Multilevel Monte Carlo Path Simulation

Mike Giles

giles@comlab.ox.ac.uk

Oxford University Computing Laboratory

First IMA Conference on Computational Finance

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 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.

Other Research

- In Dec. 2005, Ahmed Kebaier published an article in *Annals of Applied Probability* describing a two-level method which reduces the cost to $O\left(\varepsilon^{-2.5}\right)$.
- Also in Dec. 2005, Adam Speight wrote a working paper describing a similar multilevel use of control variates, but without an analysis of its complexity.
- There are also close similarities to a multilevel technique developed by Stefan Heinrich for parametric integration (*Journal of Complexity*, 1998)

Multilevel MC Approach

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

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

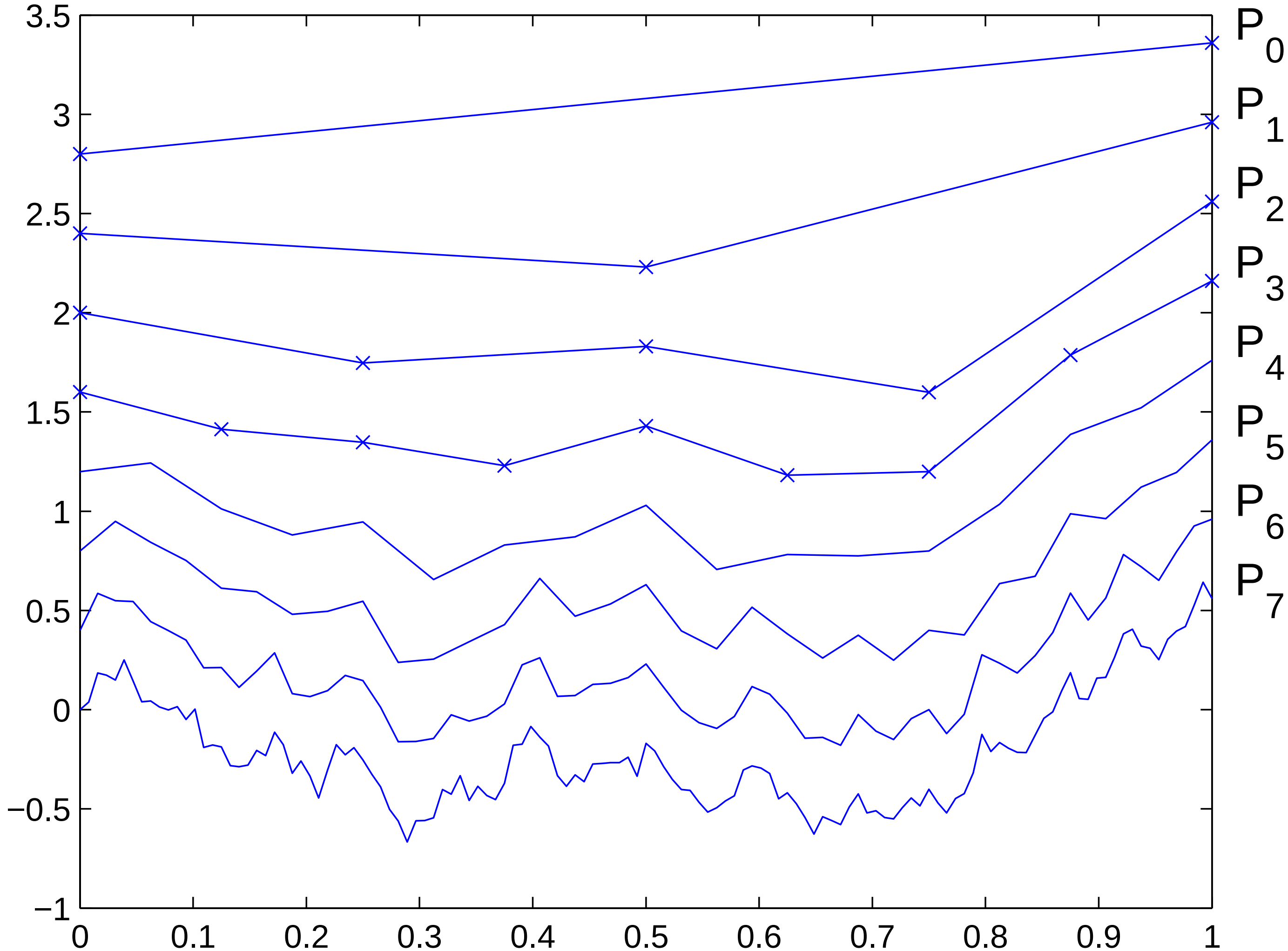
Expected value is same – aim is to reduce variance of estimator for a fixed computational cost.

Key point: approximate $E[\widehat{P}_l - \widehat{P}_{l-1}]$ 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

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

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

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

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

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:

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

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

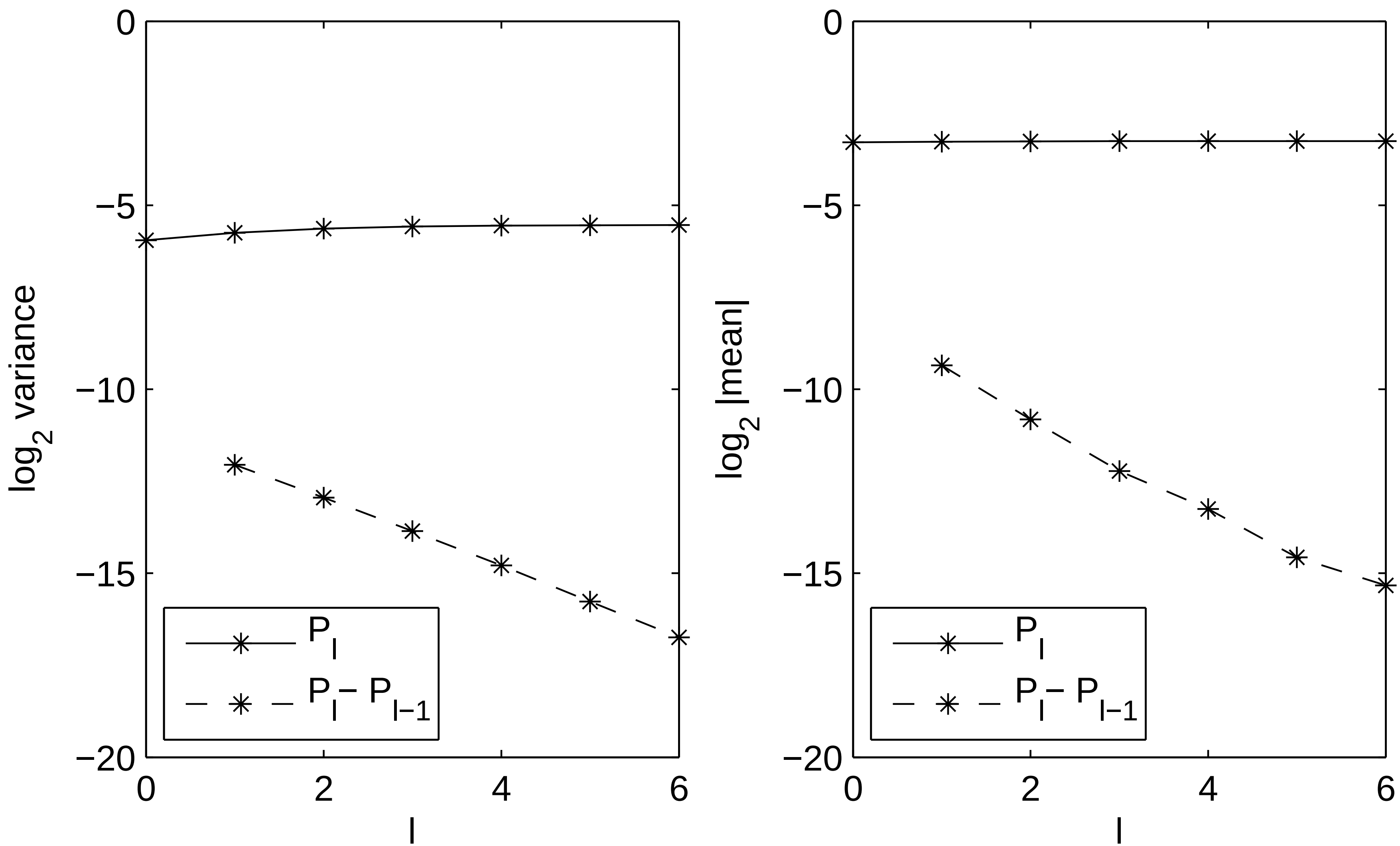
European call option with discounted payoff

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

with $K\!=\!1$.

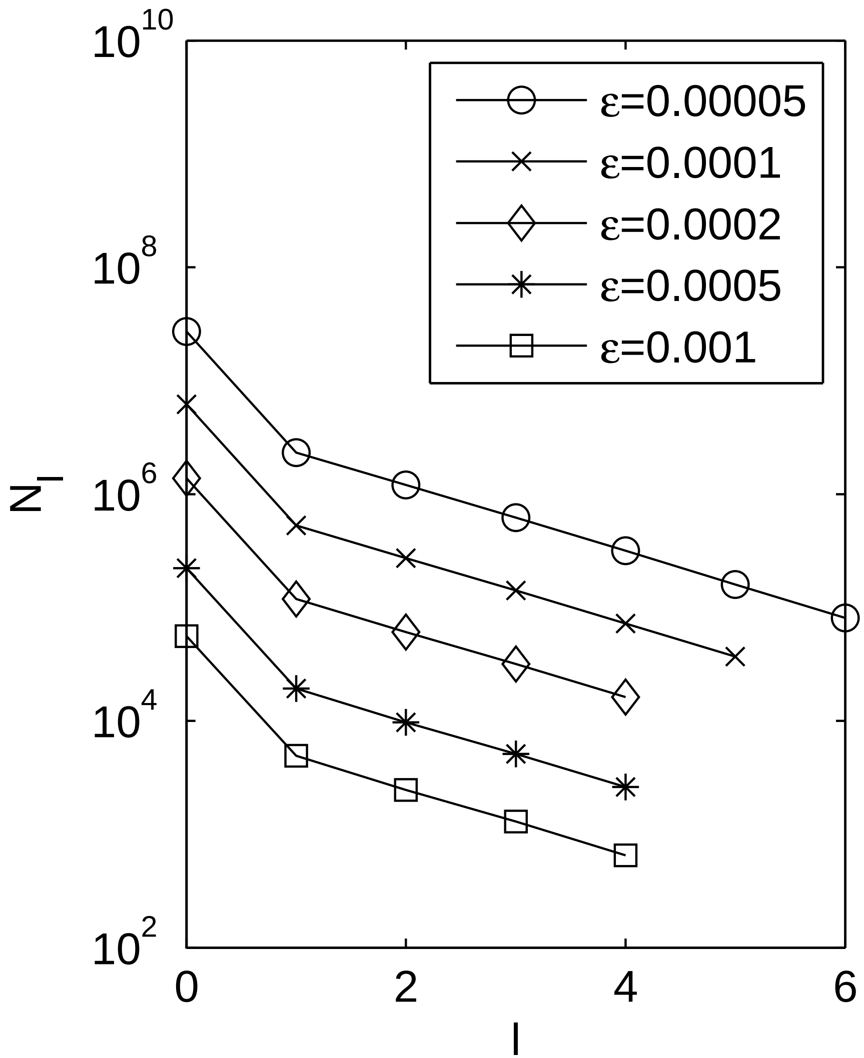
Results

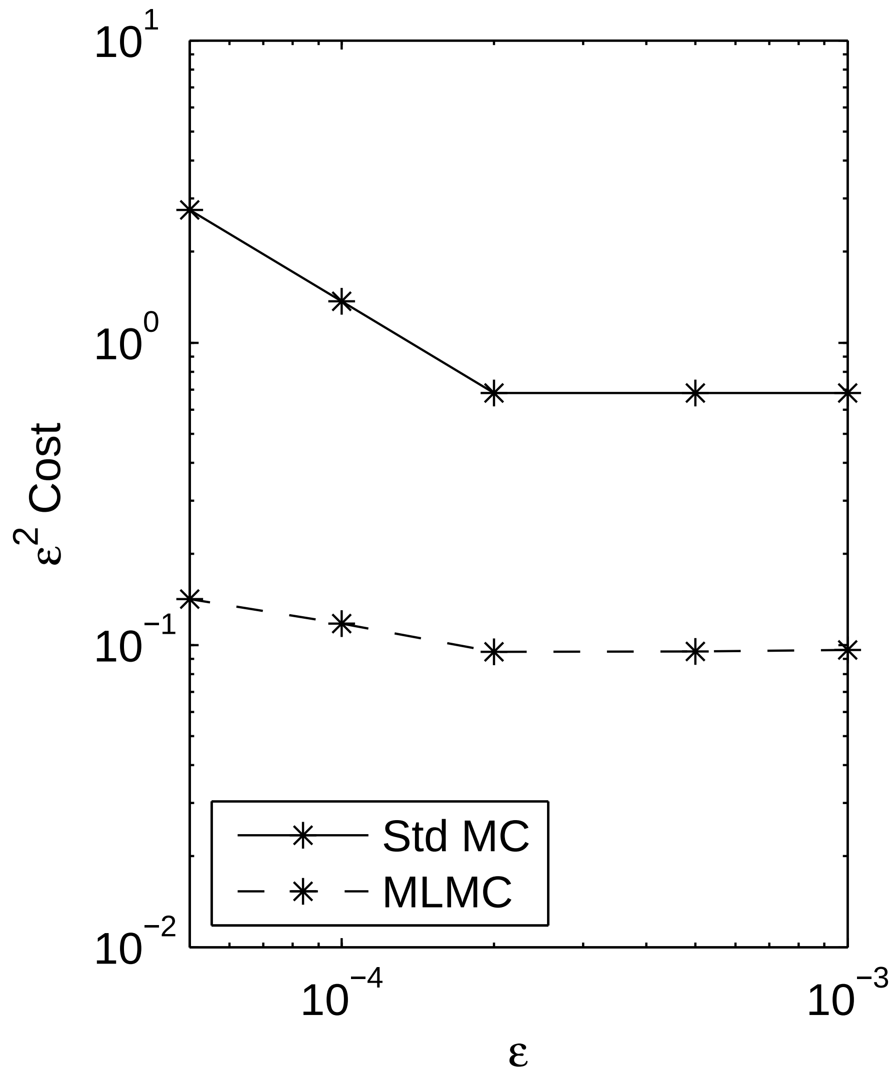
GBM: European call, $\exp(-rT)\ \max(S(T)-K,0)$



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

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

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

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

Generic scalar SDE:

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

Milstein scheme:

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

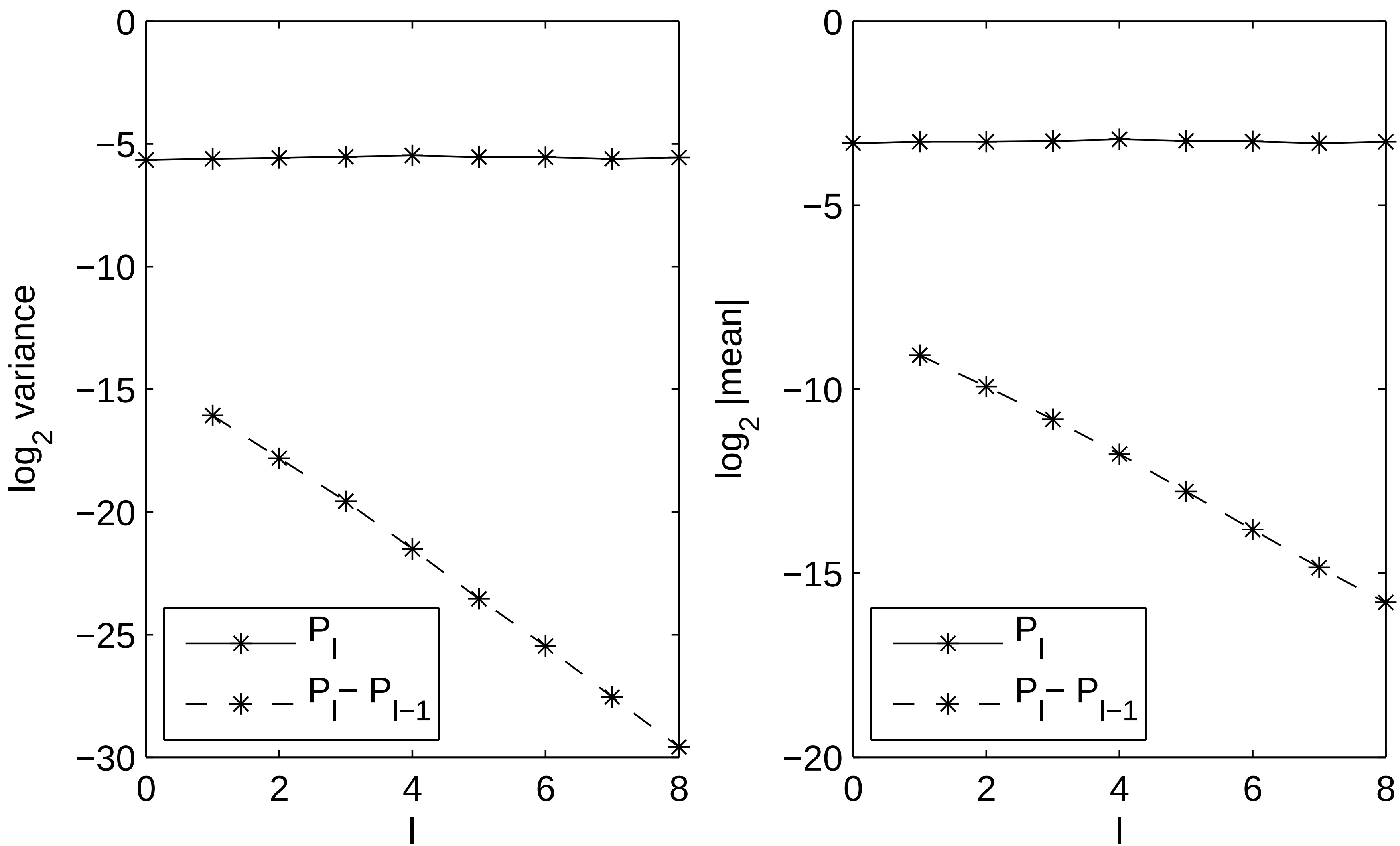
Milstein Scheme

In scalar case:

- $O(h)$ strong convergence
- $O(\varepsilon^{-2})$ complexity for Lipschitz payoffs – trivial
- $O(\varepsilon^{-2})$ complexity for Asian, lookback, barrier and digital options using carefully constructed estimators, based on Brownian interpolation
- key idea: within each timestep, model the behaviour as simple Brownian motion conditional on the two end-points – analytic results exist for distribution of min/max/average

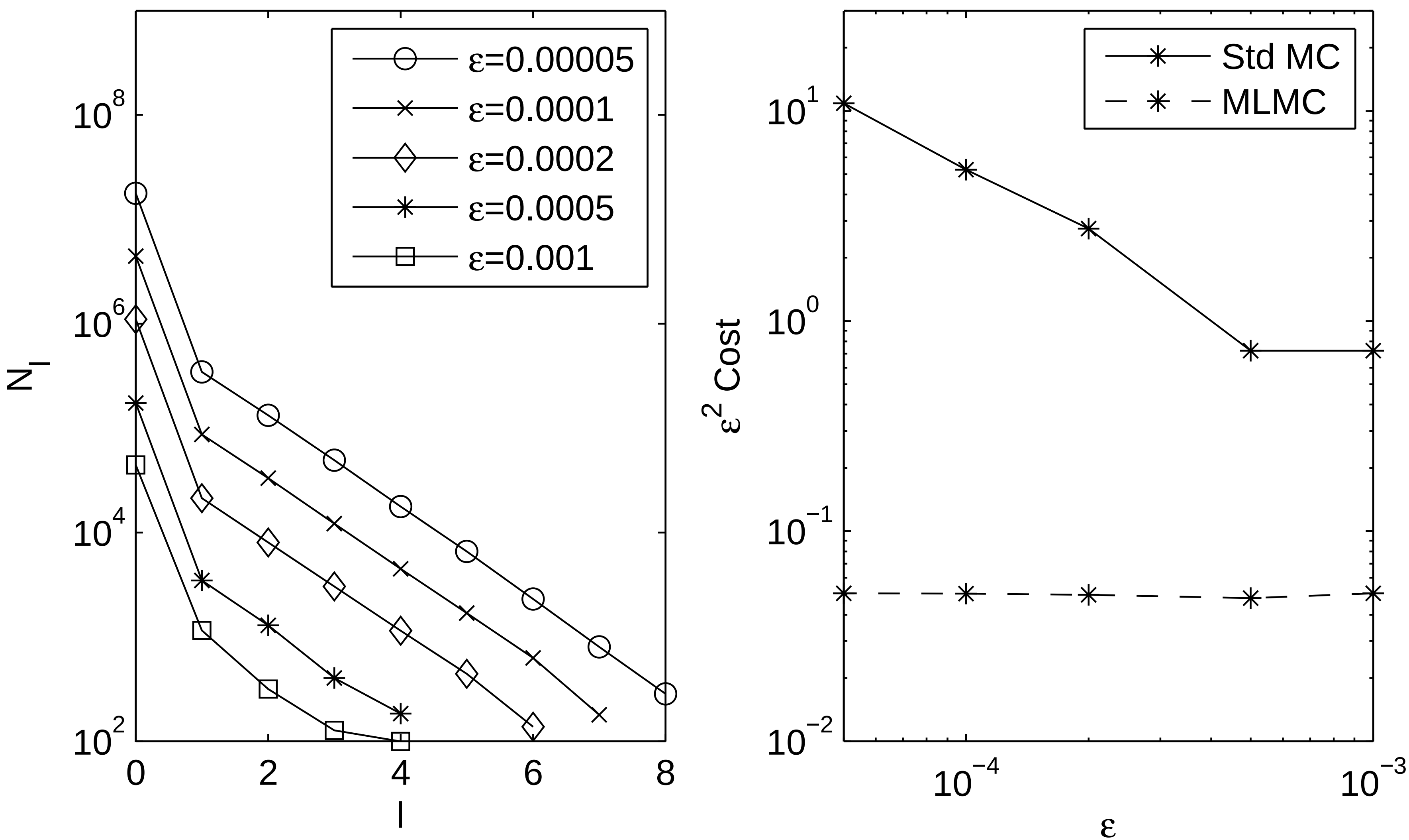
Results

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



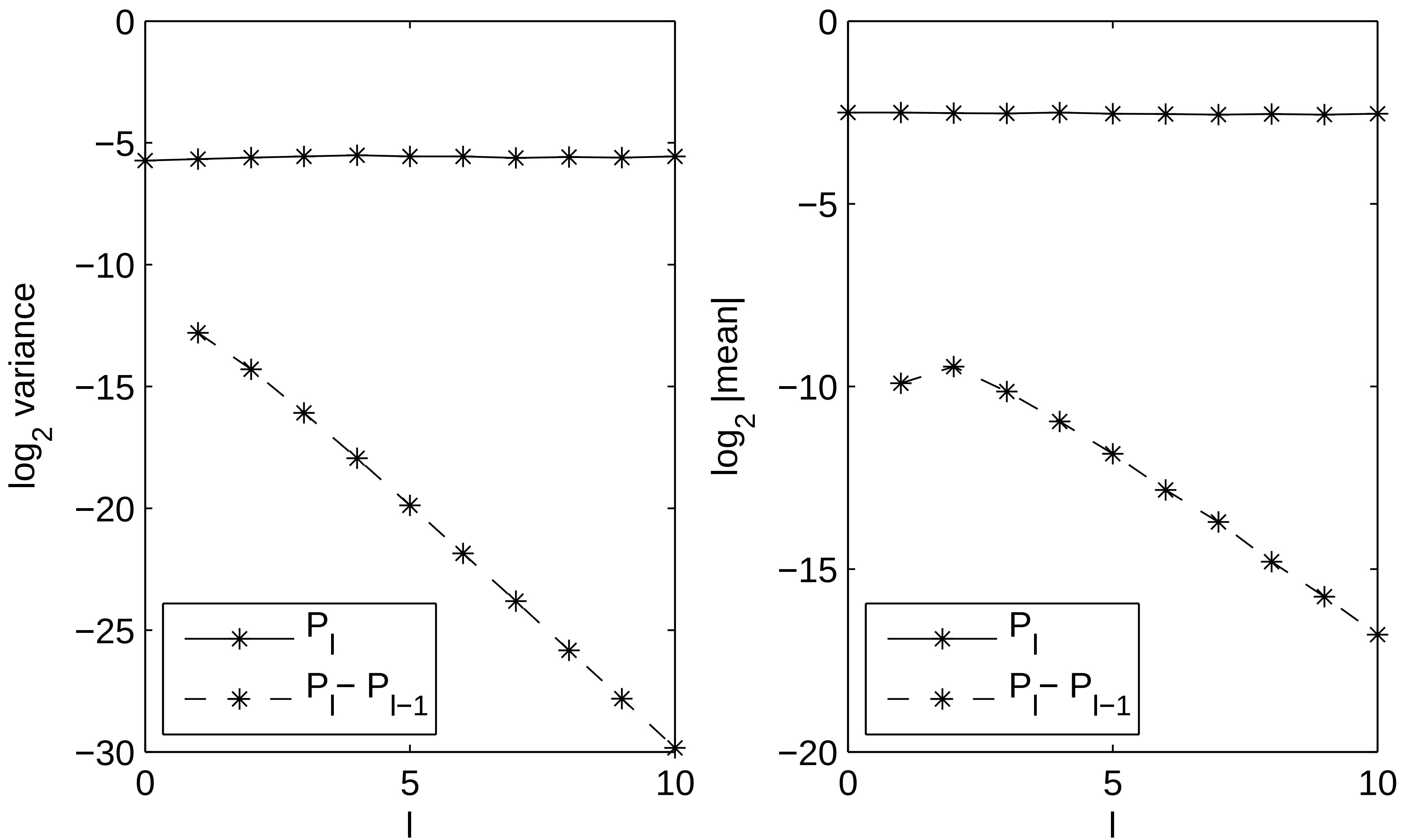
Results

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



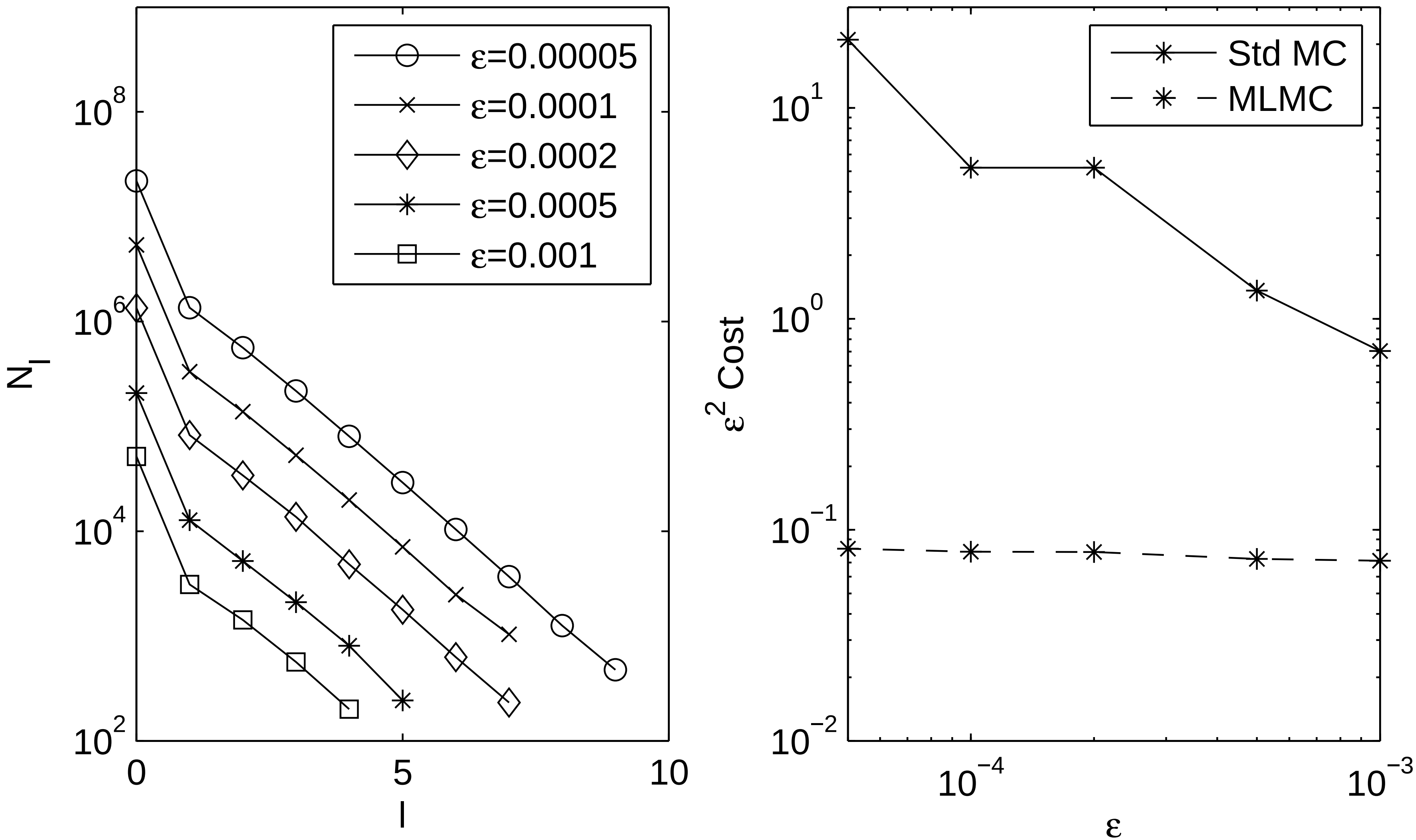
Results

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



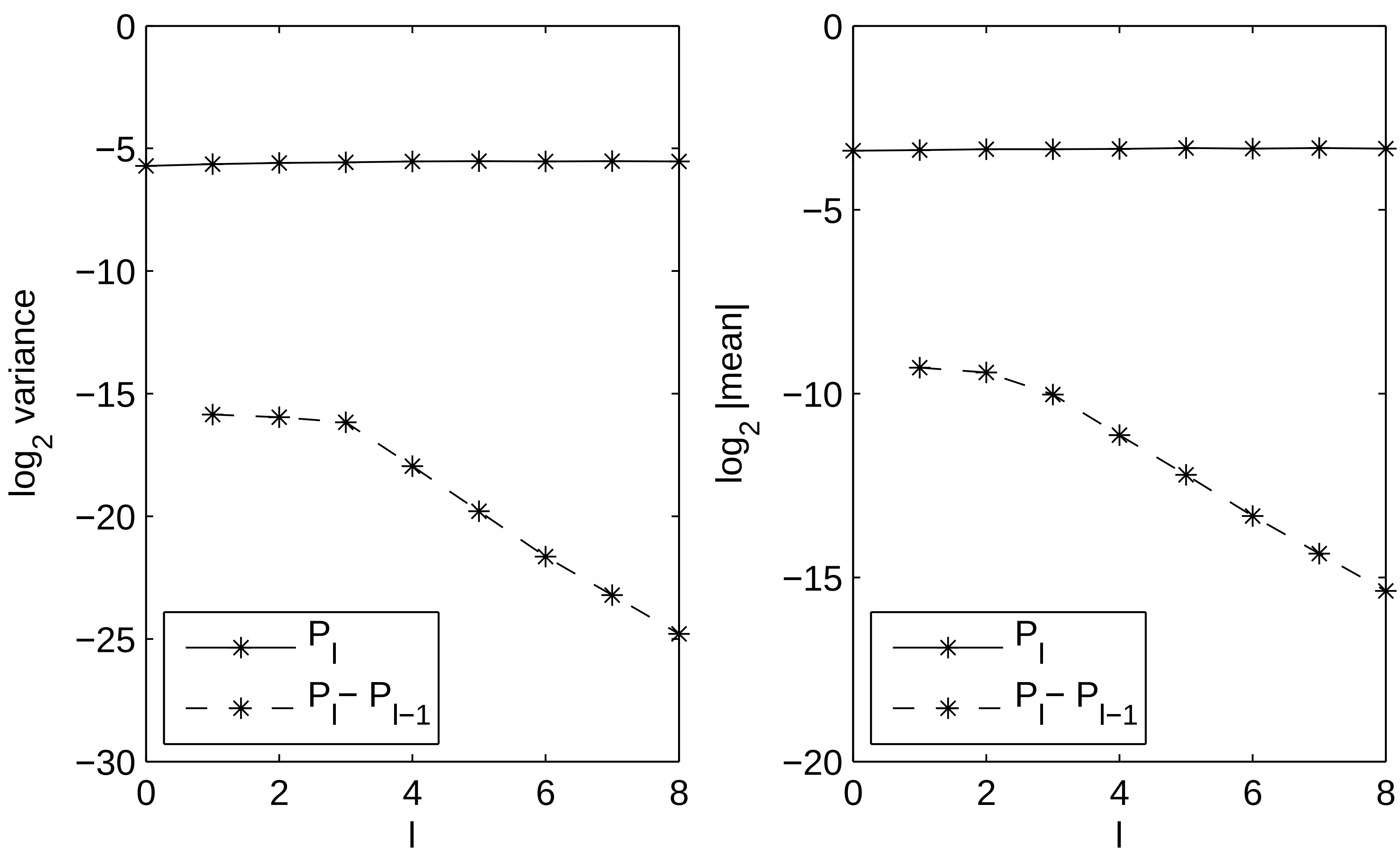
Results

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



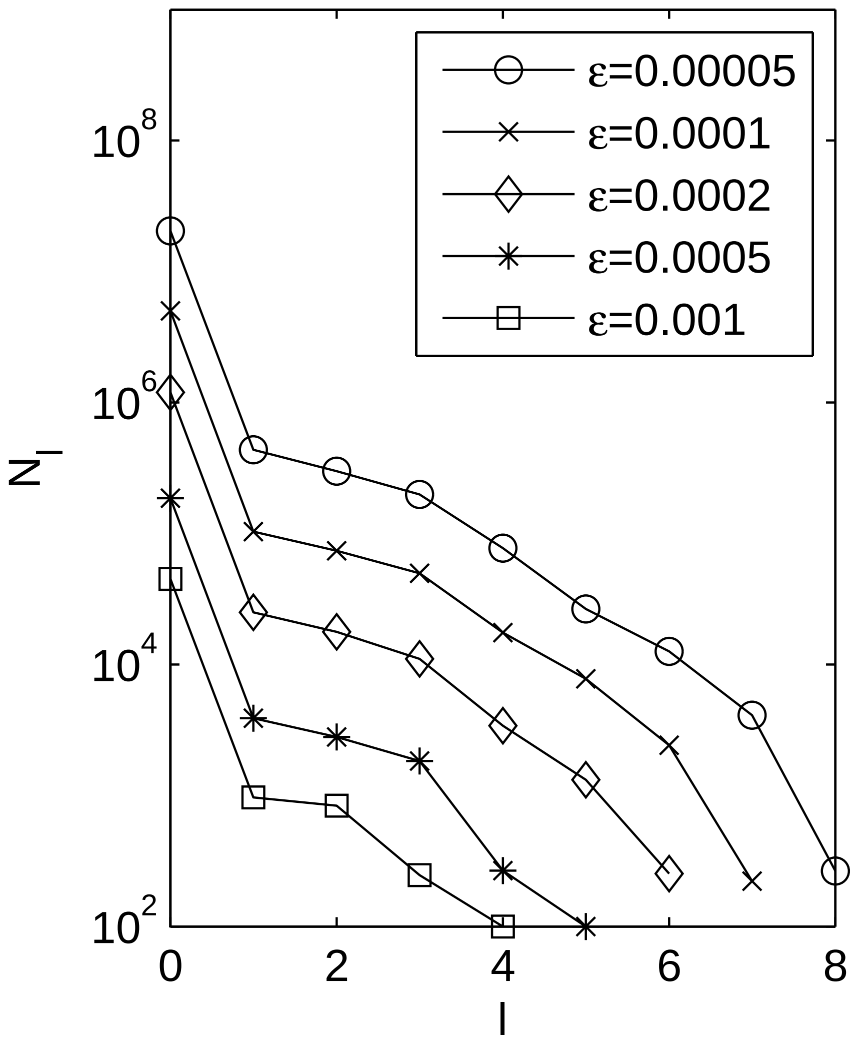
Results

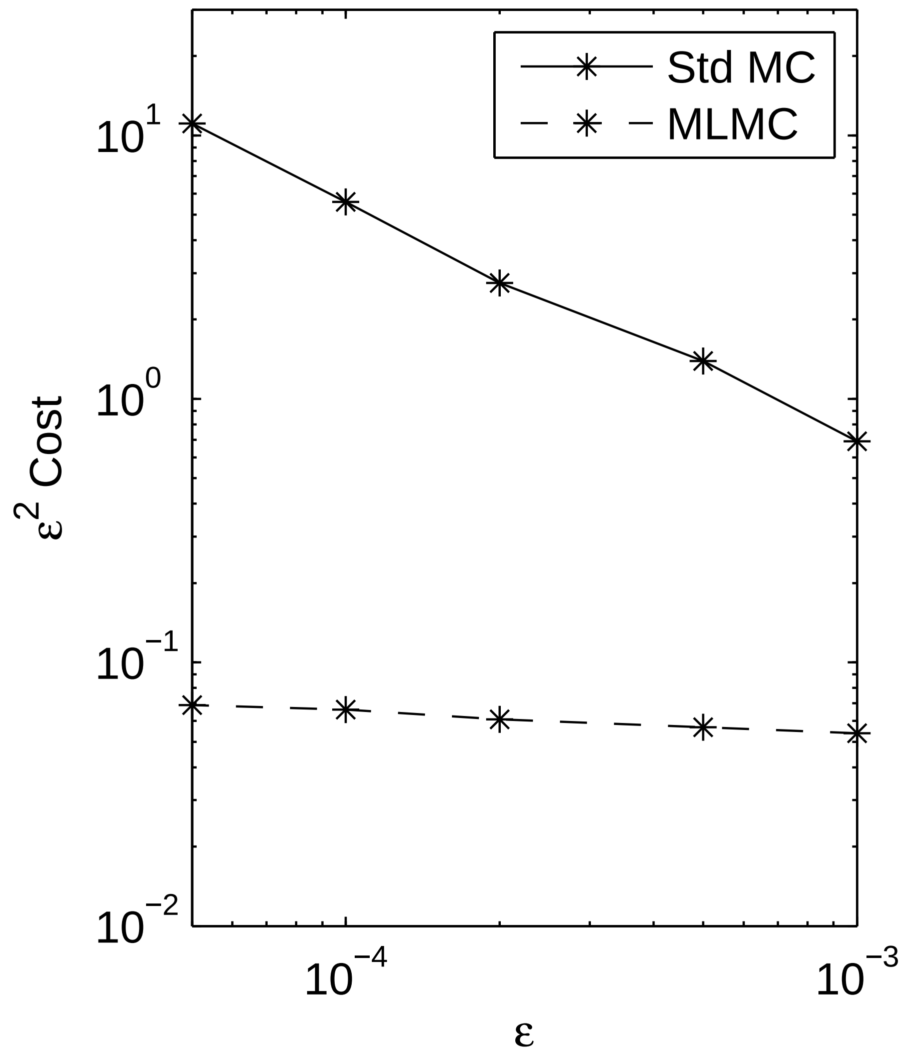
GBM: barrier option, $\exp(-rT)\,\mathbf{1}_{\min S(t)>B}\;\max(S(T)-K,0)$



Results

GBM: barrier option, $\exp(-rT)\, \mathbf{1}_{\min S(t)>B}\; \max(S(T)-K,0)$





Milstein Scheme

Generic vector SDE:

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

with correlation matrix $\Omega(S,t)$ between elements of $\mathrm{d}W(t)$.

Milstein scheme:

$$\begin{aligned}\widehat{S}_{i,n+1} &= \widehat{S}_{i,n} + a_i\,h + b_{ij}\,\Delta W_{j,n} \\ &\quad + \tfrac{1}{2}\,\frac{\partial b_{ij}}{\partial S_l}\,b_{lk}\left(\Delta W_{j,n}\,\Delta W_{k,n} - h\,\Omega_{jk} - A_{jk,n}\right)\end{aligned}$$

with implied summation, and Lévy areas defined as

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

Milstein Scheme

In vector case:

- $O(h)$ 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
- Other cases may require approximate simulation of Lévy areas

Milstein Scheme

If $\widehat{S}^c_n$ satisfies

$$\widehat{S}^c_{n+1} = R(\widehat{S}^c_n),$$

and $\widehat{S}^f_n$ satisfies

$$\widehat{S}^f_{n+1} = R(\widehat{S}^f_n) + g_n.$$

then if $g_n \ll 1$, putting $\widehat{S}^f_n = \widehat{S}^c_n + \widehat{D}_n$ and linearising gives

$$\widehat{D}_{n+1} = \frac{\partial R}{\partial S}\, \widehat{D}_n + g_n.$$

- $\widehat{S}^c_n$ represents calculation using timestep $2h$
- $\widehat{S}^f_n$ represents calculation using two timesteps of size h

Milstein Scheme

To leading order, error analysis gives

$$g_{i,n} = \frac{h}{2} \, \frac{\partial b_{ij}}{\partial S_l} \, b_{lk} \left(Y_{j,n} Z_{k,n} - Y_{k,n} Z_{j,n} \right).$$

where

$$\Delta W^f_n = \tfrac{1}{2}\sqrt{2h}\,(Y_n + Z_n), \quad \Delta W^f_{n+\frac{1}{2}} = \tfrac{1}{2}\sqrt{2h}\,(Y_n - Z_n).$$

i.e. Y_n is standard $N(0,1)$ variable used to construct coarse path, and Z_n is $N(0,1)$ variable for Brownian Bridge construction of fine path.

Note: independence implies that

$$E[g_n] = 0 \quad \Longrightarrow \quad E[\widehat{D}_n] = 0.$$

Milstein Scheme

Option 1: use control variate

Define

$$\widehat{Y}_l = N_l^{-1} \sum_{i=1}^{N_l} \left(\widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)} - \frac{\partial f}{\partial S}\, \widehat{D}_{T/2h}^{(i)} \right),$$

The control variate has zero mean and cancels out the leading order variation so that

$$V\left[\widehat{P}_l - \widehat{P}_{l-1} - \frac{\partial f}{\partial S}\, \widehat{D}_{T/2h} \right] = O(h^2)$$

for twice differentiable payoffs (and $O(h^{3/2})$ for usual Lipschitz payoffs?)

Milstein Scheme

Option 2: use antithetic variables

Define

$$\widehat{Y}_l = N_l^{-1} \sum_{i=1}^{N_l} \left(\tfrac{1}{2} \left(\widehat{P}_l^{(i)} + \widehat{P}_l^{(i)*} \right) - \widehat{P}_{l-1}^{(i)} \right),$$

where $\widehat{P}_l^{(i)*}$ is based on the same coarse path with Z_n replaced by $-Z_n$, which leads to cancellation of leading order error proportional to Z_n.

Very simple to implement (but slightly more costly?)

Results

Heston model:

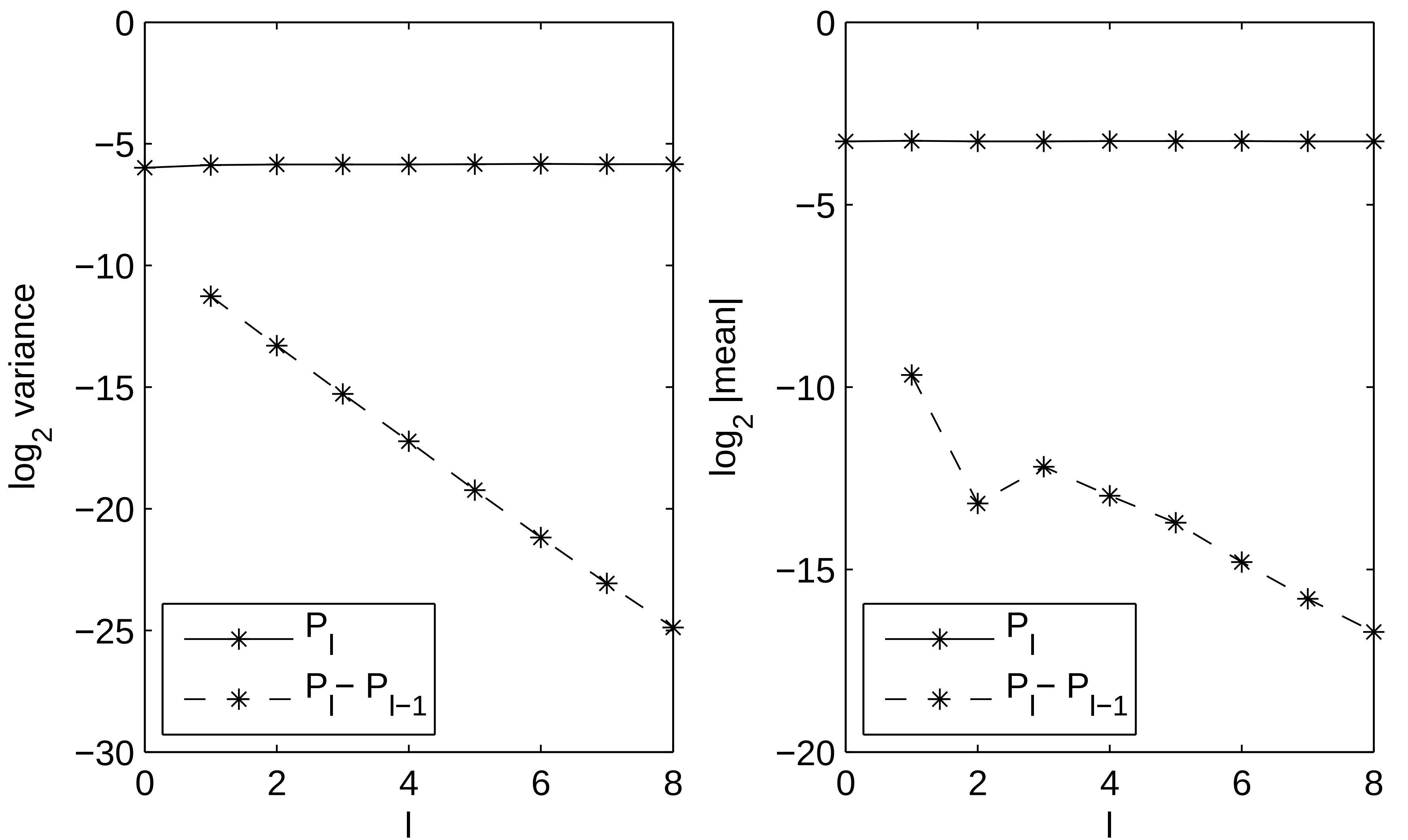
$$\begin{aligned} \mathrm{d}S &= r\, S\, \mathrm{d}t + \sqrt{V}\, S\, \mathrm{d}W_1, \qquad 0<t<T \\ \mathrm{d}V &= \lambda\, (\sigma^2 - V)\, \mathrm{d}t + \xi\, \sqrt{V}\; \mathrm{d}W_2, \end{aligned}$$

$T=1, \quad S(0)=1, \quad V(0)=0.04, \quad r=0.05,$

$\sigma=0.2, \quad \lambda=5, \quad \xi=0.25, \quad \rho=-0.5$

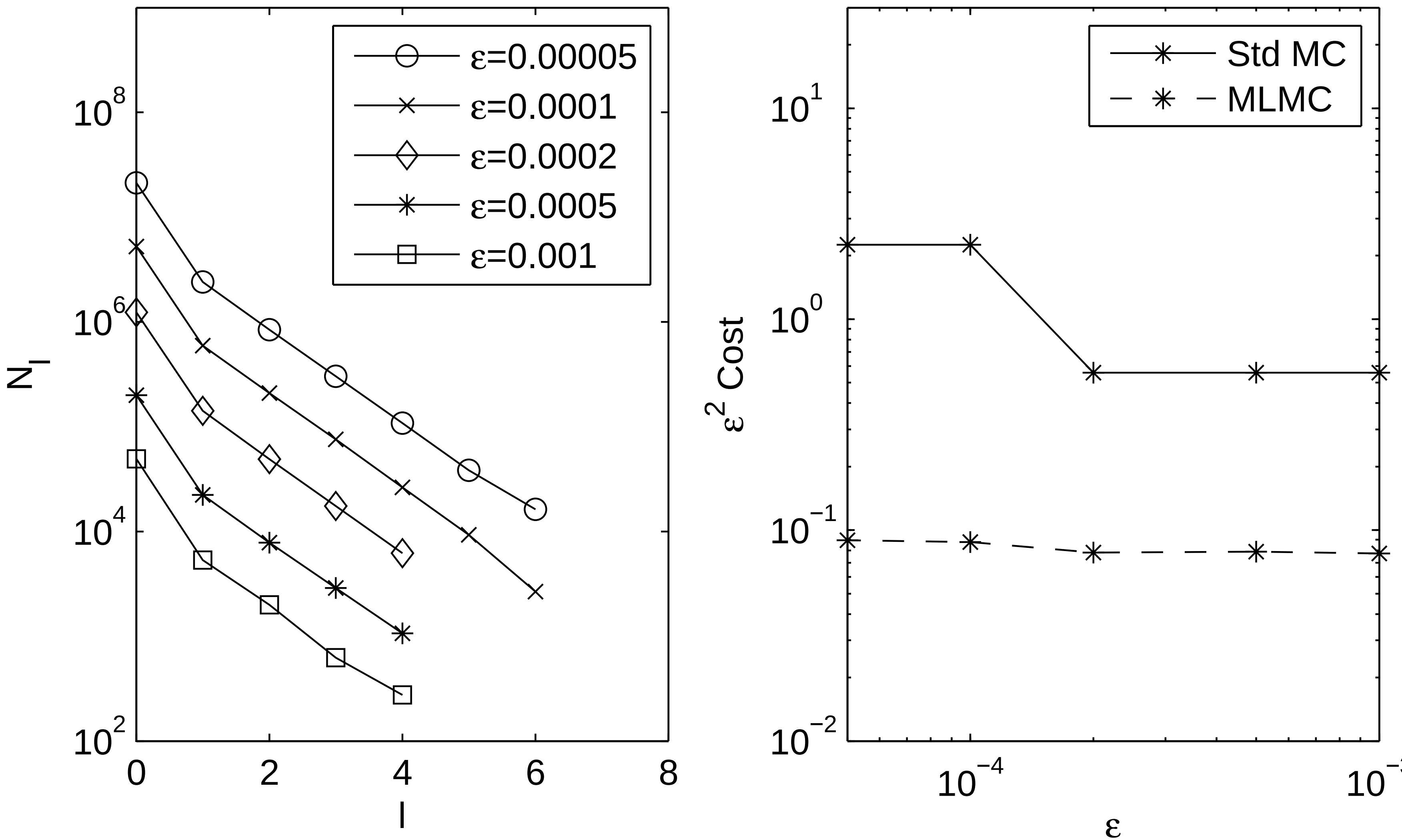
Results

Heston model: European call



Results

Heston model: European call



Conclusions

Results so far:

- (much) improved order of complexity
- (fairly) easy to implement
- significant benefits for model problems

However:

- lots of scope for further improvement
- need to test ideas on “real” finance applications

Future Work

- multi-dimensional SDEs with barrier and digital options
- quasi-Monte Carlo integration
  (F. Kuo, I. Sloan – UNSW)
- Greeks and calibration
  (P. Glasserman – Columbia Business School)
- numerical analysis
  (D. Higham, X. Mao – Strathclyde)
- real finance applications
- parallel implementation on hyper-core chips
  (ClearSpeed, nVidia – 96-128 cores)

Working Papers

M.B. Giles, "Multilevel Monte Carlo path simulation", Numerical Analysis Report NA-06/03

M.B. Giles, "Improved multilevel convergence using the Milstein scheme", Numerical Analysis Report NA-06/22

`www.comlab.ox.ac.uk/mike.giles/finance.html`

Email: `giles@comlab.ox.ac.uk`

Acknowledgements:

- Paul Glasserman and Mark Broadie for early feedback
- Microsoft for current research funding