Multilevel Monte Carlo path simulation

Mike Giles

giles@comlab.ox.ac.uk

Oxford University Computing Laboratory

Oxford-Man Institute of Quantitative Finance

Generic Problem

Stochastic differential equation with general drift and volatility terms:

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

For simple European options, 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:

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

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

$$\widehat{Y} = N^{-1} \sum_{i=1}^{N} \widehat{P}^{(i)}$$

where $\widehat{P} \equiv f(\widehat{S}_{T/h})$ is an approximation to $P \equiv f(S(T))$ for a given Brownian path $W(t)$.

Standard MC Approach

The mean square error is defined as

$$
\begin{aligned}
\mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[P]\right)^2\right] &= \mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[\widehat{P}] \;+\; \mathbb{E}[\widehat{P}] - \mathbb{E}[P]\right)^2\right] \\
&= \mathbb{E}\left[(\widehat{Y} - \mathbb{E}[\widehat{P}])^2\right] + (\mathbb{E}[\widehat{P}] - \mathbb{E}[P])^2 \\
&= N^{-1}\mathbb{V}[\widehat{P}] + \left(\mathbb{E}[\widehat{P}] - \mathbb{E}[P]\right)^2
\end{aligned}
$$

- first term is due to variance of estimator
- second term is due to bias due to finite timestep
  – weak convergence

Standard MC Approach

Weak convergence:

- error in the expected value, $\mathbb{E}[\widehat{P}] - \mathbb{E}[P]$
- most important error in most applications
- $O(h)$ for the Euler discretisation

Strong convergence:

- error in path approximation

$$\sqrt{\mathbb{E}\left[\left\|\widehat{S}_{T/h} - S(T)\right\|^2\right]} \quad \text{or} \quad \sqrt{\mathbb{E}\left[\max_{0<t<T}\left\|\widehat{S}(t) - S(t)\right\|^2\right]}$$

- usually not relevant, but important for multilevel method
- $O(h^{1/2})$ for the Euler discretisation

Standard MC Approach

Combined mean-square-error is $O(N^{-1} + h^2)$.

To make this equal to ε^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 $\widehat{P}_l$

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

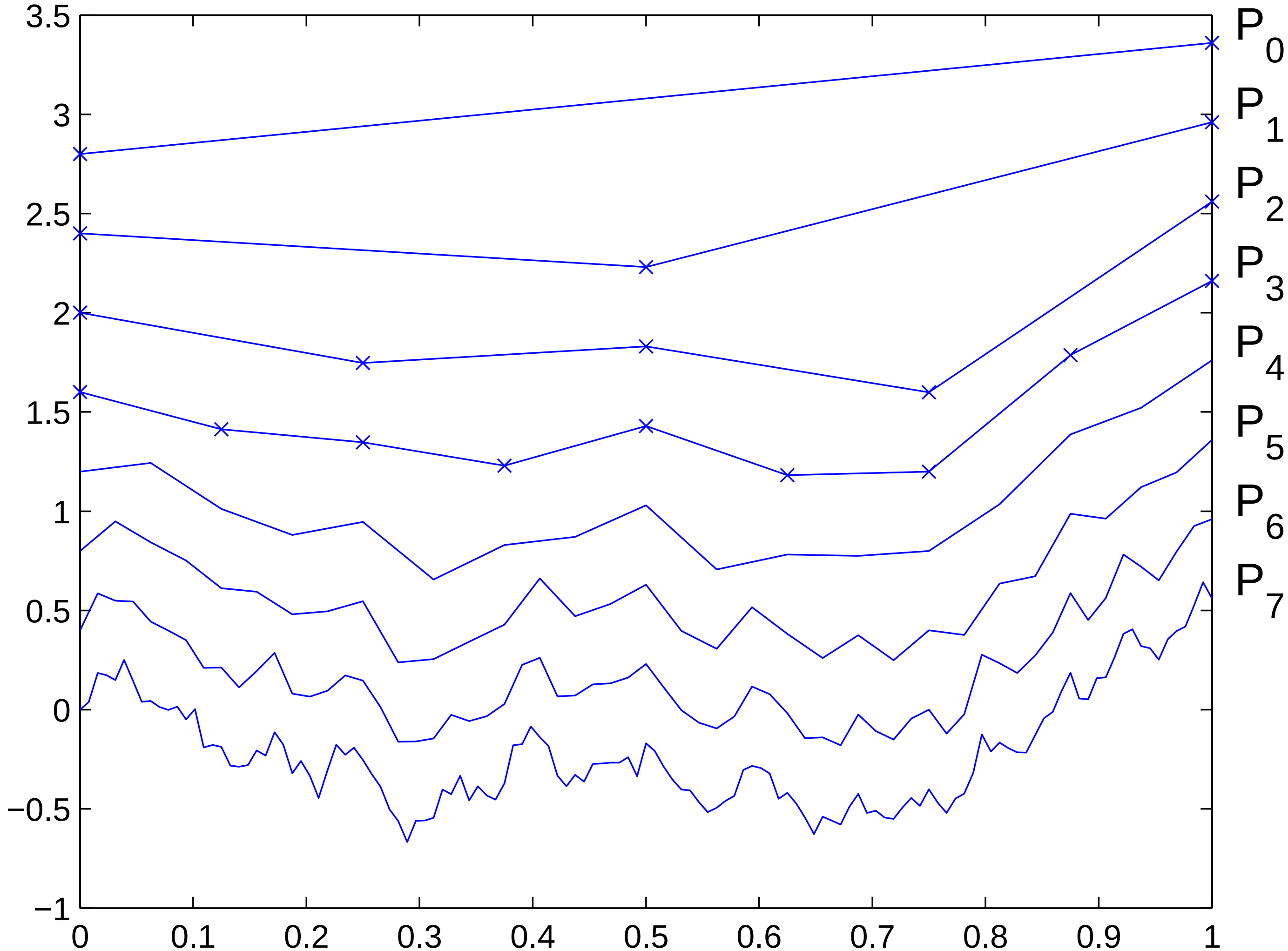
Expected value is same – aim is to reduce variance of estimator for a fixed computational cost.

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

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

$$\left|\widehat{P} - P\right| \;\leq\; c\,\left\|\widehat{S}_{T/h} - S(T)\right\| \quad\Longrightarrow\quad \mathbb{V}[\widehat{P}_l - P] = O(h_l)$$

Also, if $c = a - b$ then

$$\sqrt{\mathbb{V}[c]} \;\leq\; \sqrt{\mathbb{V}[a]} + \sqrt{\mathbb{V}[b]}$$

and so, putting

$$\widehat{P}_l - \widehat{P}_{l-1} = (\widehat{P}_l - P) - (\widehat{P}_{l-1} - P)$$

it follows that

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

Multilevel MC Approach

Hence, the optimal N_l is asymptotically proportional to h_l, and 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 ε^2 MSE for a computational cost which is $O(\varepsilon^{-2} L^2) = O(\varepsilon^{-2} (\log \varepsilon)^2)$.

Results

Geometric Brownian motion:

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

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

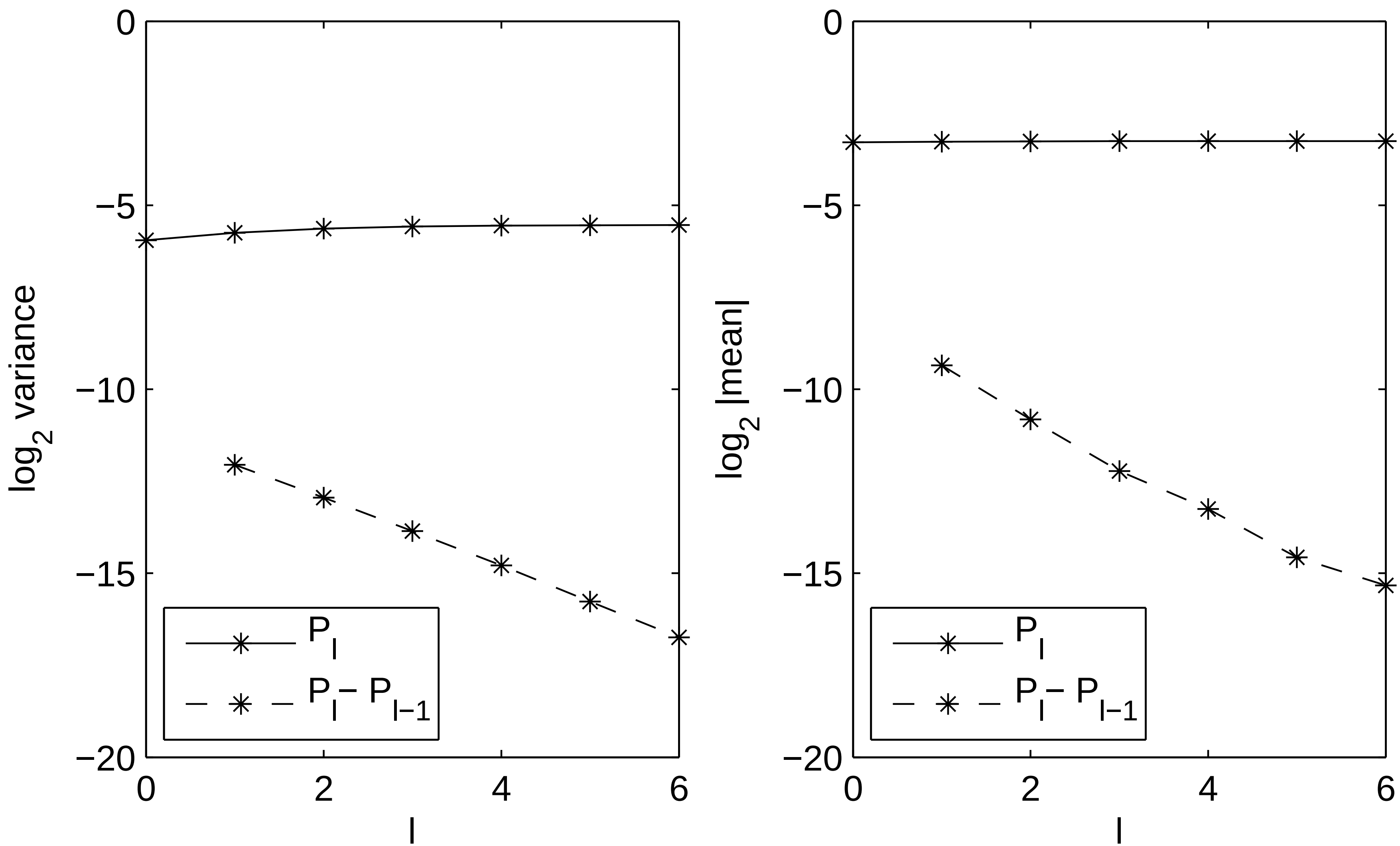
European call option with discounted payoff

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

with $K\!=\!1$.

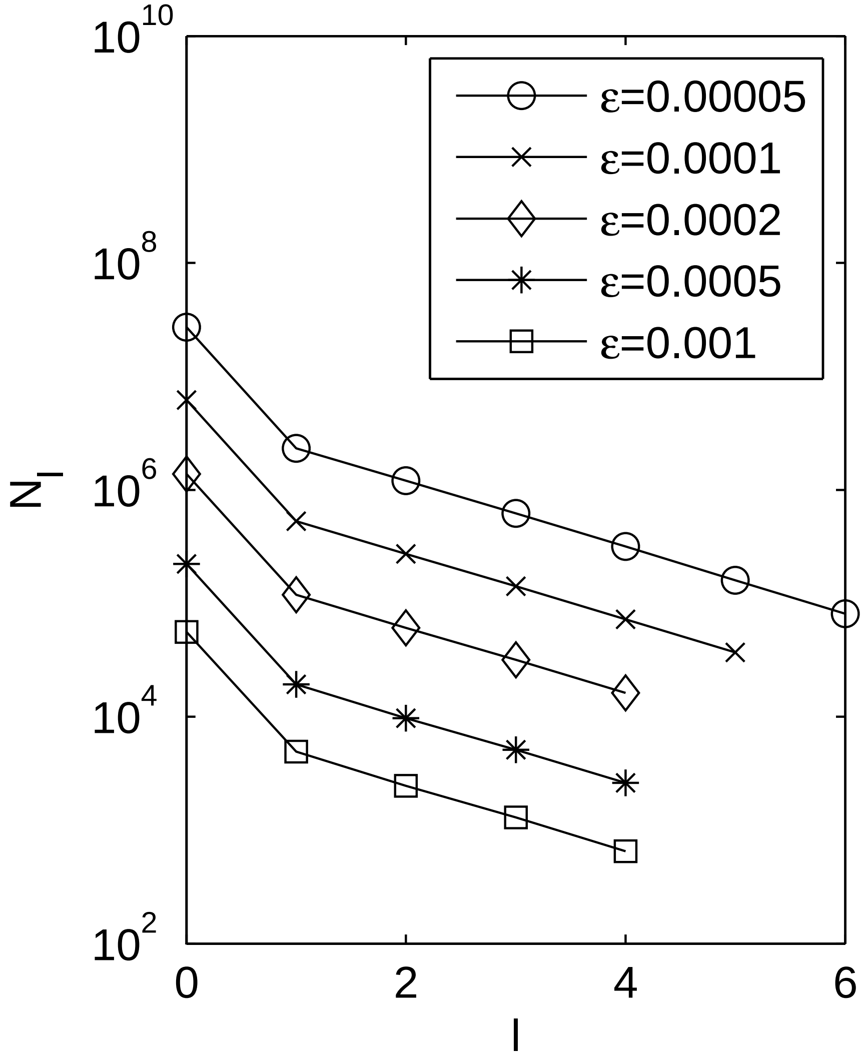
MLMC Results

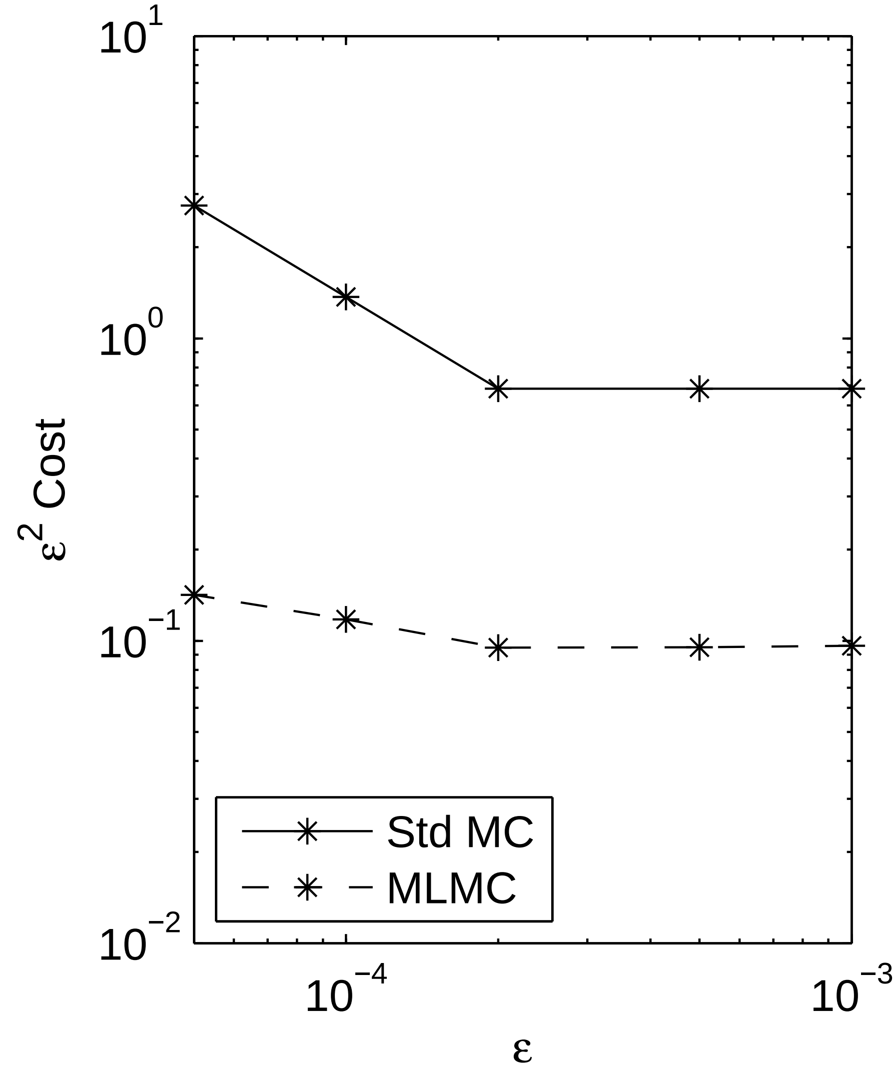
GBM: European call, $\exp(-rT)\ \max(S(T)-K, 0)$



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

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 scheme —
better strong convergence, same weak convergence

Generic scalar SDE:

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

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 or extrapolation

Milstein Scheme

Key idea: within each timestep, model the behaviour as simple Brownian motion conditional on the two end-points

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

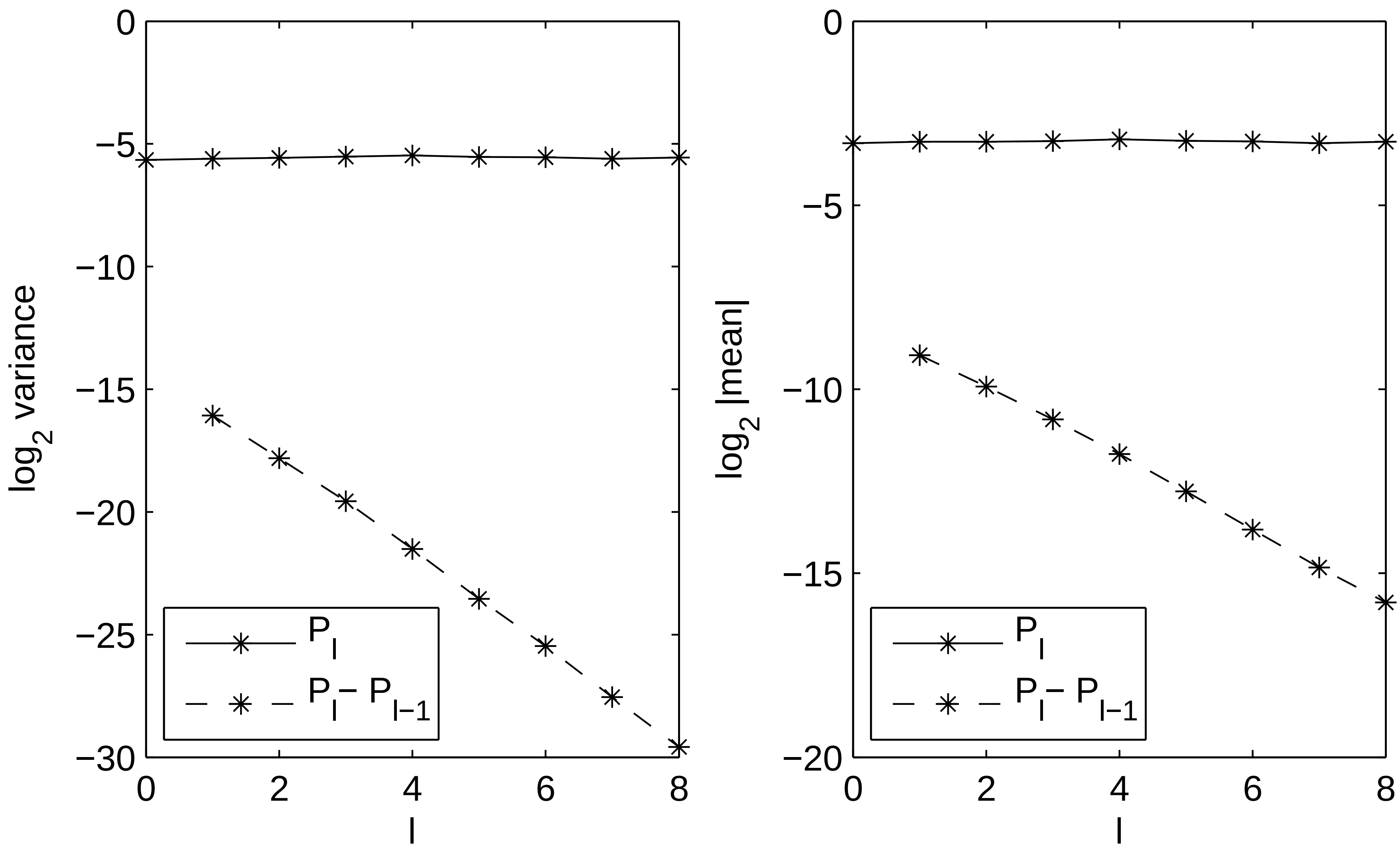
where

$$
\lambda(t) = \frac{t - t_n}{t_{n+1} - t_n}
$$

There then exist analytic results for the distribution of the min/max/average over each timstep.

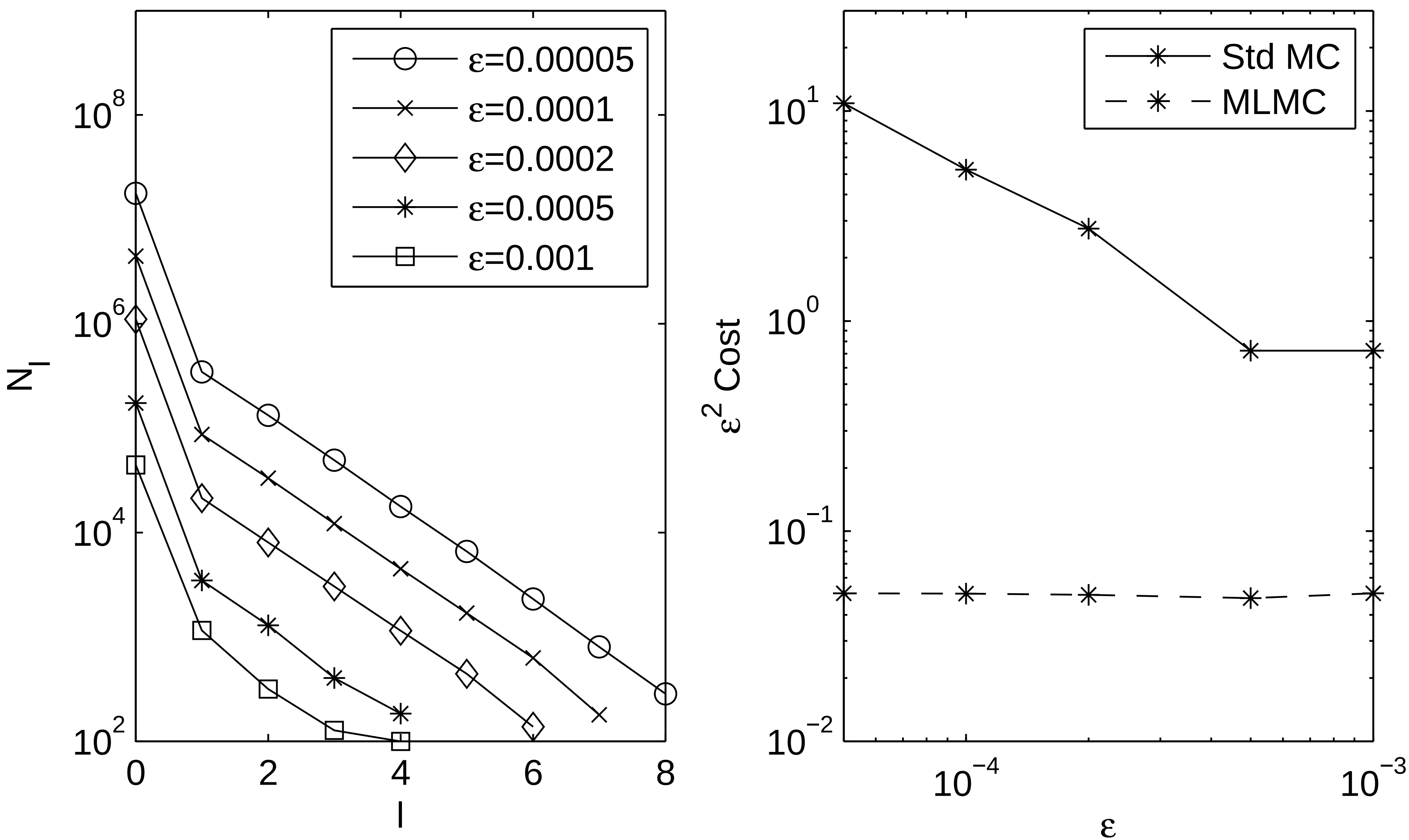
MLMC Results

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



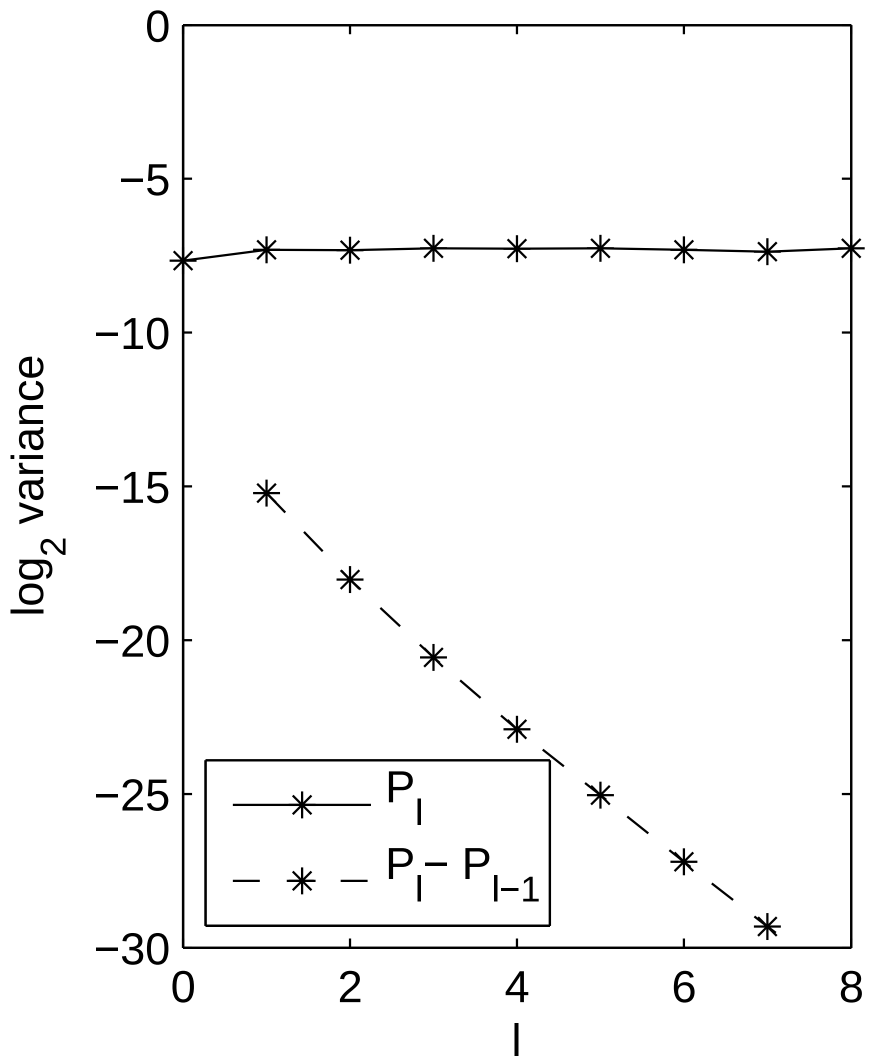
MLMC Results

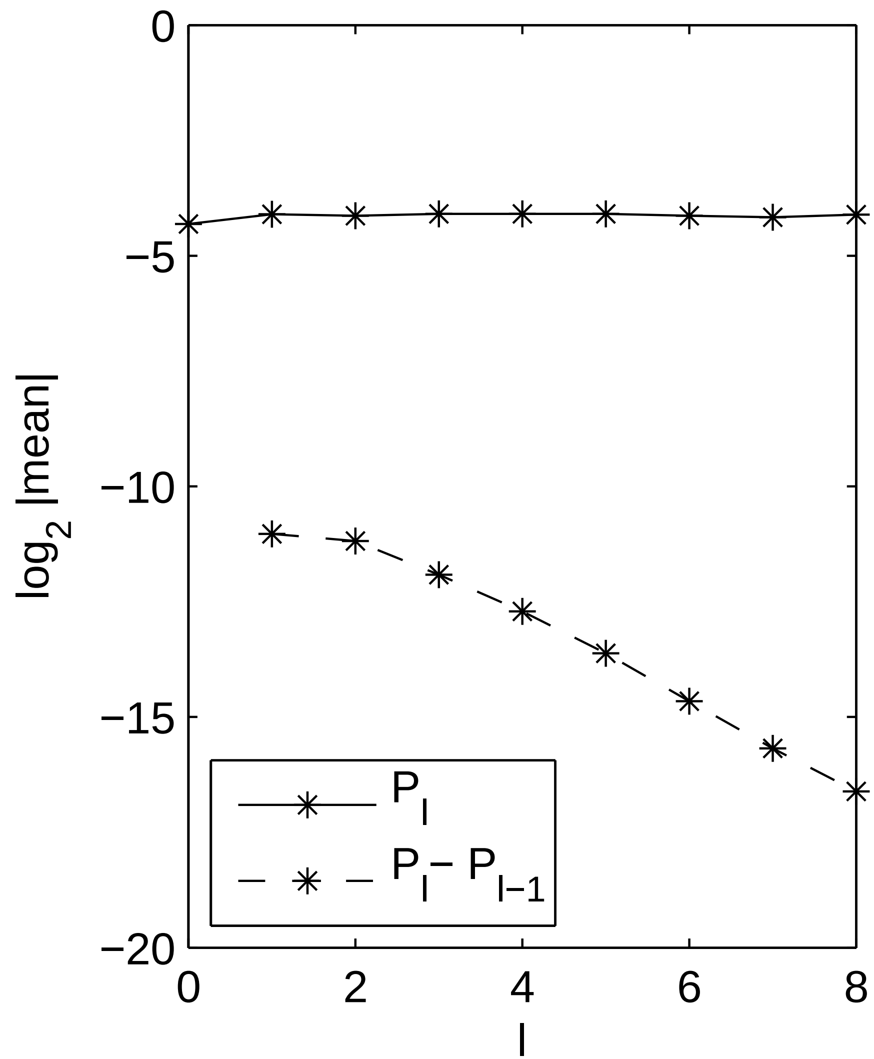
GBM: European call, $\exp(-rT)\,\max(S(T)-K,0)$



MLMC Results

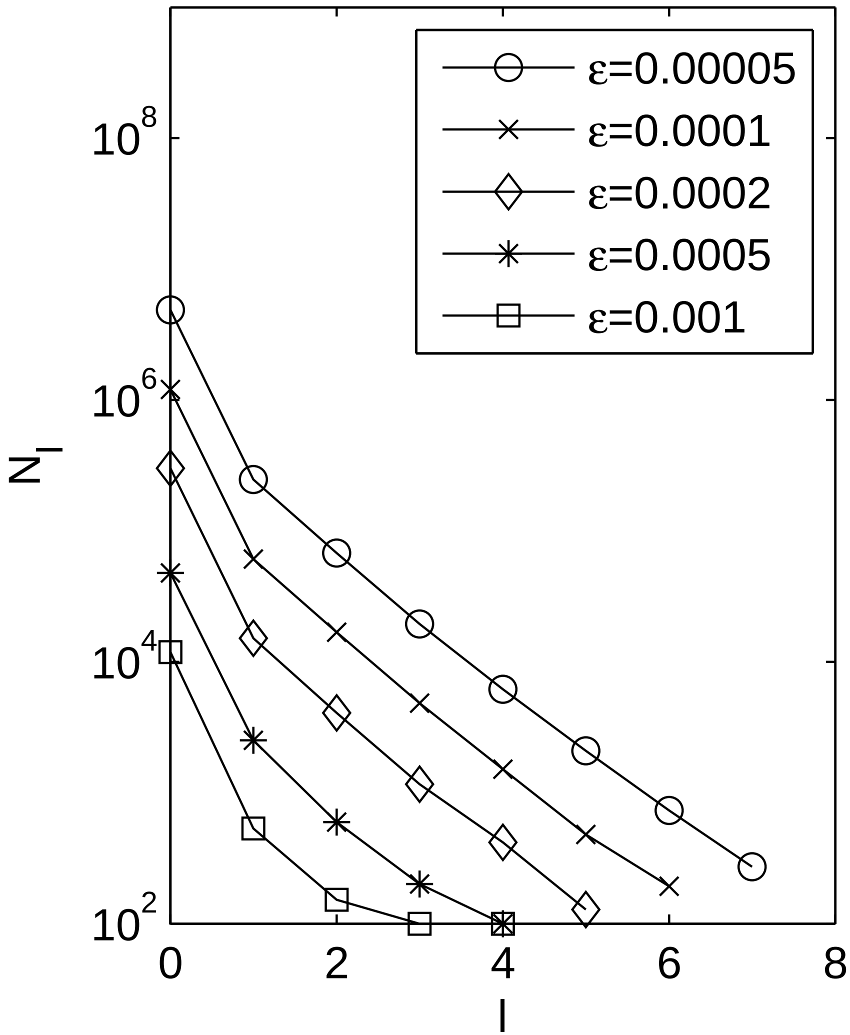
GBM: Asian option, $\exp(-rT)\ \max(T^{-1}\int_0^T S(t)\,\mathrm{d}t - 1, 0)$

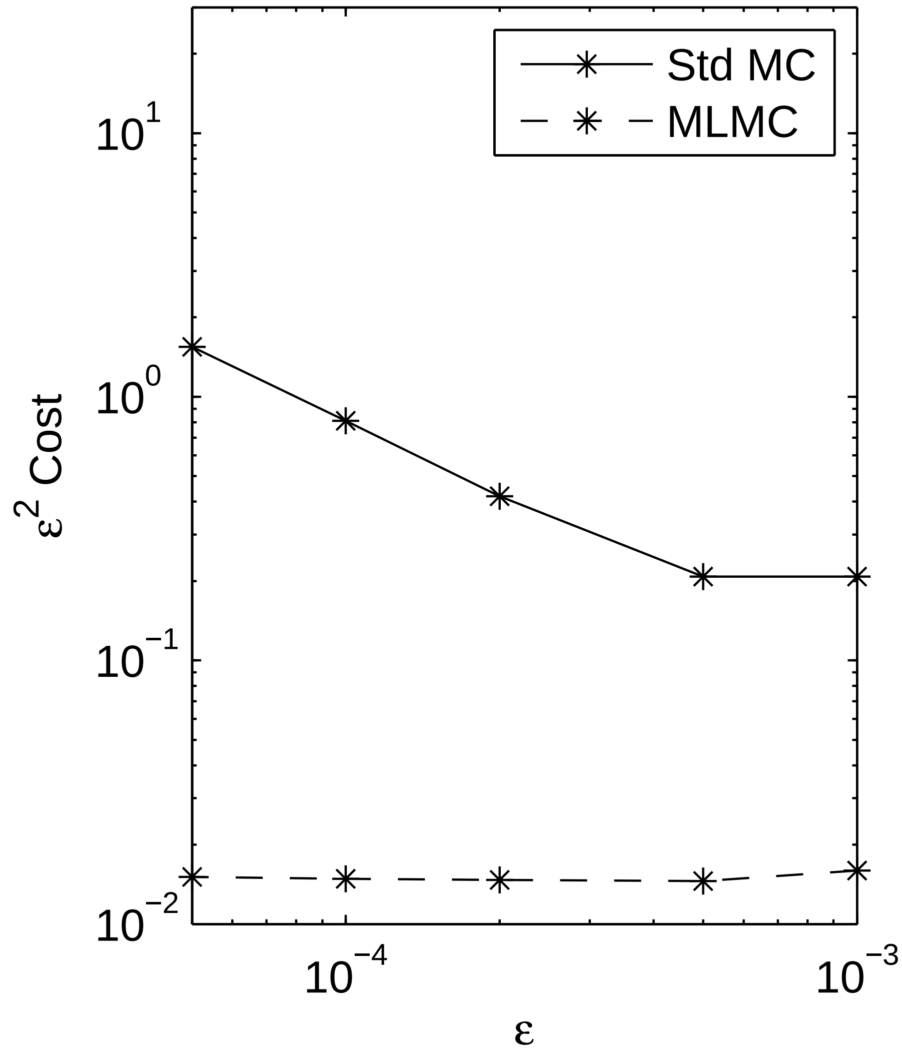




MLMC Results

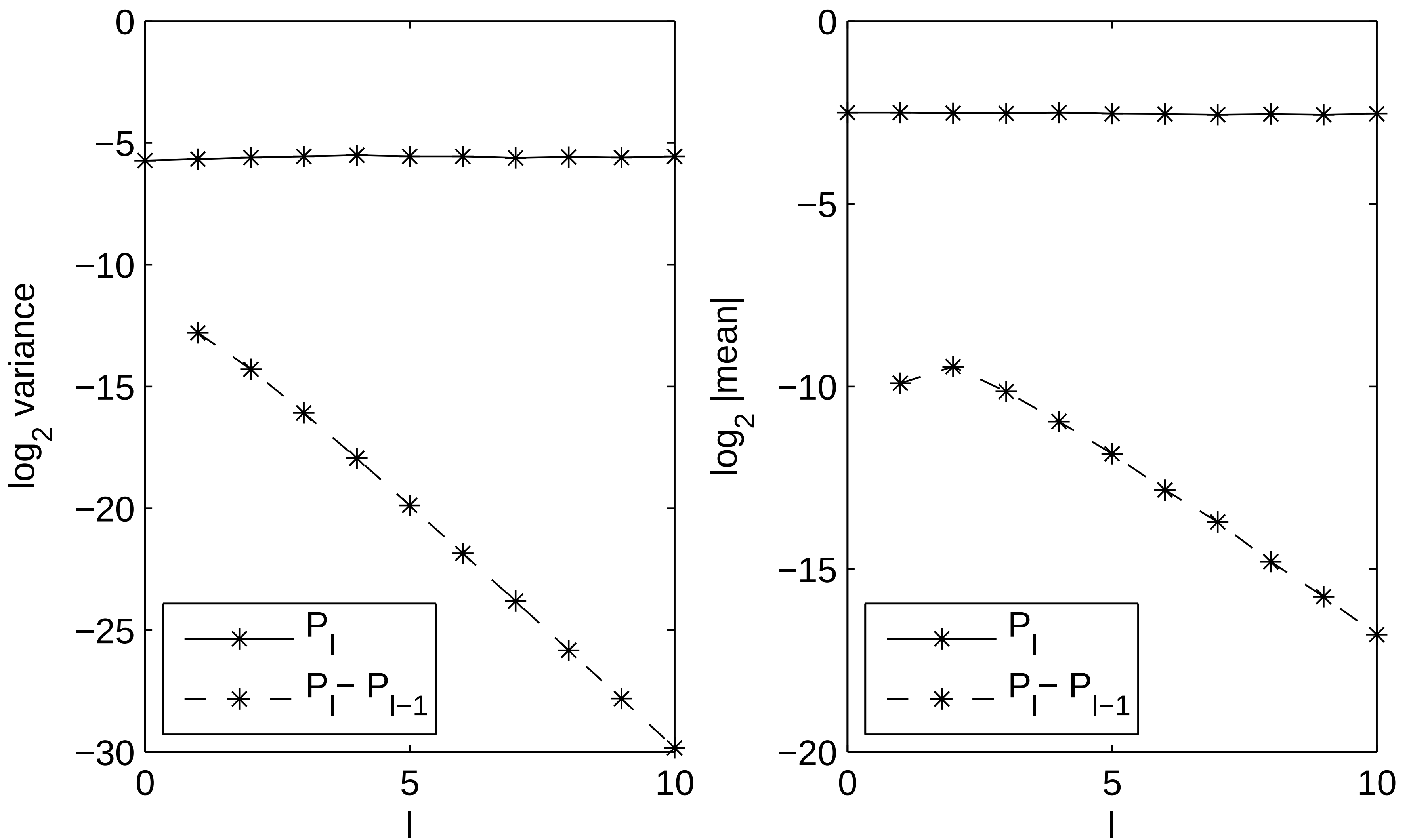
GBM: Asian option, $\exp(-rT)\ \max(T^{-1}\int_0^T S(t)\,\mathrm{d}t - 1, 0)$





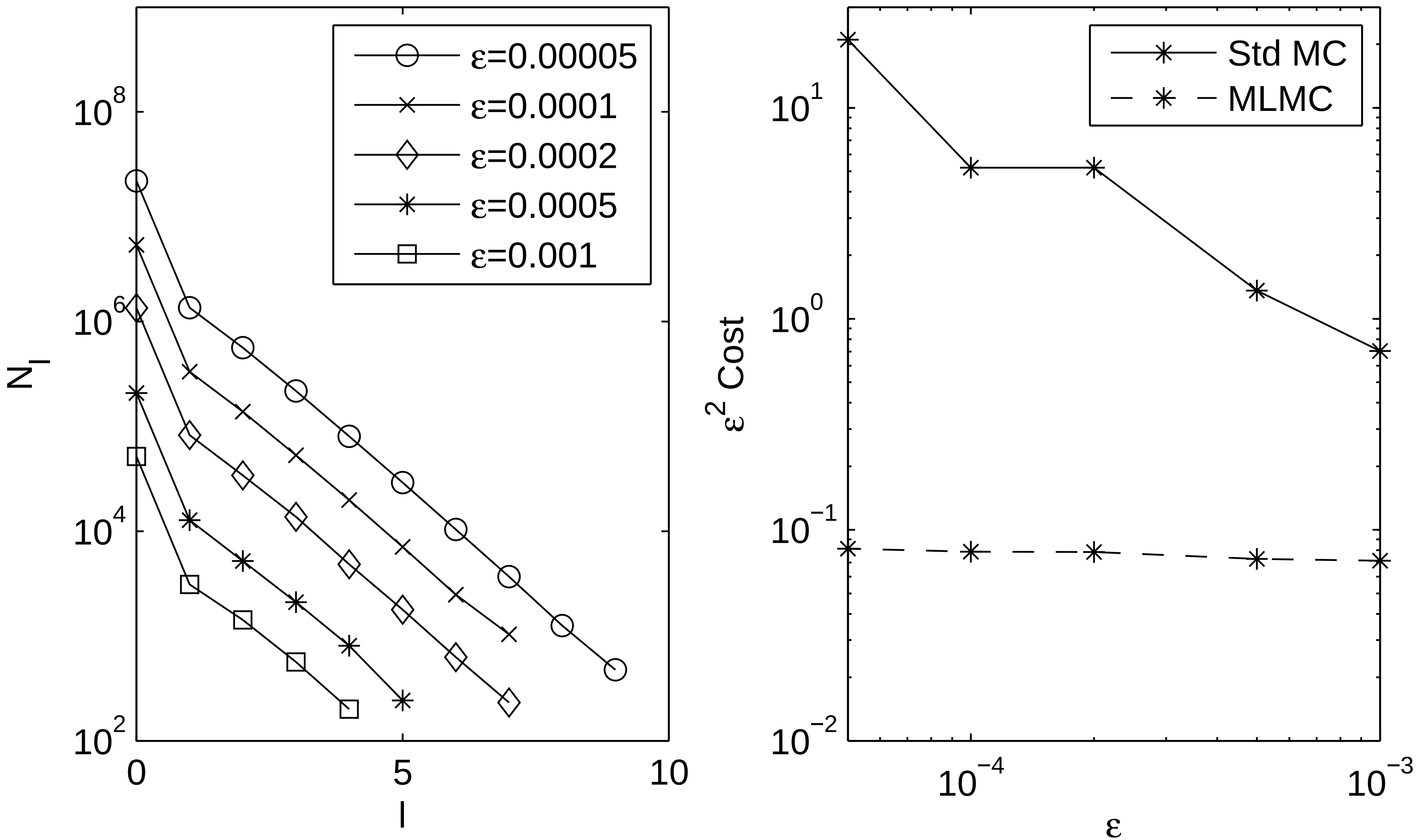
MLMC Results

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



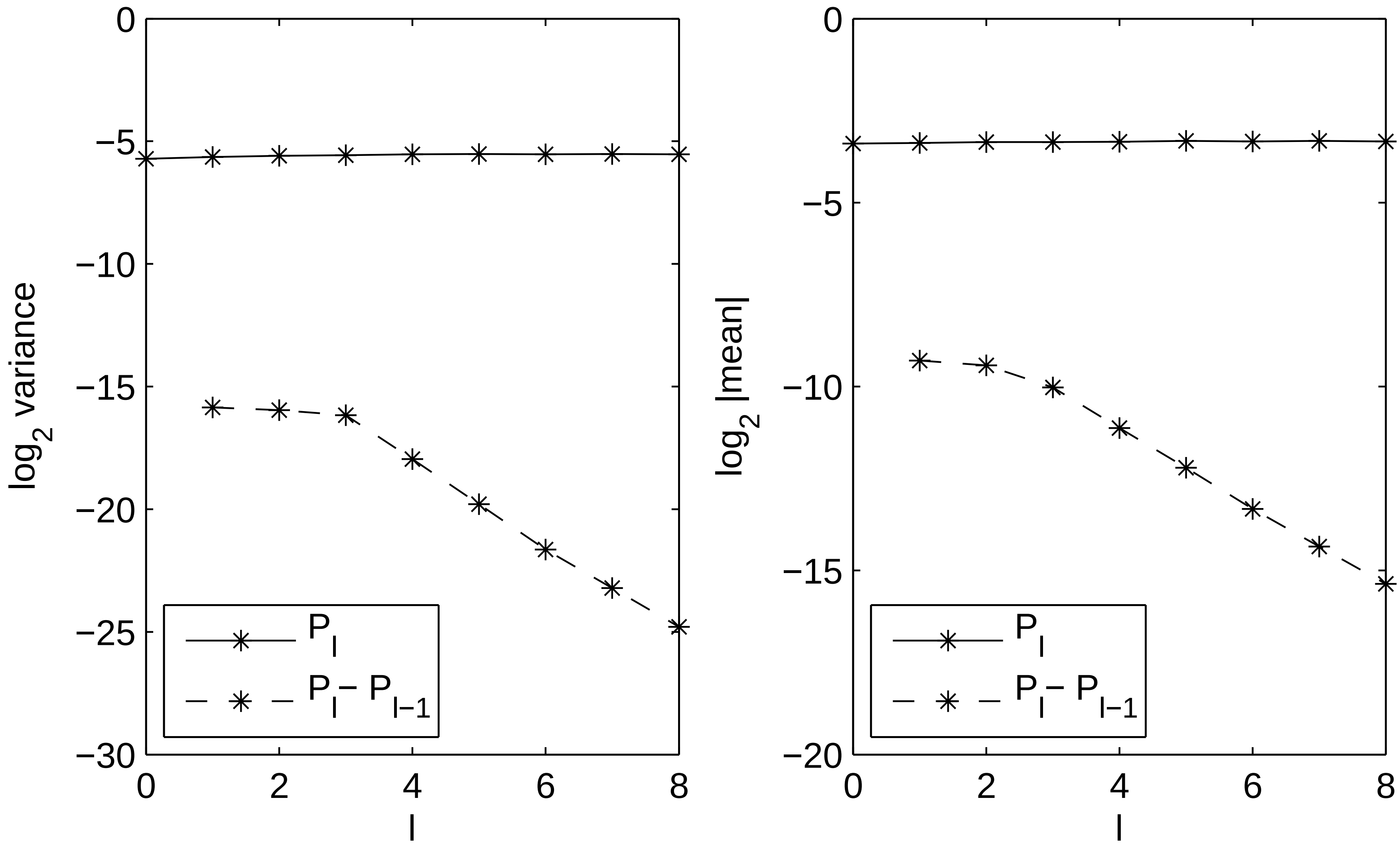
MLMC Results

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



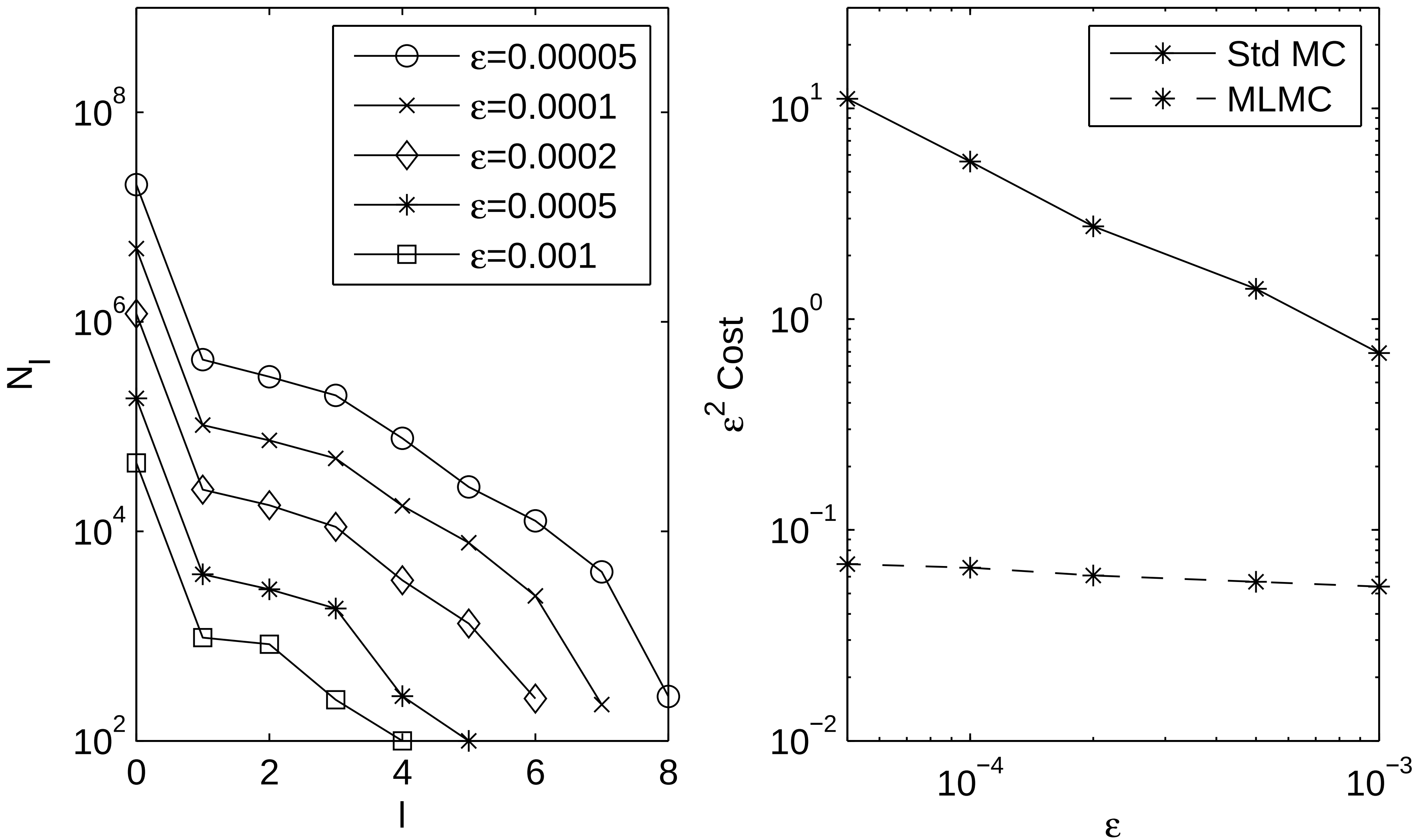
MLMC Results

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



MLMC 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

Results

Heston model:

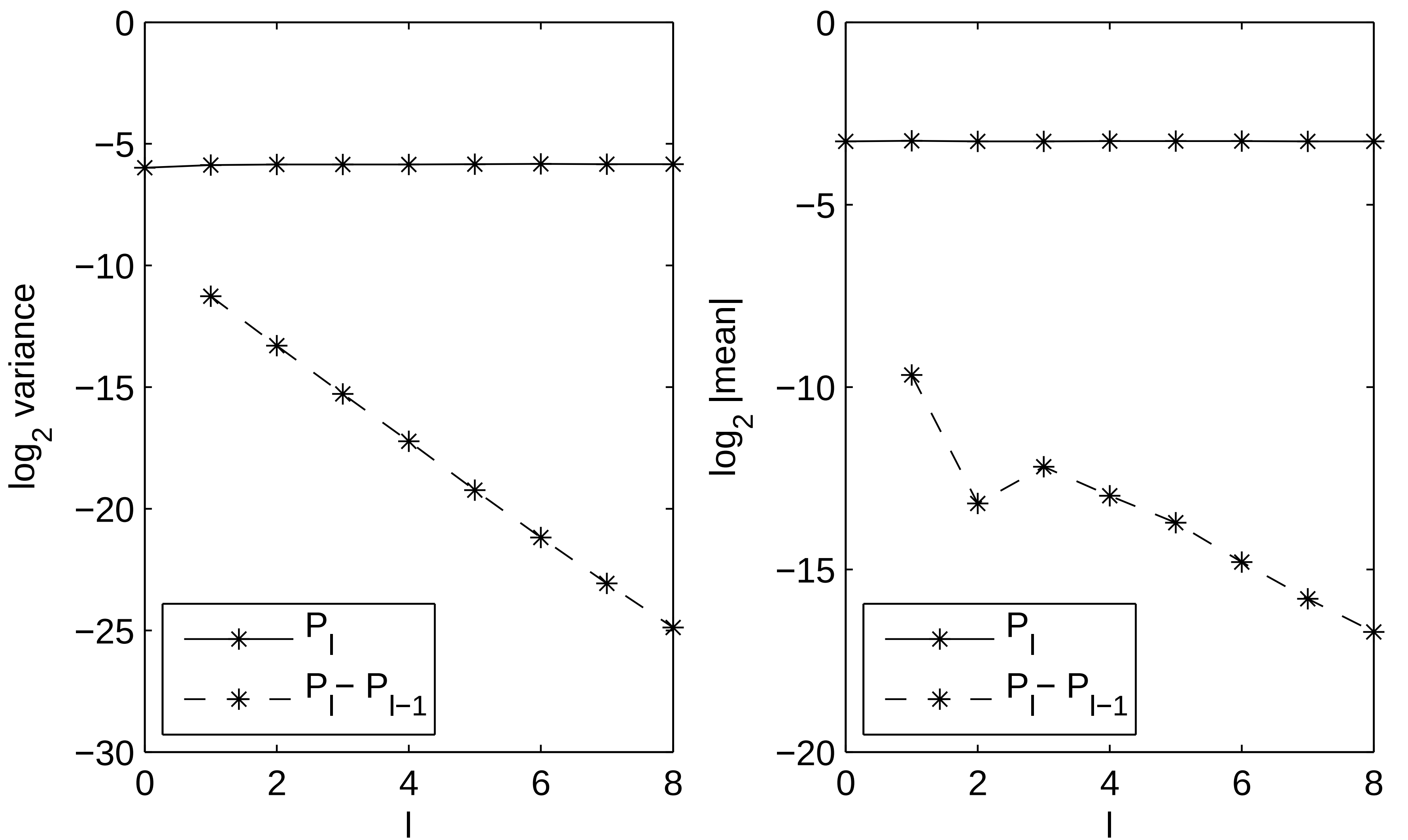
$$\begin{aligned} \mathrm{d}S &= r\,S\,\mathrm{d}t + \sqrt{V}\,S\,\mathrm{d}W_1, \qquad 0<t<T \\ \mathrm{d}V &= \lambda\,(\sigma^2 - V)\,\mathrm{d}t + \xi\,\sqrt{V}\,\mathrm{d}W_2, \end{aligned}$$

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

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

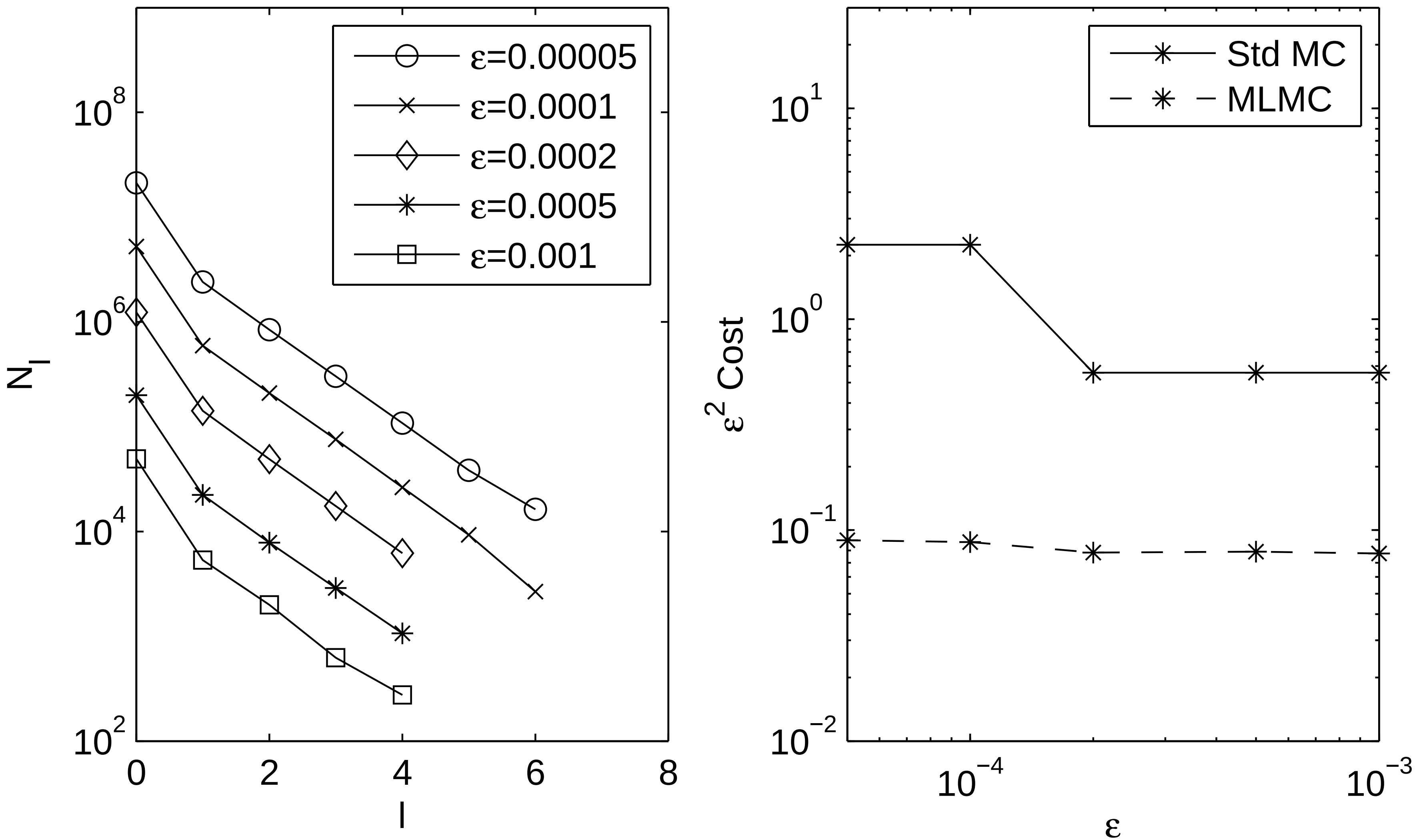
MLMC Results

Heston model: European call



MLMC Results

Heston model: European call



Quasi-Monte Carlo

- well-established technique for approximating high-dimensional integrals
- for finance applications see papers by l'Ecuyer and book by Glasserman
- Sobol sequences are perhaps most popular;
  we use lattice rules (Sloan & Kuo)
- two important ingredients for success:
  - randomized QMC for confidence intervals
  - good identification of "dominant dimensions" (Brownian Bridge and/or PCA)

Quasi-Monte Carlo

Approximate high-dimensional hypercube integral

$$\int_{[0,1]^d} f(x)\, \mathrm{d}x$$

by

$$\frac{1}{N} \sum_{i=0}^{N-1} f(x^{(i)})$$

where

$$x^{(i)} = \left[\frac{i}{N}\, z \right]$$

and z is a d-dimensional “generating vector”.

Quasi-Monte Carlo

In the best cases, error is $O(N^{-1})$ instead of $O(N^{-1/2})$ but without a confidence interval.

To get a confidence interval, let

$$x^{(i)} = \left[\frac{i}{N}\, z \; + \; x_0 \right].$$

where x_0 is a random offset vector.

Using 32 different random offsets gives a confidence interval in the usual way.

Quasi-Monte Carlo

For the path discretisation we can use

$$\Delta W_n = \sqrt{h}\, \Phi^{-1}(x_n),$$

where Φ^{-1} is the inverse cumulative Normal distribution.

Much better to use a Brownian Bridge construction:

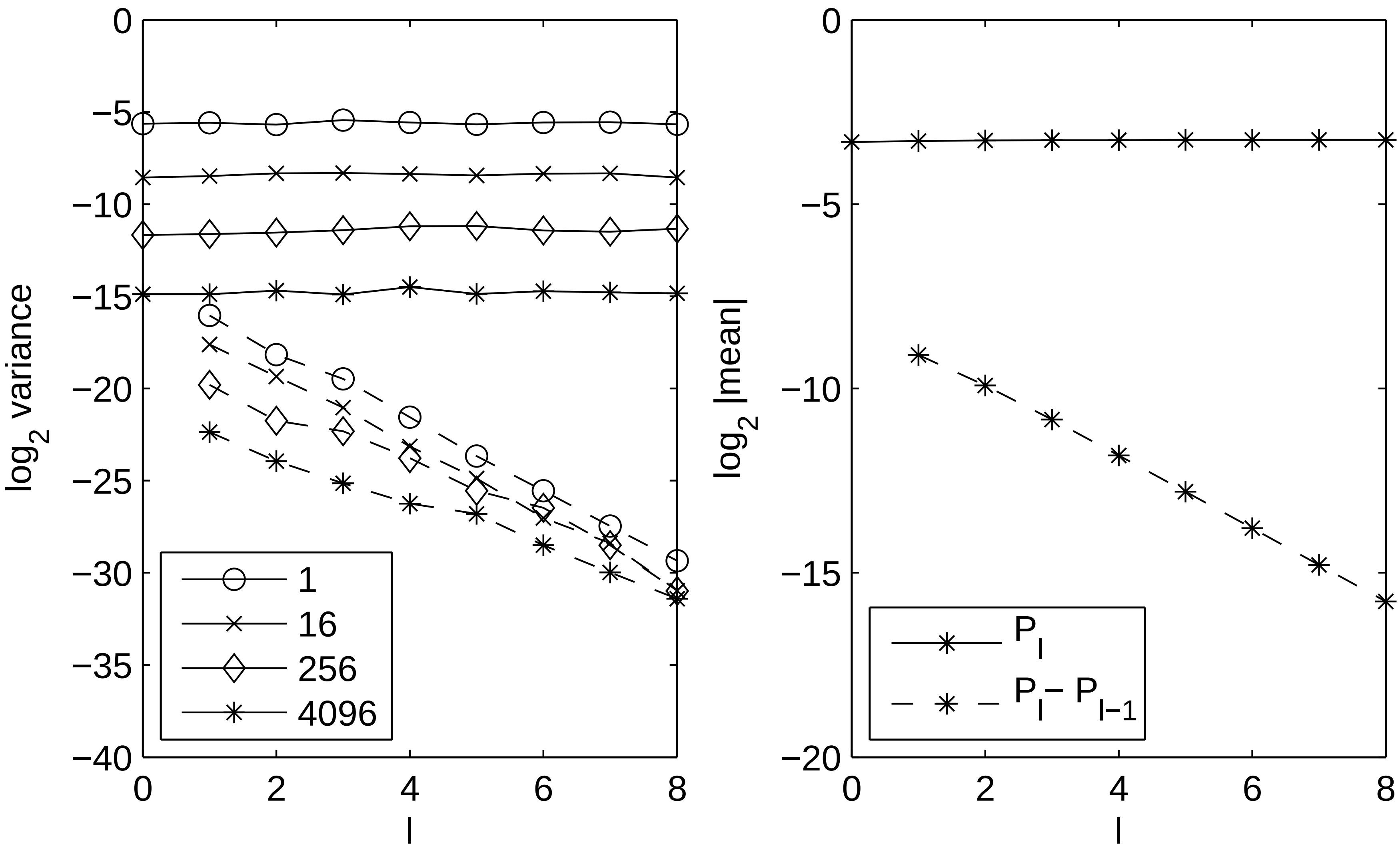
- $x_1 \quad \longrightarrow \quad W(T)$
- $x_2 \quad \longrightarrow \quad W(T/2)$
- $x_3, x_4 \quad \longrightarrow \quad W(T/4), W(3T/4)$
- $\ldots$ and so on by recursive bisection

Multilevel QMC

- rank-1 lattice rule developed by Sloan, Kuo & Waterhouse at UNSW
- 32 randomly-shifted sets of QMC points
- number of points in each set increased as needed to achieved desired accuracy, based on confidence interval estimate
- results show QMC to be particularly effective on lowest levels with low dimensionality

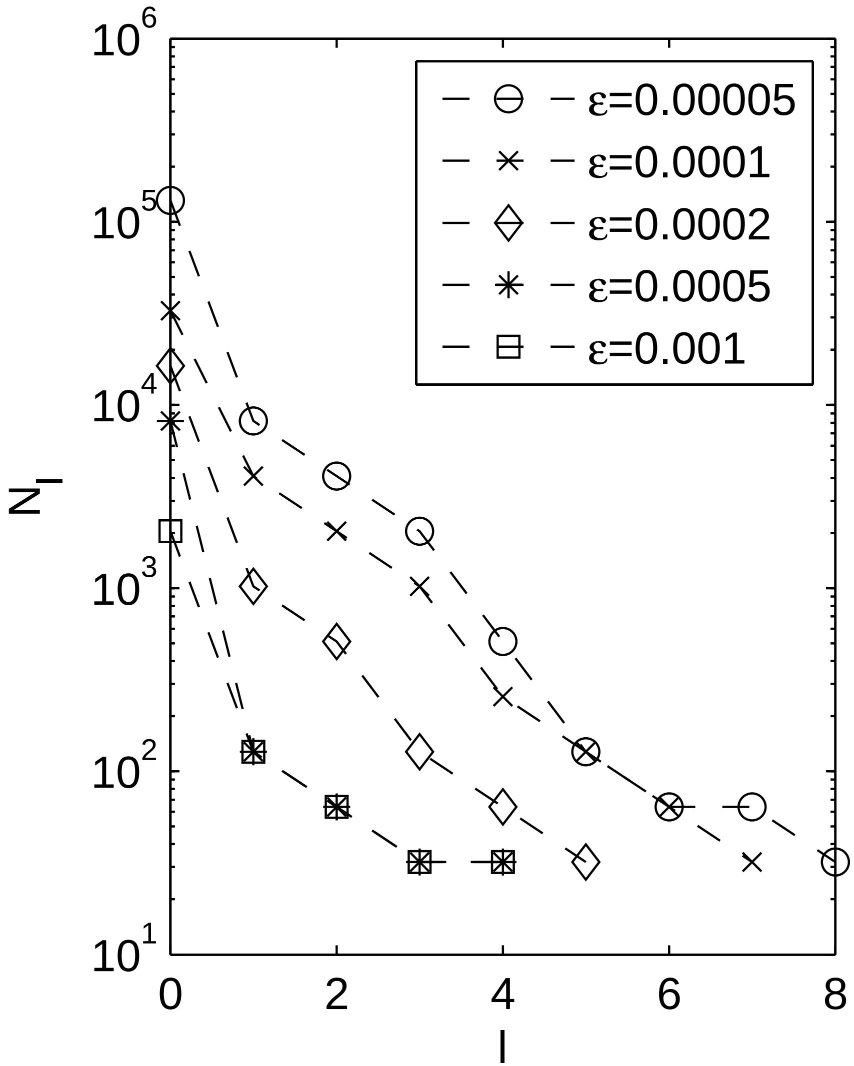
MLQMC Results

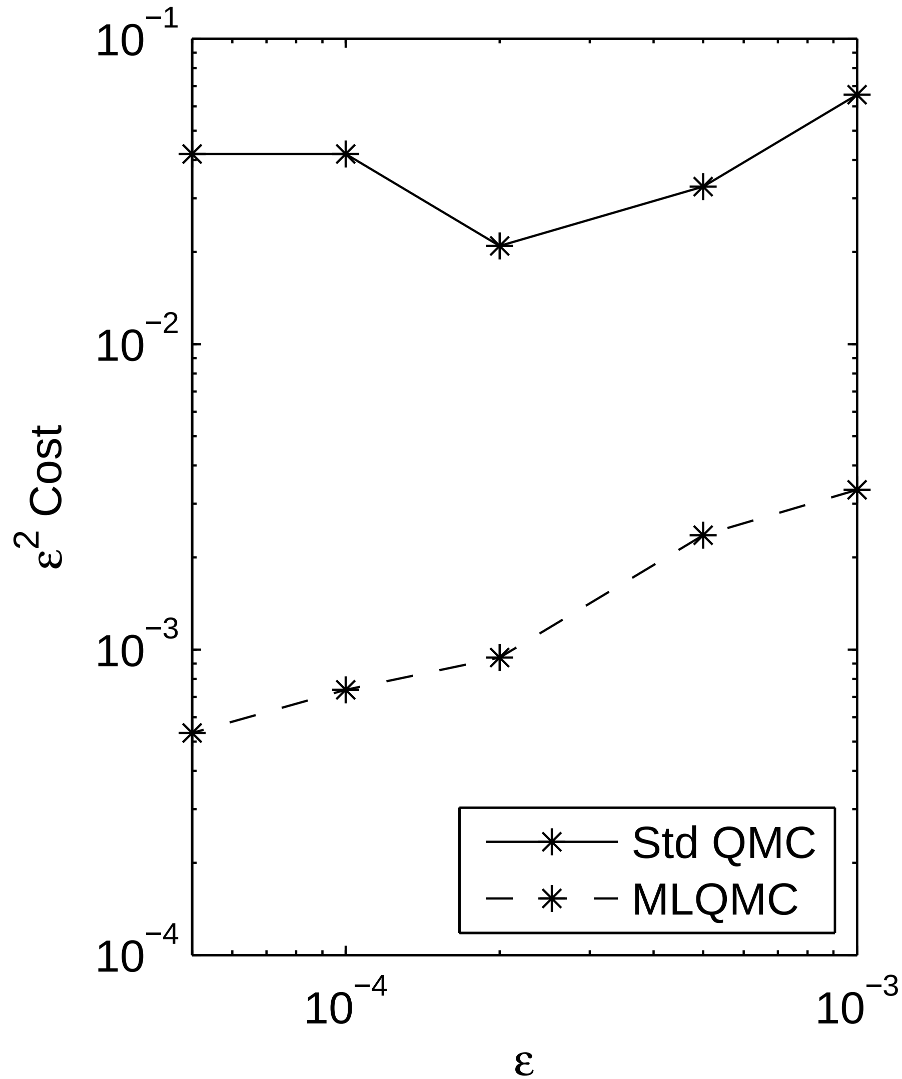
GBM: European call



MLQMC Results

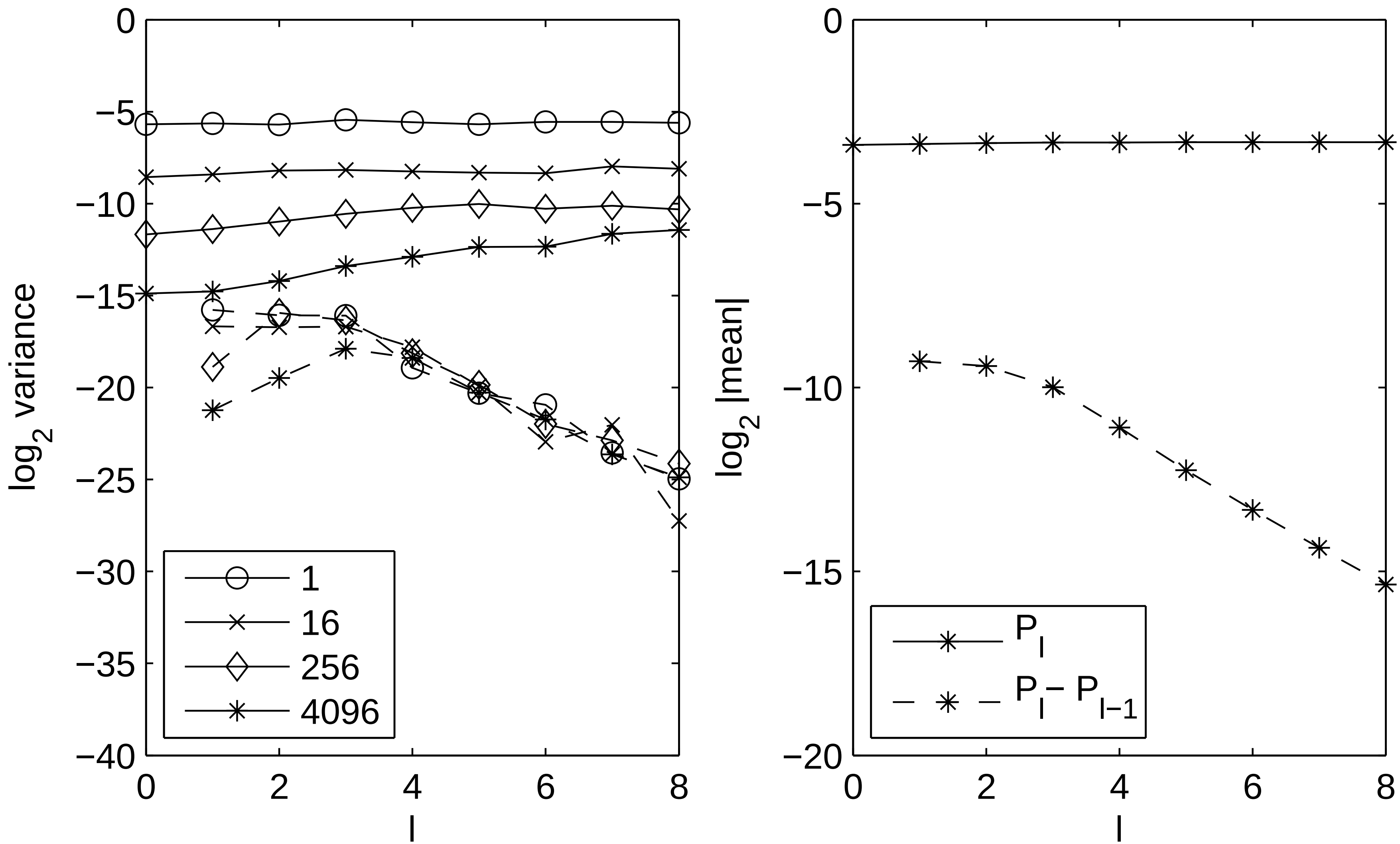
GBM: European call





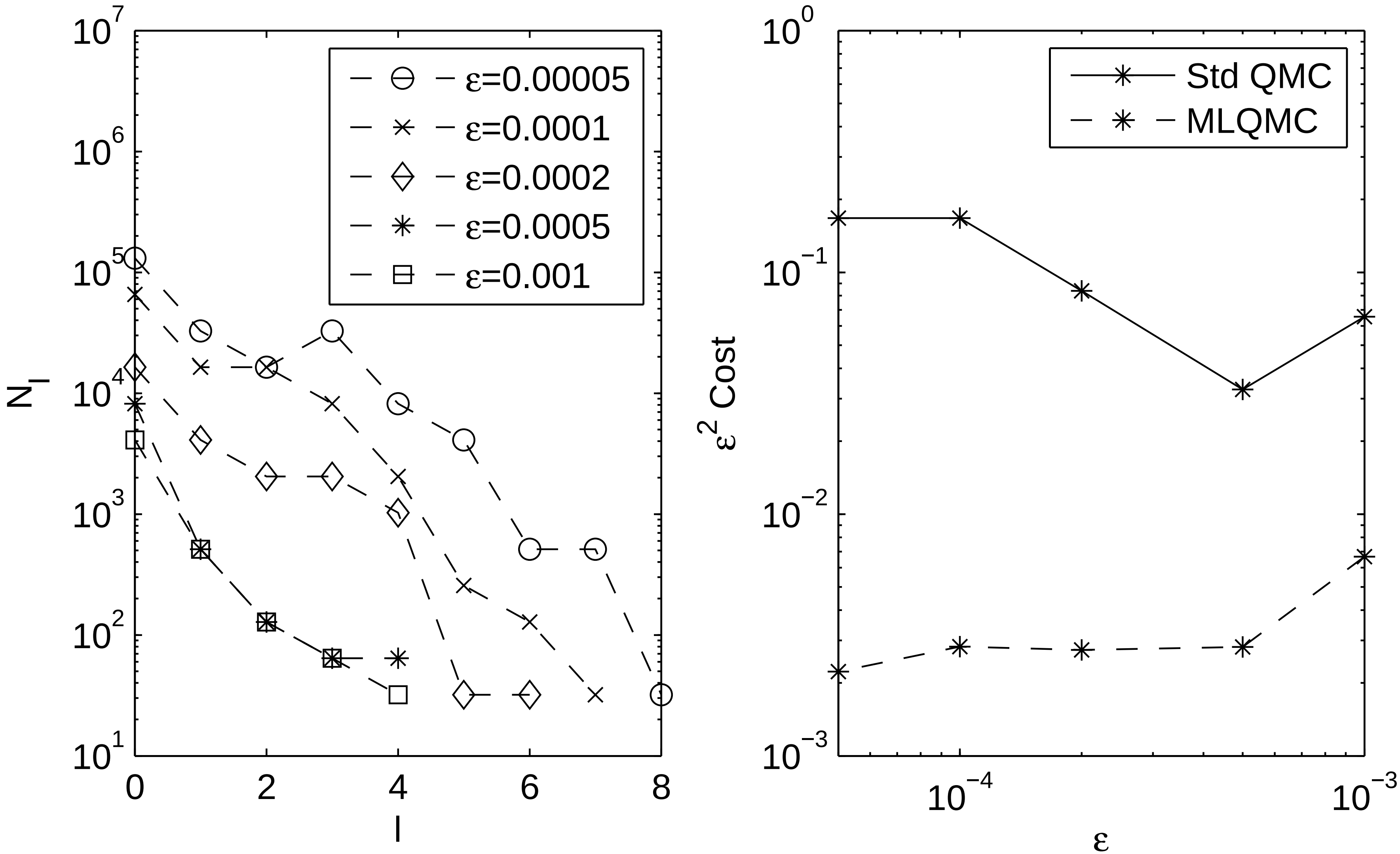
MLQMC Results

GBM: barrier option



MLQMC Results

GBM: barrier option



Conclusions

Results so far:

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

However:

- scope for further improvement
- need to test ideas on real finance applications

Current/Future Work

- numerical analysis
  (D. Higham, X. Mao – Strathclyde)
- Greeks for hedging and risk management
  (P. Glasserman – Columbia)
- parallel implementation on NVIDIA graphics cards
  (96 – 128 cores)
- real-world finance applications

Papers

M.B. Giles, “Multilevel Monte Carlo path simulation”, to appear in Operations Research, 2007

M.B. Giles, “Improved multilevel convergence using the Milstein scheme”, to appear in MCQMC06 proceedings, Springer-Verlag, 2007.

M.B. Giles, “Multilevel quasi-Monte Carlo path simulation”, submitted to Journal of Computational Finance, 2007.

Acknowledgements:

- Paul Glasserman and Mark Broadie for early feedback
- Frances Kuo, Ian Sloan and Ben Waterhouse (UNSW) for collaboration on quasi-Monte Carlo integration