Multilevel Monte Carlo Analysis

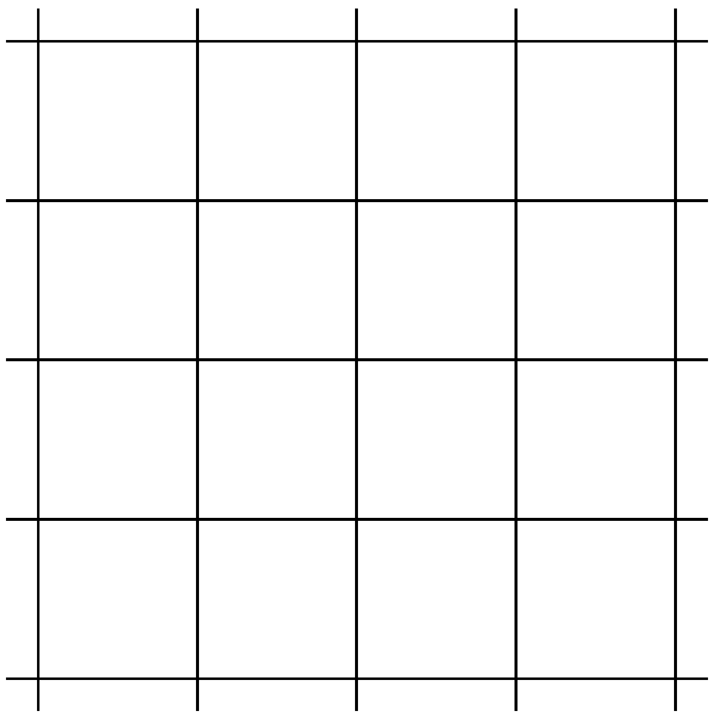
Mike Giles

giles@comlab.ox.ac.uk

Oxford University Computing Laboratory

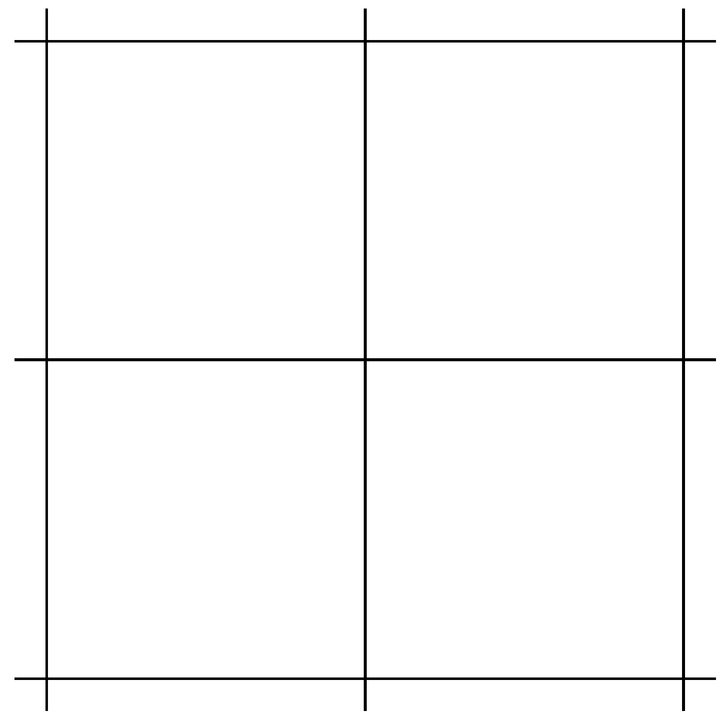
Multigrid

Multigrid is a technique which is often used in solving PDE discretisations:



Fine grid

more accurate
more expensive



Coarse grid

less accurate
less expensive

Multigrid

Multigrid combines calculations on a sequence of grids, each twice as fine as the previous, to get the accuracy of the finest grid at a much lower computational cost.

We will now apply the same concept to Monte Carlo path calculations.

Generic Problem

Stochastic ODE with general drift and volatility terms:

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

($W(t)$ is a Wiener variable with the properties that for any $q\!<\!r\!<\!s\!<\!t$, $W(t)\!-\!W(s)$ is Normally distributed with mean 0 and variance $t\!-\!s$, independent of $W(r)\!-\!W(q)$.)

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,

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

Standard MC Approach

Euler discretisation with timestep h:

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

where ΔW_n are Normal with mean 0, variance Δt.

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 $O\left(\varepsilon^{-2}(\log\varepsilon)^2\right)$

(In 2005, Ahmed Kebaier published a two-level method which reduces the cost to $O\left(\varepsilon^{-2.5}\right)$, equivalent to a single application of Richardson extrapolation.)

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

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 $\sqrt{V_l\, h_l}$.

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 $O(\varepsilon^{-2} L^2) = O(\varepsilon^{-2} (\log \varepsilon)^2)$.

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*

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

Results

Geometric Brownian motion:

$$\mathrm{d}S = r\,S\,\mathrm{d}t + \sigma\,S\,\mathrm{d}W, \qquad 0 < t < 1,$$

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

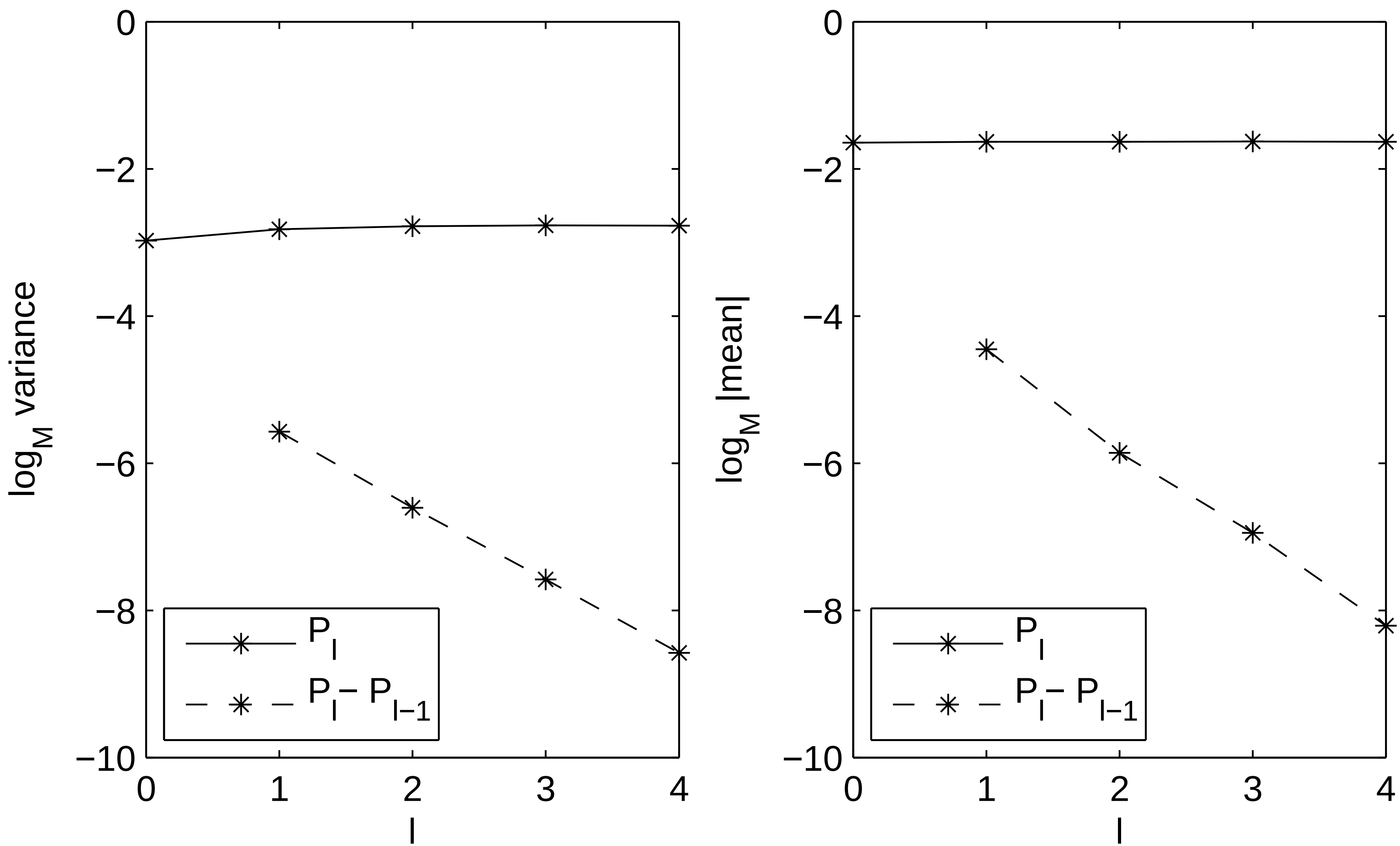
Heston model:

$$\begin{aligned} \mathrm{d}S &= r\,S\,\mathrm{d}t + \sqrt{V}\,S\,\mathrm{d}W_1, \qquad 0 < t < 1 \\ \mathrm{d}V &= \lambda\,(\sigma^2 - V)\,\mathrm{d}t + \xi\,\sqrt{V}\,\mathrm{d}W_2, \end{aligned}$$

$S(0)\!=\!1$, $V(0)\!=\!0.04$, $r\!=\!0.05$, $\sigma\!=\!0.2$, $\lambda\!=\!5$, $\xi\!=\!0.25$
and correlation $\rho\!=\!-0.5$ between $\mathrm{d}W_1$ and $\mathrm{d}W_2$

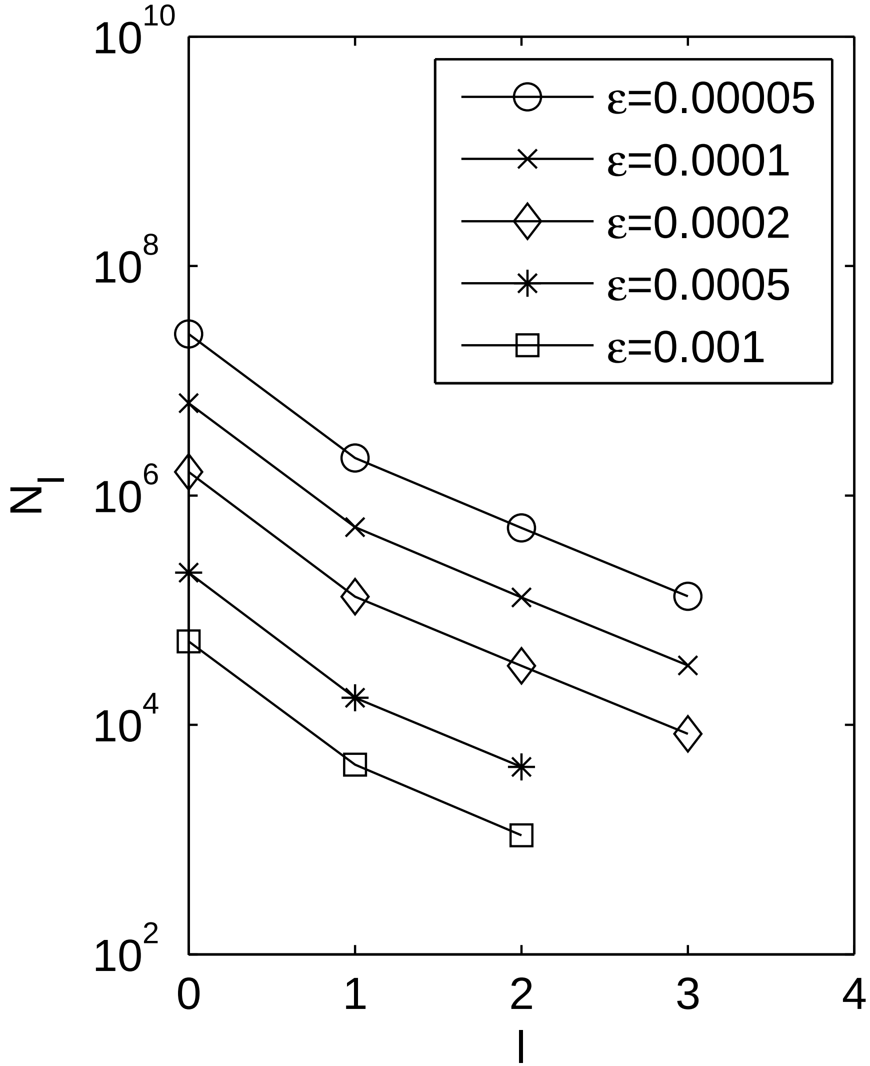
Results

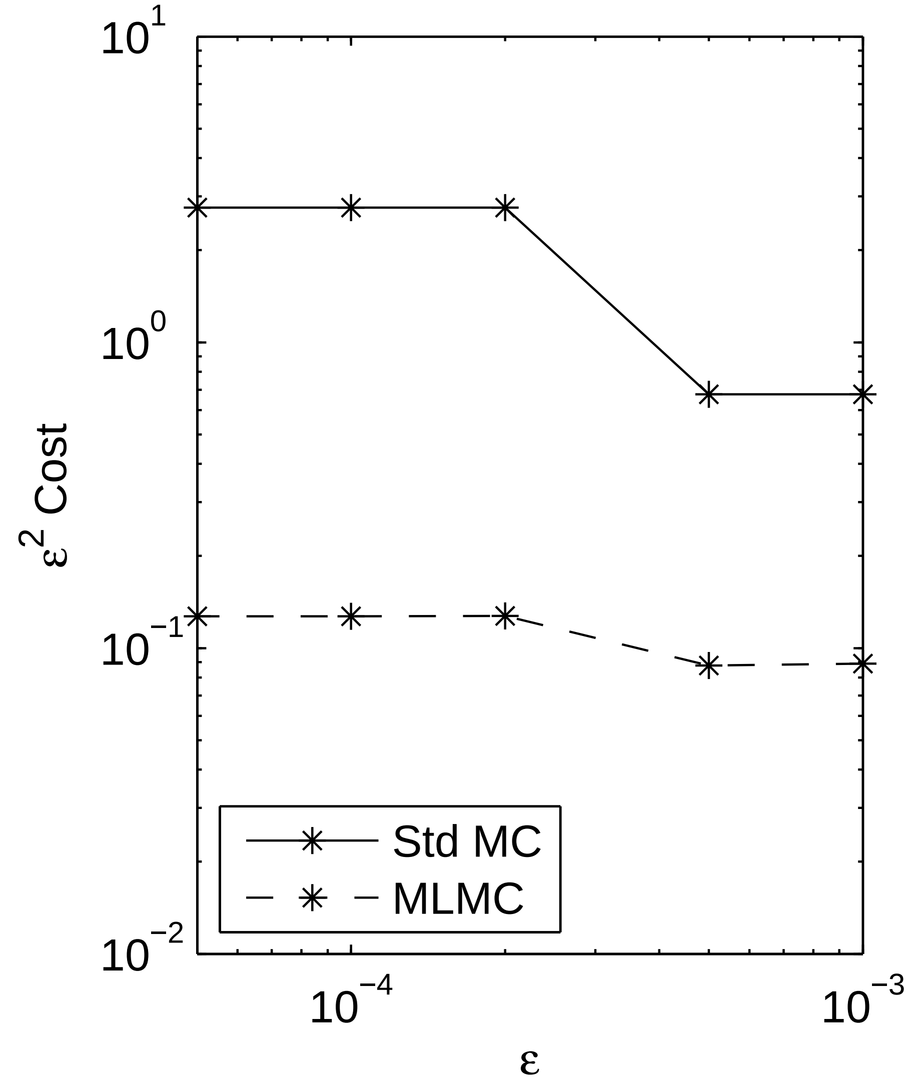
GBM: European call, $\max(S(1)-1,0)$



Results

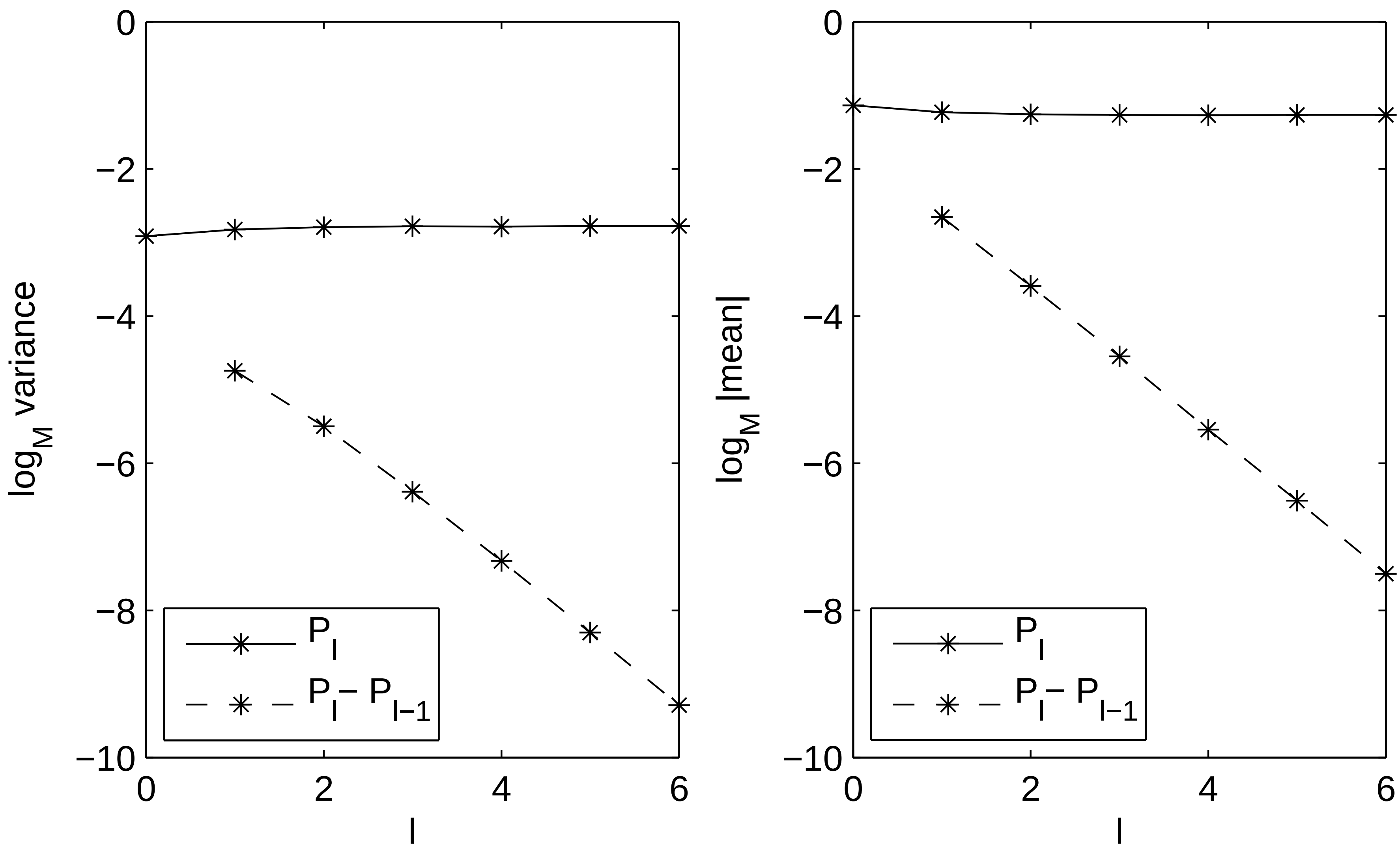
GBM: European call, $\max(S(1)-1, 0)$





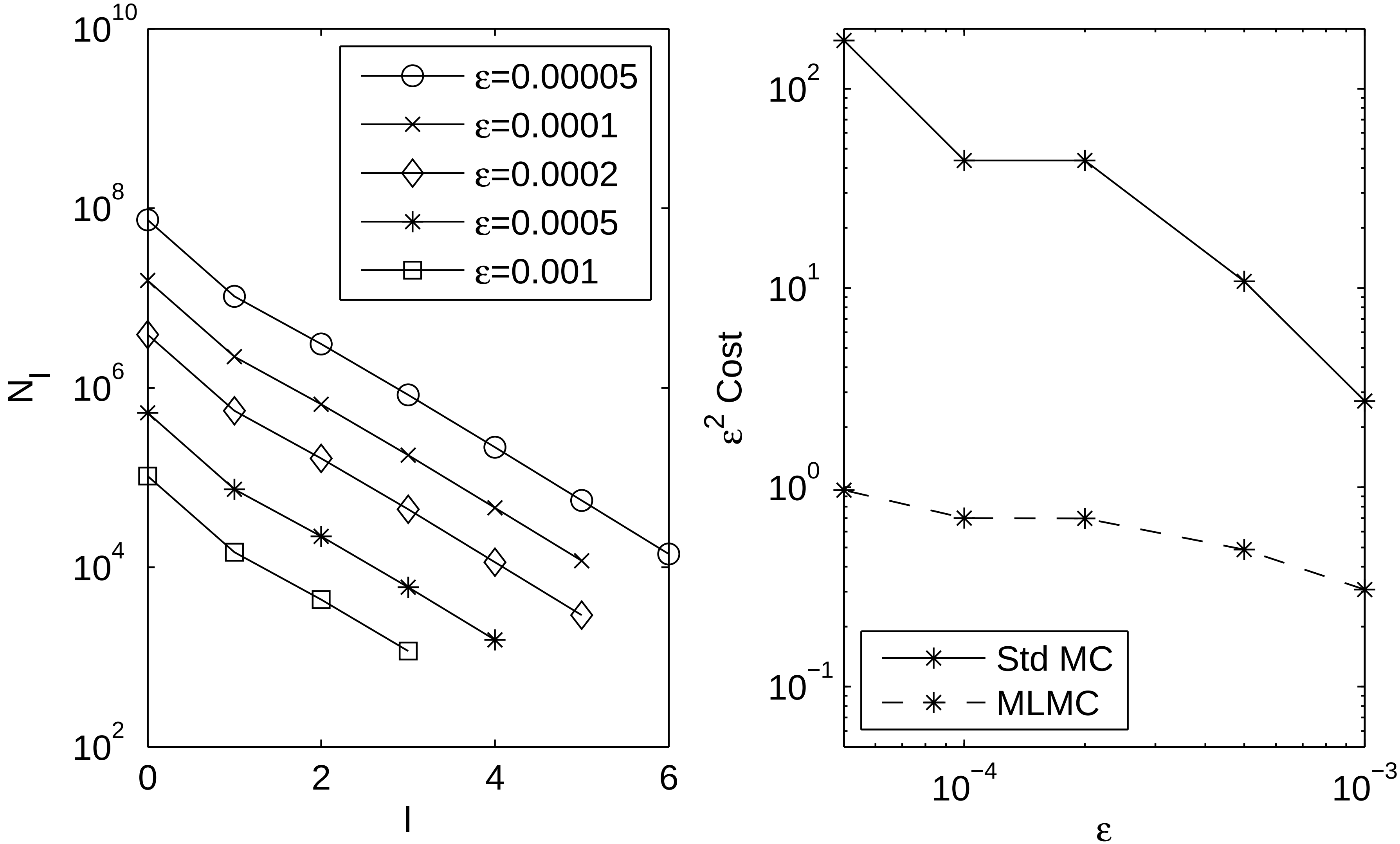
Results

GBM: lookback option, $S(1) - \min_{0<t<1} S(t)$



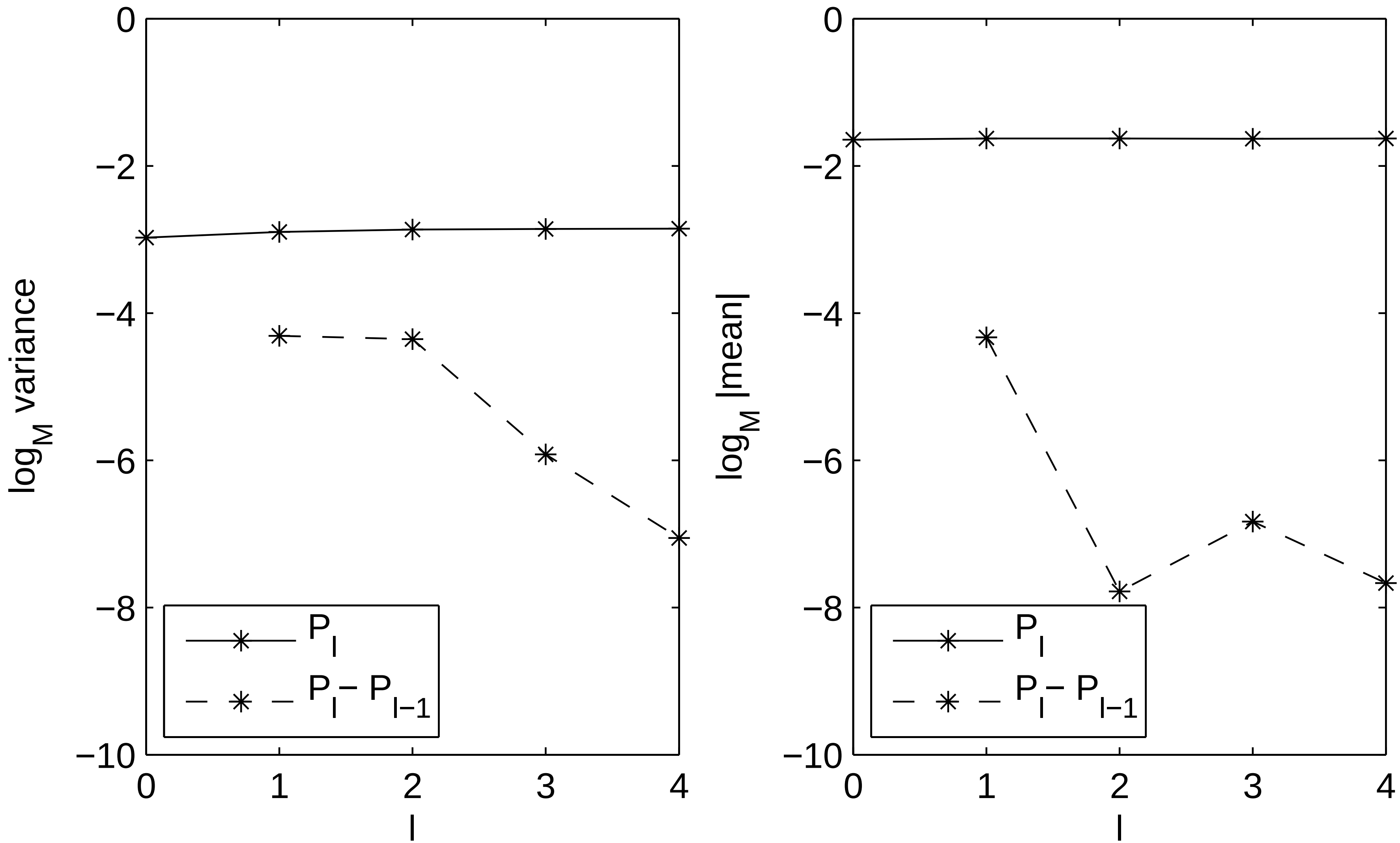
Results

GBM: lookback option, $S(1) - \min_{0<t<1} S(t)$



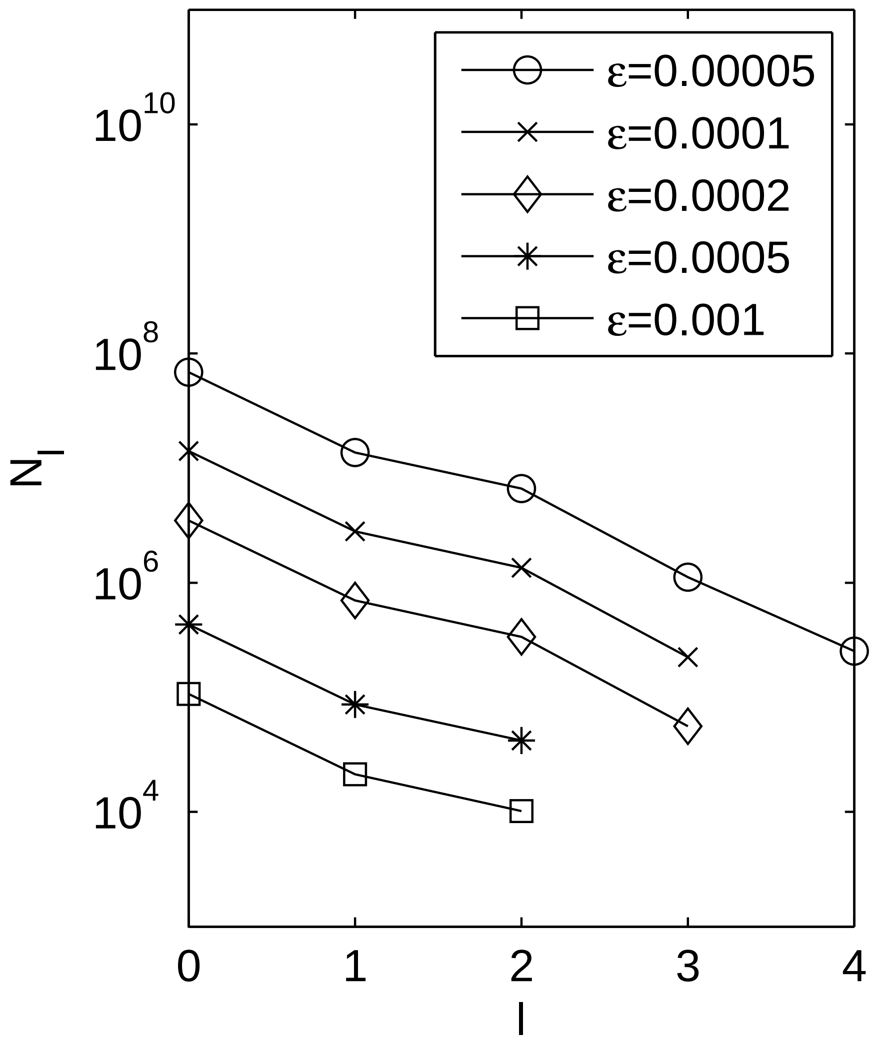
Results

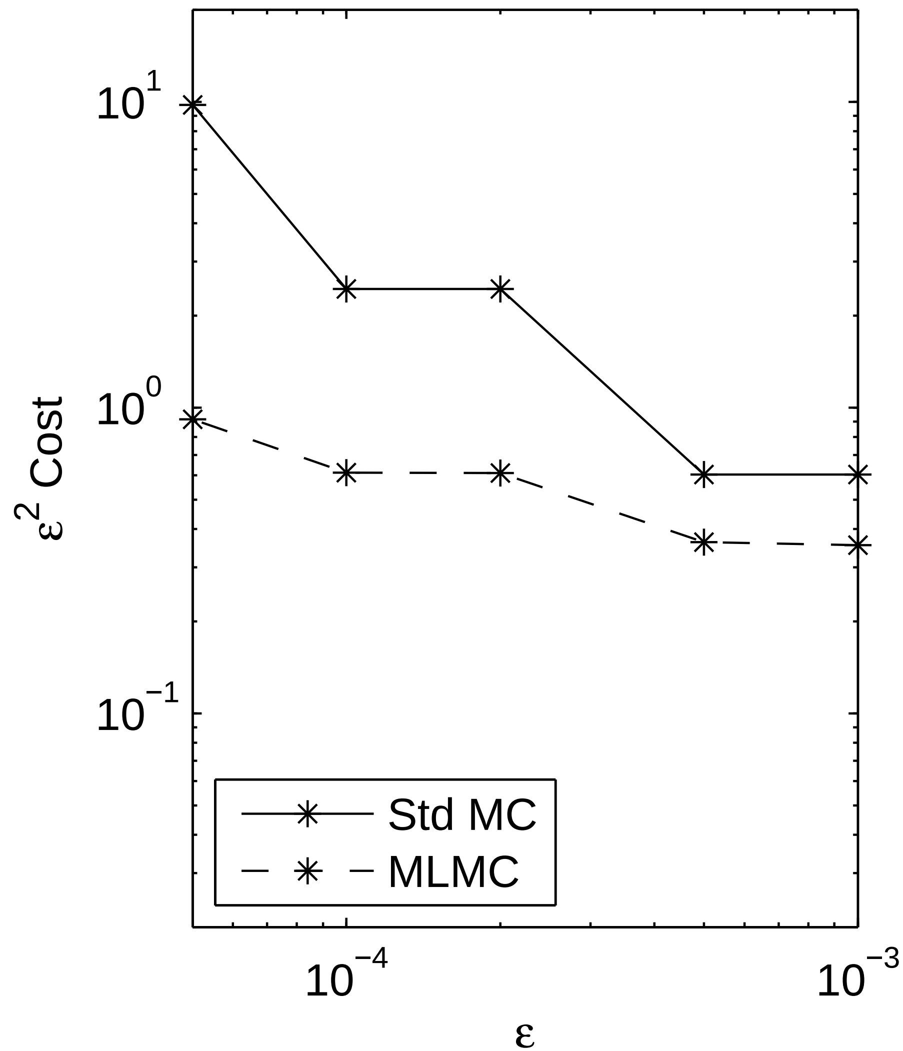
Heston model: European call



Results

Heston model: European call





Final words

Conclusions:

- improved order of complexity
- easy to implement
- significant benefits in practice

Future work:

- use of Milstein method and a control variate or antithetic variables to reduce complexity to $O(\varepsilon^{-2})$
- adaptive sampling to treat discontinuous payoffs and pathwise derivatives for Greeks
- use of quasi-Monte Carlo methods
- use of other variance reduction techniques