Numerical Methods II

Prof. Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Multilevel Monte Carlo

- new approach to achieving greater accuracy for the same computational cost
- builds on the elements we've already learned
- incorporates ideas from the numerical solution of PDEs
- an illustration of the fact that this subject is not mature – there's still plenty of scope for improvement on existing methods

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,

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

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:

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

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 $h_\ell = 2^{-\ell}\, T,\ \ell = 0, 1, \ldots, L$, and payoff $\widehat{P}_\ell$

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

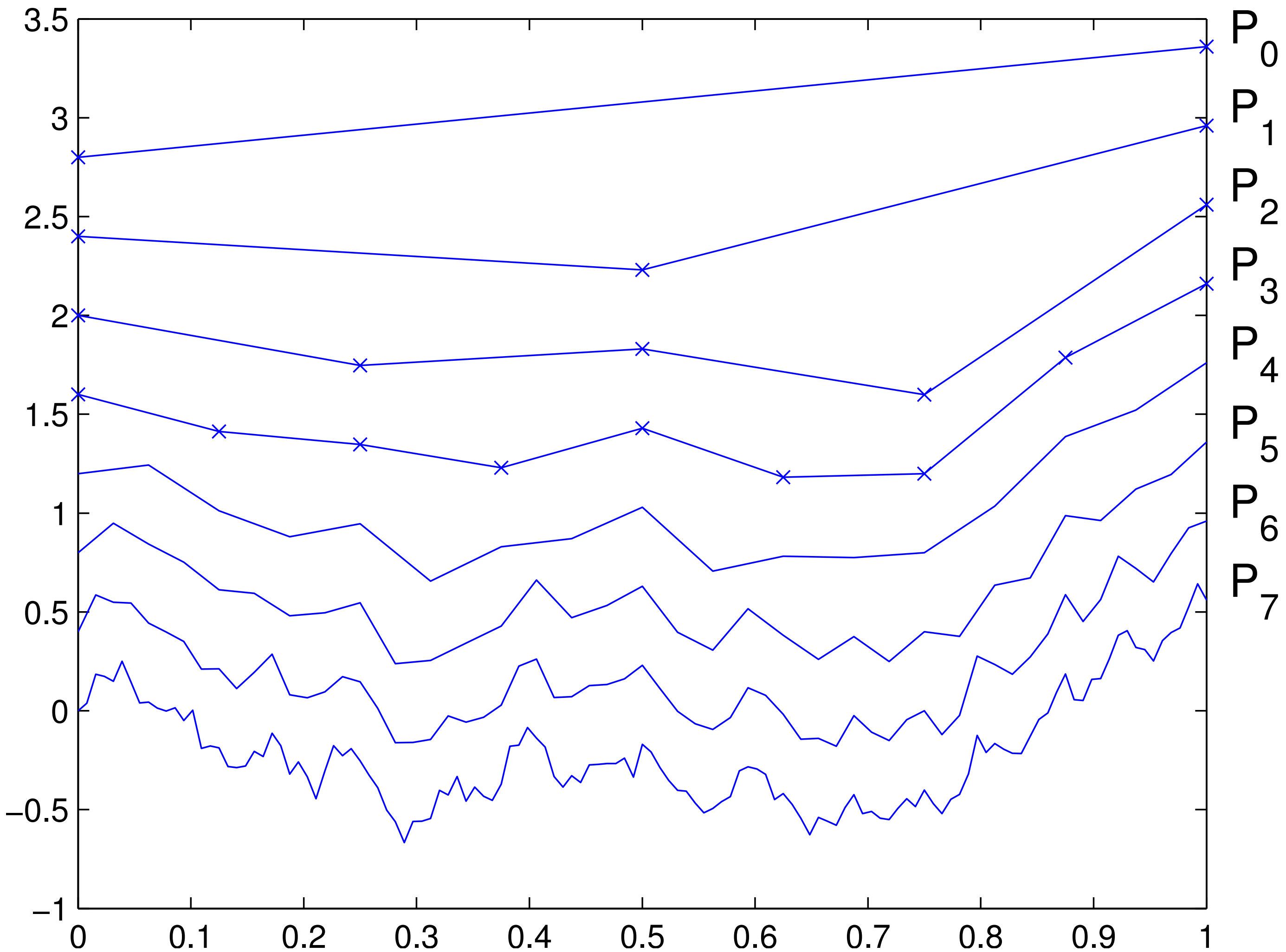
Expected value is same – aim is to reduce variance of estimator for a fixed computational cost.

Key point: approximate $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ using N_ℓ simulations with $\widehat{P}_\ell$ and $\widehat{P}_{\ell-1}$ obtained using <u>same</u> Brownian path.

$$\widehat{Y}_\ell = N_\ell^{-1} \sum_{i=1}^{N_\ell} \left(\widehat{P}_\ell^{(i)} - \widehat{P}_{\ell-1}^{(i)} \right)$$

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

$$\mathbb{V}\left[\sum_{\ell=0}^{L} \widehat{Y}_\ell\right] = \sum_{\ell=0}^{L} N_\ell^{-1} V_\ell, \qquad V_\ell \equiv \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}],$$

and the computational cost is proportional to $\displaystyle\sum_{\ell=0}^{L} N_\ell \, h_\ell^{-1}$.

Hence, by using a Lagrange multiplier, the computational cost is minimised for a fixed variance by choosing N_ℓ to be proportional to $\sqrt{V_\ell \, h_\ell}$.

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

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

Also, if $c = a - b$ then

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

and so, putting

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

it follows that

$$\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = O(h_\ell)$$

Multilevel MC Approach

Hence, the optimal N_ℓ is asymptotically proportional to h_ℓ, and to make the combined variance $O(\varepsilon^2)$ requires

$$N_\ell = O(\varepsilon^{-2} L\, h_\ell).$$

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

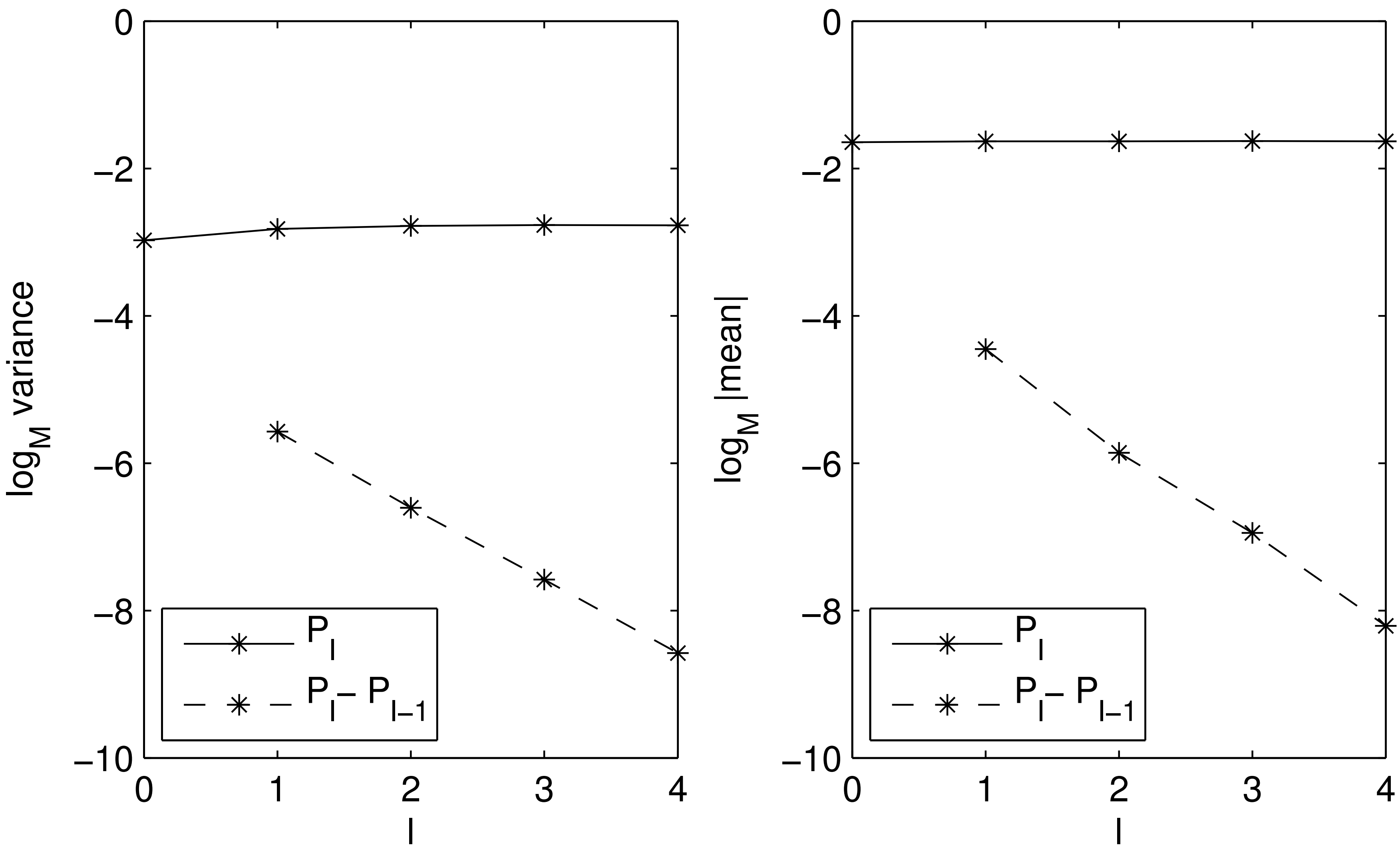
$$\exp(-rT)\;\max(S(T)\!-\!K,0)$$

with $K\!=\!1$.

Numerical calculations use factor 4 incease in number of timesteps at each level.

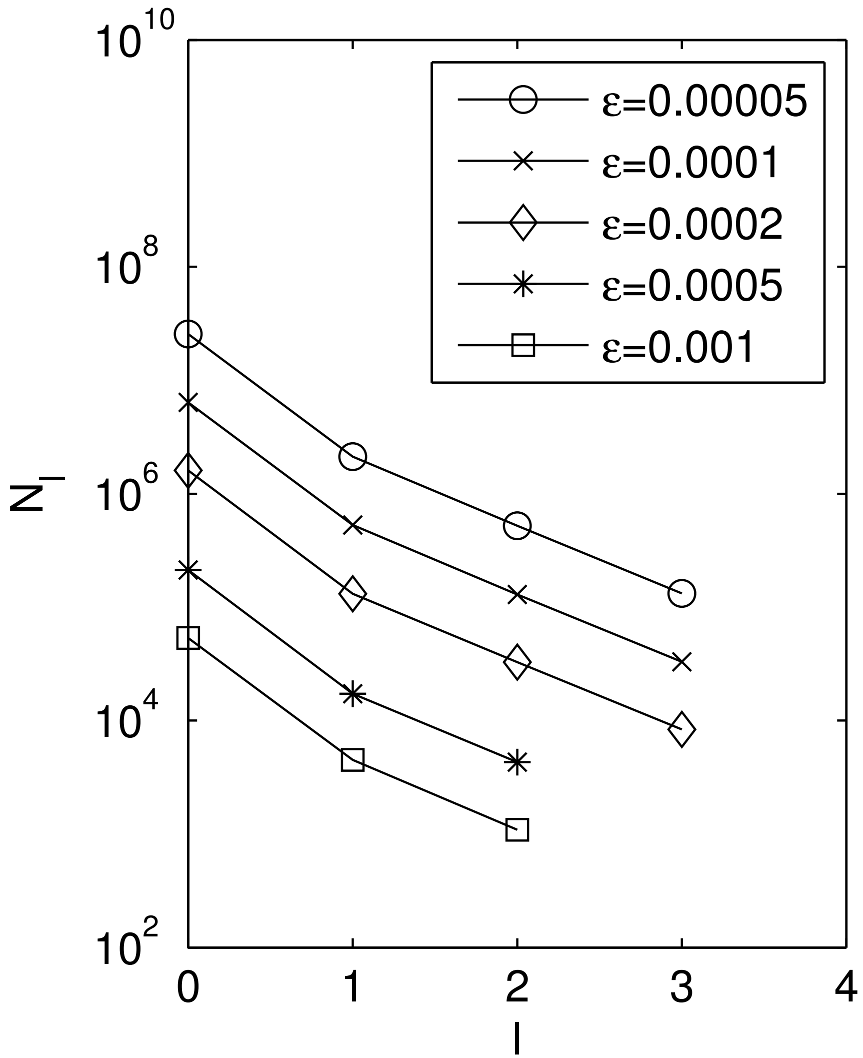
MLMC Results

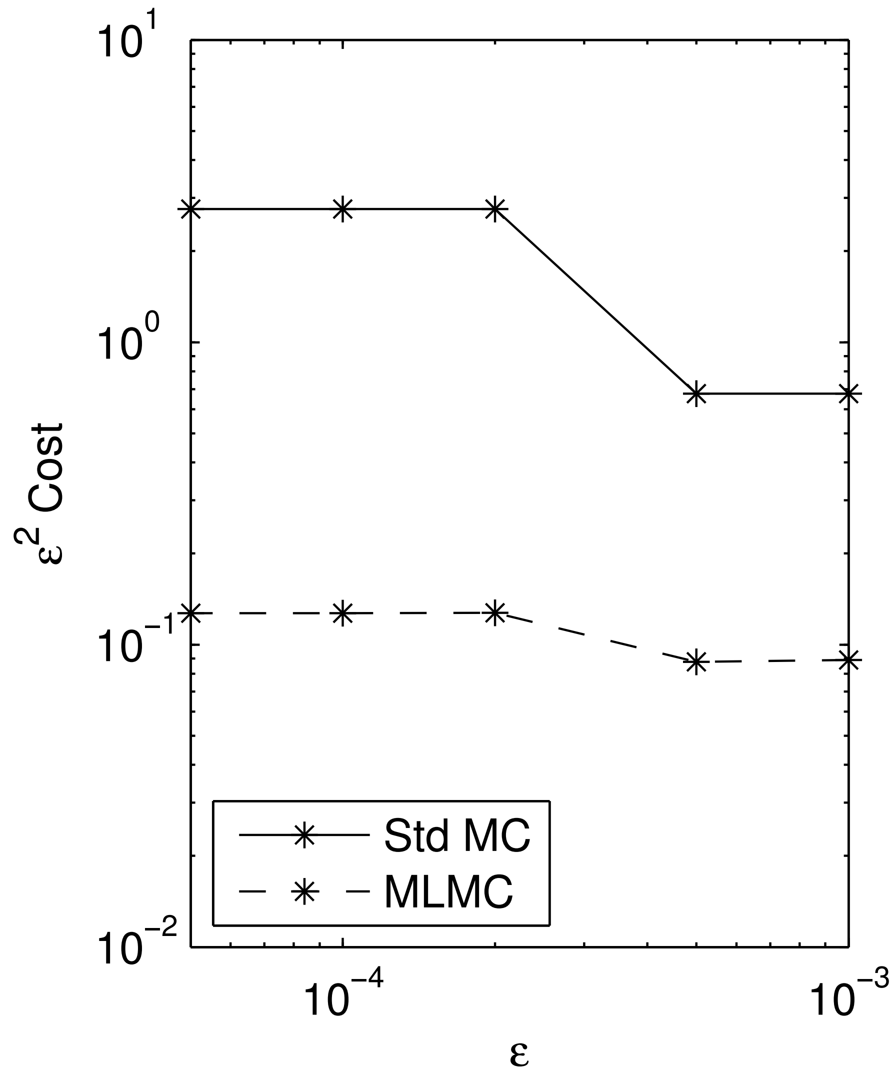
GBM: European call, $\exp(-rT)\ \max(S(T)-K, 0)$



MLMC Results

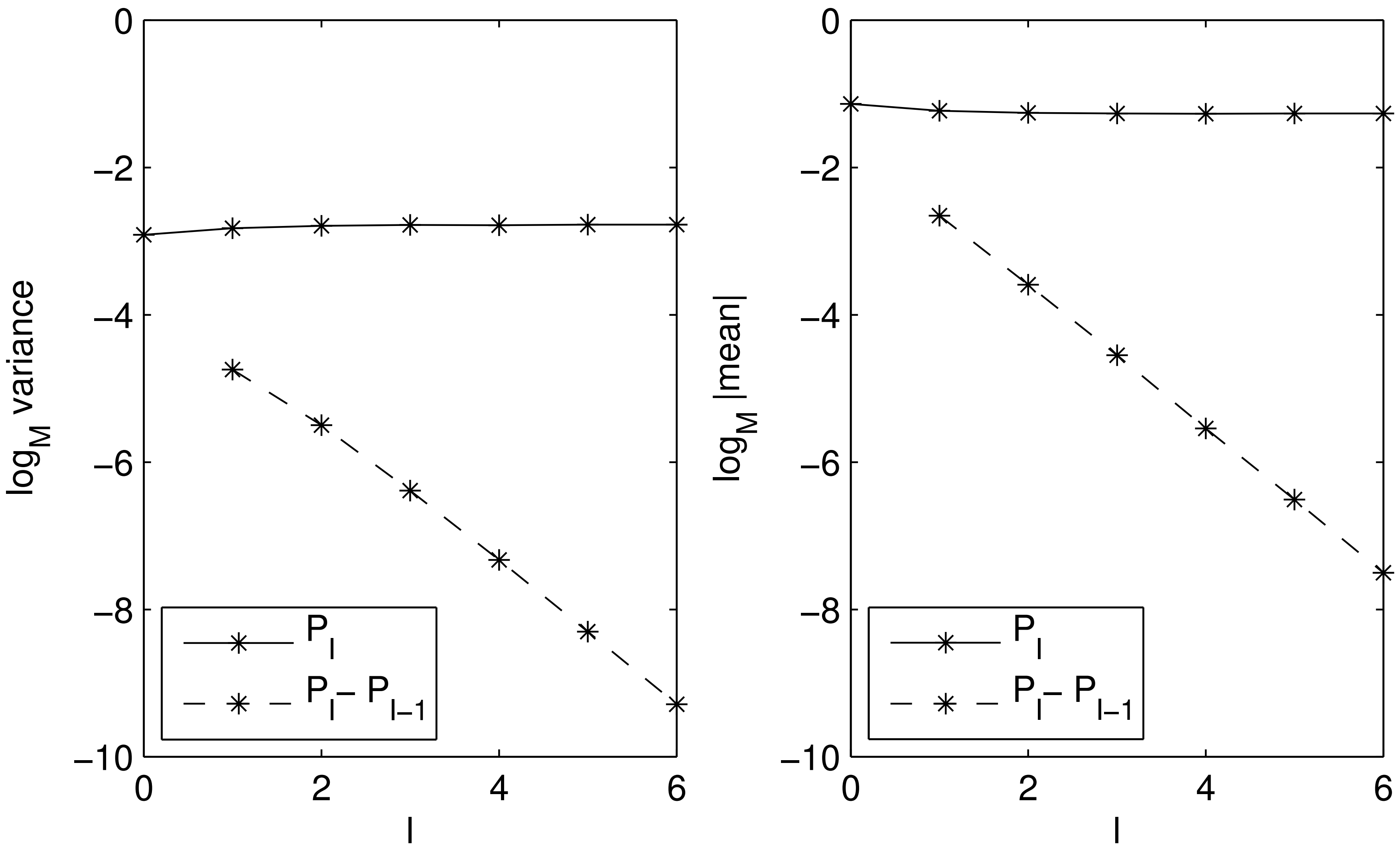
GBM: European call, $\exp(-rT)\ \max(S(T)-K, 0)$





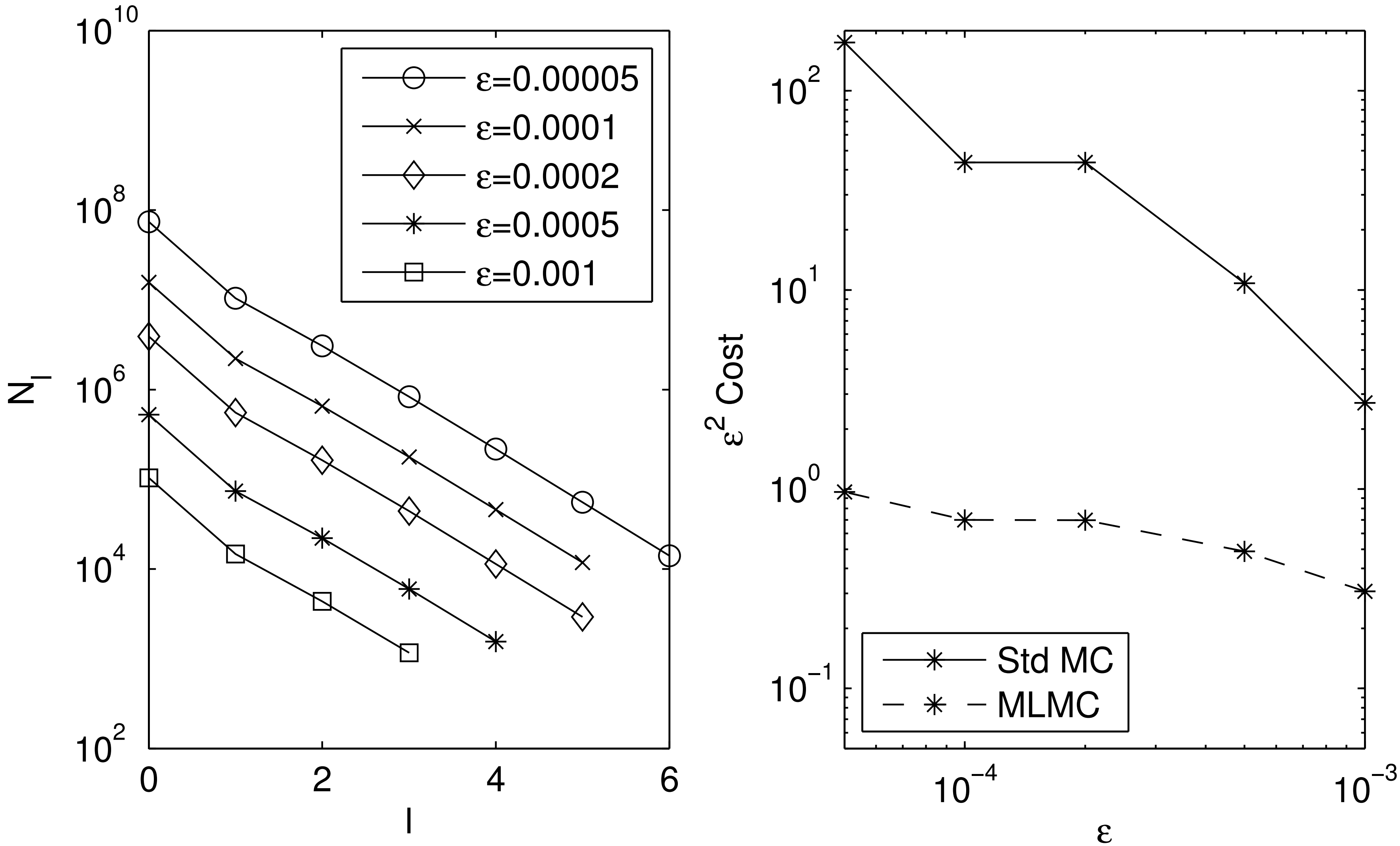
MLMC Results

GBM: lookback call, $\exp(-rT)\left(S(T) - \min_t S(t)\right)$



MLMC Results

GBM: lookback call, $\exp(-rT)\,\left(S(T) - \min_t S(t)\right)$



Extensions: I

Use of Milstein method:

- better strong convergence than Euler-Maruyama method
- able to achieve $O(\varepsilon^{-2})$ cost for
  - digital options
  - barrier options
  - lookback options
- for multi-factor SDEs needs approximate Lévy areas
  - still active research area
  - in some cases, can neglect Lévy areas and use an antithetic "trick" to get good multilevel variance without $O(h)$ strong convergence

Extensions: II

Use of Sobol points or rank-1 lattice rule for MLQMC:

- Milstein method leads to most computational effort on coarsest levels
- QMC is particularly effective at coarsest levels
- QMC doesn't provide much benefits at finer levels, but overall benefits are very significant

Extensions: III

Current research is on nested simulations relevant to Value-at-Risk (VaR) or Conditional Value-at-Risk (CVaR).

Given underlying risk factors Y, VaR is the loss level L_α such $\mathbb{P}[L > L_\alpha \,|\, Y] = \alpha$ for some small probability α.

CVaR is then $\mathbb{E}[L - L_\alpha \,|\, L > L_\alpha] = \alpha^{-1} \mathbb{E}[(L - L_\alpha)^+]$.

The complication is that the loss itself is a risk-neutral expected value – this gives a nested expectation of the form

$$\mathbb{E}\left[g\left(\mathbb{E}[f(X,Y)|Y]\right)\right]$$

In the simplest multilevel treatment, level ℓ uses 2^ℓ inner samples for the conditional expectation.

Other Work

- numerical analysis (D. Higham, X. Mao – Strathclyde)
- adaptive time-stepping (R. Tempone – KAUST)
- Greeks for hedging and risk management (S. Burgos)
- exponential Lévy processes (Y. Xia)
- multidimensional SDEs without Lévy areas (L. Szpruch – Edinburgh)
- various techniques for handling digital options
- SPDEs (C. Reisinger)
- applications to lots of other stochastic models in physics, engineering, biochemistry

See my webpages for details

Conclusions

Results:

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

However:

- still lots of scope for further developments (e.g. current research on risk analysis)
- still not taken up by banks, perhaps because they are not yet convinced savings are big enough?