted for one payoff function

Extensions

5) “vibrato” Monte Carlo

- problem with discontinuous payoffs is that small changes in path can lead to a big change in the payoff
- so far, have treated digital options using a “trick” in Paul Glasserman’s book, taking the conditional expectation one timestep before maturity, which effectively smooths the payoff
- the “vibrato” Monte Carlo idea generalises this to cases in which the conditional expectation is not known in closed form

Extensions

6) American options

- with European options, the buyer can only exercise the option at maturity, the final time T
- with American options, the buyer can exercise at any time, leading to an optimal control problem
- in PDE approaches, this is solved using a linear complementarity approach which marches backwards in time
- modifying Monte Carlo methods is much harder – an active research topic
- I have some ideas on how to incorporate the multilevel approach – hope to start a project on this soon

Extensions

7) SPDEs (stochastic PDEs)

- working with a colleague Christoph Reisinger on a financial SPDE which is a convection-diffusion PDE with a stochastic convection "velocity":

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

  – preliminary results look good

- working with Rob Scheichl (Bath) on an elliptic SPDE where the diffusivity is a log-normal stochastic field:

$$\nabla \cdot \big(\kappa(x) \nabla p\big) = 0.$$

  – again, preliminary 1D results look good.

Extensions

8) CUDA implementation on NVIDIA graphics cards

- advances in computer hardware/software are important as well as advances in mathematics
- graphics cards are very powerful parallel processors, with up to 240 cores per graphics chip (GPU)
- 2 years ago, NVIDIA introduced the CUDA development environment which uses minor extension to C/C++
- with a visiting student, Xiaoke Su, achieved $100\times$ speedup on a Monte Carlo application using 128 cores
- (more recently, achieved $50\times$ speedup for simple PDE applications, including implicit ADI time-marching)

Conclusions

Multilevel Monte Carlo method has already achieved

- improved order of complexity
- significant benefits for model problems

but much more research is needed, both theoretical and applied.

Acknowledgements:

- Paul Glasserman, Mark Broadie, Terry Lyons for early discussions
- Microsoft, EPSRC, Oxford-Man Institute for funding