MLMC for reflected diffusions

Mike Giles [1]

Eike Müller   Rob Scheichl   Tony Shardlow [2]

Kavita Ramanan [3]

[1]Mathematical Institute, University of Oxford

[2]Dept. of Mathematical Sciences, University of Bath

[3]Division of Applied Mathematics, Brown University

MCFG internal seminar

March 4, 2016

Outline

- multilevel Monte Carlo
  - current research interests
- 1D particles with mass
  - standard treatment
  - expanded domain
  - new treatment
  - results
- 1D massless particles
  - new treatment
  - results
  - financial modelling example
- multi-dimensional generalisations

Multilevel Monte Carlo

MLMC is based on the telescoping sum

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

where $\widehat{P}_\ell$ represents an approximation of some output P on level ℓ.

In SDE applications with uniform timestep $h_\ell = 2^{-\ell}\, h_0$, if the weak convergence is

$$\mathbb{E}[\widehat{P}_\ell - P] = O(2^{-\alpha\,\ell}),$$

and $\widehat{Y}_\ell$ is an unbiased estimator for $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$, based on N_ℓ samples, with variance

$$\mathbb{V}[\widehat{Y}_\ell] = O(N_\ell^{-1}\, 2^{-\beta\,\ell}),$$

and expected cost

$$\mathbb{E}[C_\ell] = O(N_\ell\; 2^{\gamma\,\ell}), \quad \ldots$$

Multilevel Monte Carlo

... then the finest level L and the number of samples N_ℓ on each level can be chosen to achieve an RMS error of ε at an expected cost

$$C = \begin{cases} O\left(\varepsilon^{-2}\right), & \beta > \gamma, \\ O\left(\varepsilon^{-2}(\log \varepsilon)^2\right), & \beta = \gamma, \\ O\left(\varepsilon^{-2-(\gamma-\beta)/\alpha}\right), & 0 < \beta < \gamma. \end{cases}$$

Multilevel Monte Carlo

The standard estimator for SDE applications is

$$\widehat{Y}_\ell = N_\ell^{-1} \sum_{n=0}^{N_\ell} \left(\widehat{P}_\ell(W^{(n)}) - \widehat{P}_{\ell-1}(W^{(n)}) \right)$$

using the same Brownian motion $W^{(n)}$ for the n^{th} sample on the fine and coarse levels.

However, there is some freedom in how we construct the coupling provided $\widehat{Y}_\ell$ is an unbiased estimator for $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$.

Also, uniform timestepping is not required – it is fairly straightforward to implement MLMC using non-nested adaptive timestepping.

(G, Lester, Whittle: MCQMC14 proceedings)

MLMC – current research

- adaptive timestepping for SDEs with non-globally Lipschitz drift (Wei Fang – talk next term?)
- long-chain molecules in solution (Endre Süli)
- stochastic biochemical reactions (Ruth Baker)
- Langevin dynamics for Big Data machine learning (Sebastian Vollmer)
- Stopped diffusions – Feynman-Kac (Francisco Bernal – IST Lisbon)
- MLMC + QMC (Frances Kuo, Ian Sloan – UNSW)
- CDF estimation (Klaus Ritter – TU Kaiserslautern)
- VaR (Ralf Korn – TU Kaiserslautern)

1D particles with mass

Position x_t and velocity u_t, subject to deterministic and stochastic forcing:

$$\begin{aligned} \mathrm{d}u_t &= a(x_t, u_t, t)\, \mathrm{d}t + b(x_t, t)\, \mathrm{d}w_t \\ \mathrm{d}x_t &= u_t\, \mathrm{d}t \end{aligned}$$

Domain $x \geq 0$, with reflection so that when it hits $x\!=\!0$ at time τ then the velocity is reflected, so

$$u_{\tau^+} = -\, u_{\tau^-}.$$

1D particles with mass

Euler-Maruyama treatment with uniform timestep h:

$$\begin{aligned} \widehat{u}_{n+1} &= s_n \left(\widehat{u}_n + a(\widehat{x}_n, \widehat{u}_n, t)\, h + b(\widehat{x}_n, t_n)\, \Delta w_n \right) \\ \widehat{x}_{n+1} &= s_n \left(\widehat{x}_n + \widehat{u}_n\, h \right) \end{aligned}$$

with $s_n = \pm 1$ chosen so that $\widehat{x}_{n+1} \geq 0$.

Problem: only $O(h^{1/2})$ strong convergence

Reason: doesn't account for reflection occurring part-way through a timestep.

1D particles with mass

Idea: if $A(X,U,t)$, $B(X,t)$ are sufficiently smooth, get $O(h)$ convergence using an extended domain:

$$\begin{aligned} \mathrm{d}U_t &= A(X_t,U_t,t)\,\mathrm{d}t + B(X_t,t)\,\mathrm{d}W_t \\ \mathrm{d}X_t &= U_t\,\mathrm{d}t, \end{aligned}$$

with

$$A(X,U,t) = \begin{cases} a(X,U,t), & X \geq 0 \\ -a(-X,-U,t), & X < 0 \end{cases}$$

$$B(X,t) = \begin{cases} b(X,t), & X \geq 0 \\ b(-X,t), & X < 0 \end{cases}$$

and then take $x = |X|$ as output.

1D particles with mass

Why does that give $O(h)$ strong convergence, but the original doesn't?

If we define

$$\begin{pmatrix} u_t \\ x_t \end{pmatrix} = S(X_t) \begin{pmatrix} U_t \\ X_t \end{pmatrix},$$

where $S(X) \equiv \mathrm{sign}(X)$, then u_t, x_t satisfy

$$\begin{aligned} \mathrm{d}u_t &= a(x_t, u_t, t)\,\mathrm{d}t + b(x_t, t)\,S(X_t)\,\mathrm{d}W_t \\ \mathrm{d}x_t &= u_t\,\mathrm{d}t, \end{aligned}$$

By setting $\mathrm{d}w_t = S(X_t)\,\mathrm{d}W_t$, we see that this is equivalent in distribution to the original model problem.

Note: strong convergence is now at fixed W_t – not the same as fixed w_t.

1D particles with mass

New MLMC treatment:

$$\begin{aligned} \widehat{u}^p_{n+1} &= \widehat{u}_n + a(\widehat{x}_n, \widehat{u}_n, t_n)\, h + b(\widehat{x}_n, t_n)\, \widehat{s}_n\, \Delta W_n \\ \widehat{x}^p_{n+1} &= \widehat{x}_n + \widehat{u}_n\, h \end{aligned}$$

followed by a correction/reflection step:

$$\begin{aligned} \widehat{u}_{n+1} &= \mathrm{sign}(\widehat{x}^p_{n+1})\ \widehat{u}^p_{n+1} \\ \widehat{x}_{n+1} &= \mathrm{sign}(\widehat{x}^p_{n+1})\ \widehat{x}^p_{n+1} \\ \widehat{s}_{n+1} &= \mathrm{sign}(\widehat{x}^p_{n+1})\ \widehat{s}_n \end{aligned}$$

with <u>same</u> Brownian path for coarse and fine levels.

Can show that when a and b are both constant, the coarse and fine paths are identical at coarse timesteps.

1D particles with mass

Test case 1:

$$x_0 = 0.2, \quad u_0 = -0.2, \quad a(x,t) = 0, \quad b(x,t) = 0.5.$$

in domain $0 \leq x \leq 1$, with reflection at both boundaries.

Output of interest: $\int_0^1 x_t \, \mathrm{d}t$ approximated by $\displaystyle\sum_{n=1}^{2^\ell} h_\ell \, \widehat{x}_n$.

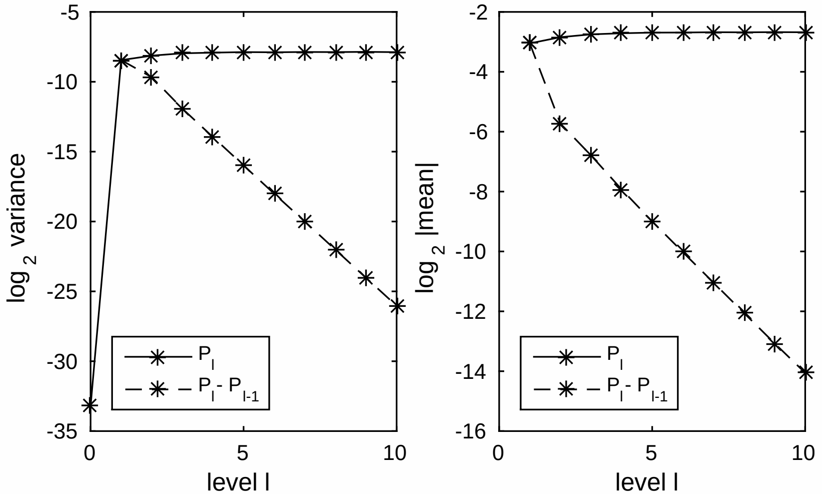
Test case 2: changes drift, volatility to

$$a(x,t) = -0.2, \quad b(x,t) = 0.5 + 0.5\,x.$$

– standard $O(h)$ numerical analysis no longer applies

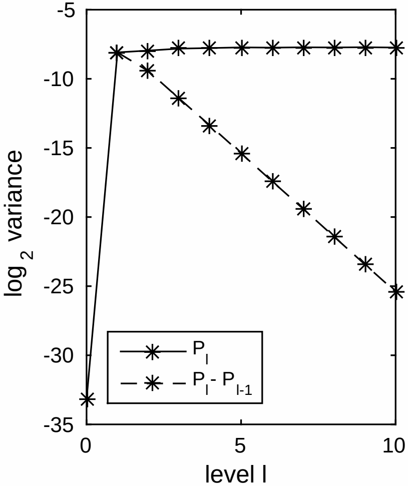
1D particles with mass

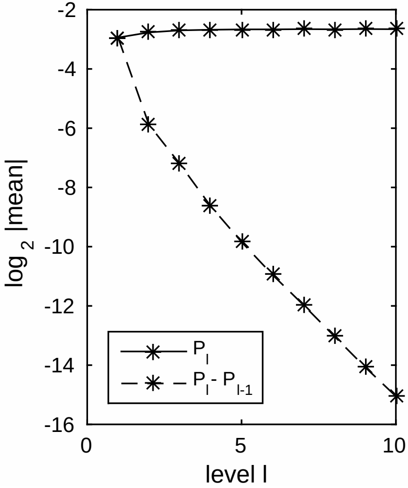
Test case 1: $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2$ $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell$



1D particles with mass

Test case 2: $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2$ $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell$





1D massless particles

Without mass, the SDE is

$$\mathrm{d}x_t = a(x_t, t)\,\mathrm{d}t + b(x_t, t)\,\mathrm{d}w_t$$

and if the domain is $x \geq 0$, particles are prevented from crossing $x = 0$.

Euler-Maruyama treatment with uniform timestep h:

$$\widehat{x}_{n+1} = \left| \widehat{x}_n + a(\widehat{x}_n, t)\,h + b(\widehat{x}_n, t_n)\,\Delta w_n \right|$$

Again only $O(h^{1/2})$ strong convergence, even when b is uniform

1D massless particles

Thinking about the extended domain leads to

$$\mathrm{d}x_t = a(x_t, t)\,\mathrm{d}t + b(x_t, t)\,S(X_t)\,\mathrm{d}W_t$$

where $S(X) \equiv \mathrm{sign}(X)$, and hence the numerical approximation is

$$\widehat{x}^p_{n+1} = \widehat{x}_n + a(\widehat{x}_n, t_n)\,h + b(\widehat{x}_n, t_n)\,\widehat{s}_n\,\Delta W_n$$

followed by a correction/reflection step:

$$\begin{aligned} \widehat{x}_{n+1} &= \mathrm{sign}(\widehat{x}^p_{n+1})\;\widehat{x}^p_{n+1} \\ \widehat{s}_{n+1} &= \mathrm{sign}(\widehat{x}^p_{n+1})\;\widehat{s}_n \end{aligned}$$

with <u>same</u> Brownian path for coarse and fine levels.

Note: if b is not uniform then we need to use first order Milstein approximation to get $O(h)$ strong convergence.

1D massless particles

Test case 1:

$$x_0 = 0.2, \quad a(x,t) = 0, \quad b(x,t) = 0.5.$$

in domain $0 \leq x \leq 1$, with reflection at both boundaries.

Output of interest: $\int_0^1 x_t \, \mathrm{d}t$ approximated by $\displaystyle\sum_{n=1}^{2^\ell} h_\ell \, \widehat{x}_n$.

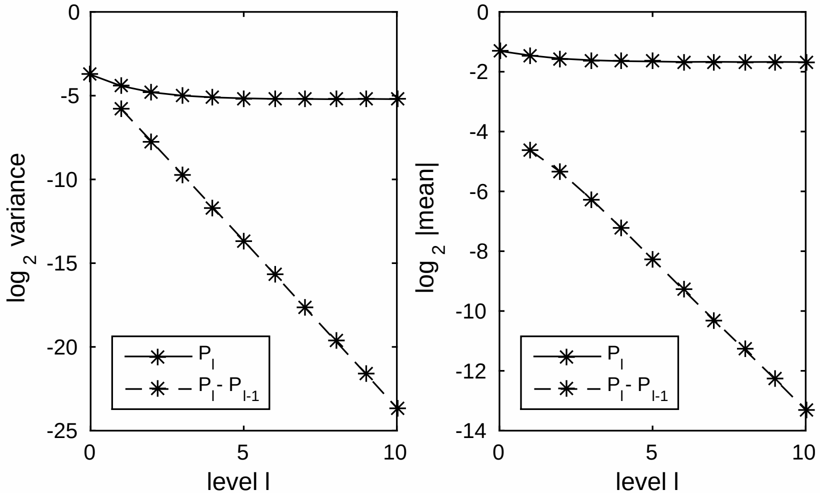
Test case 2: changes drift, volatility to

$$a(x,t) = -0.2, \quad b(x,t) = 0.5 + 0.5\,x.$$

– standard $O(h)$ numerical analysis no longer applies

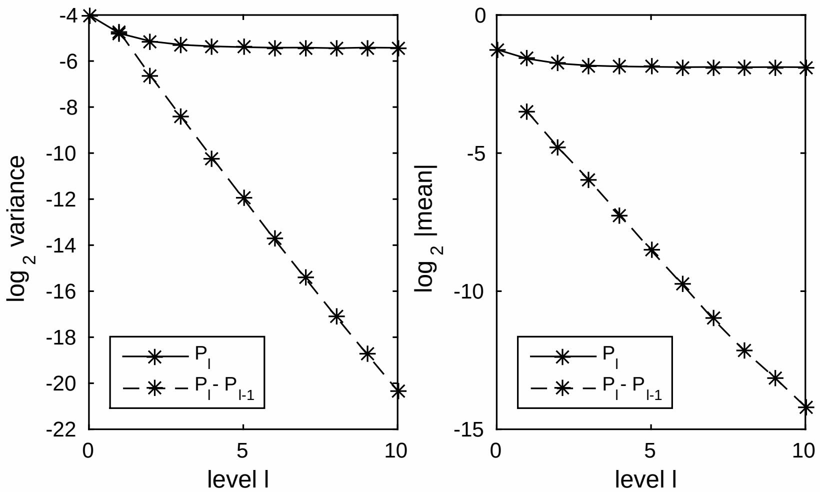
1D massless particles

Test case 1: $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2$ $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell$



1D massless particles

Test case 2: $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^{3/2}$ $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell$



1D massless particles

Why is the variance $O(h^{3/2})$?

Ad-hoc explanation:

- $O(1)$ path density near $x=0$
- $O(h^{1/2})$ movement in each timestep
- $\Longrightarrow O(h^{1/2})$ probability of crossing boundary in each timestep
- $\Longrightarrow O(h^{-1/2})$ total crossings per path
- each crossing gives error which is $O(h)$ but has near-zero mean
- if crossings are approximately independent, then
  $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = O(h^{-1/2} \times h^2) = O(h^{3/2})$

Note: in the case with mass, the velocity is $O(1)$, the movement in each timestep is $O(h)$, so the number of crossings is $O(1) \implies V_\ell = O(h^2)$.

Financial modelling example

If a central bank acts to keep an exchange rate x within a given range $[x_1, x_2]$, this can be modelled by a reflected Ornstein-Uhlenbeck process:

$$\mathrm{d}x_t = \kappa\,(x_{equil} - x_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + \mathrm{d}L_{1,t} - \mathrm{d}L_{2,t}$$

where $x_1 < x_{equil} < x_2$ is the equlibrium value, $L_{1,t}$ is a local time which increases only when $x_t = x_1$, and $L_{2,t}$ is a local time which increases only when $x_t = x_2$.

The local times correspond here to the sale/purchase of currency by the central bank to keep the rate within limits. (Yang *et al*, 2012)

A new MSc project will look at this model, its MLMC implementation, and other financial applications.

Multi-dimensional extensions

In simple cases:

- isotropic volatility
- normal reflection

the 1D ideas extend fairly naturally to multi-dimensional applications

Good for engineering applications (e.g. 3D atmospheric pollutant dispersal)

However, in general multi-dimensional applications are much more complicated.

MLMC for reflected diffusions

Joint research with Kavita Ramanan (Brown University)

Motivation comes from network queue analysis, approximated by a reflected Brownian diffusion within a domain D, with SDE

$$\mathrm{d}x_t = a(x_t)\,\mathrm{d}t + b\,\mathrm{d}W_t + \nu(x_t)\,\mathrm{d}L_t$$

where L_t is a local time which increases when x_t is on the boundary ∂D.

$\nu(x_t)$ can be normal to the boundary (pointing inwards), but in other cases it is not and reflection from the boundary includes a tangential motion.

A penalised version is

$$\begin{aligned} \mathrm{d}x_t &= a(x_t)\,\mathrm{d}t + b\,\mathrm{d}W_t + \nu(x_t)\,\mathrm{d}L_t \\ \mathrm{d}L_t &= -\lambda \min(0, d(x_t))\,\mathrm{d}t, \qquad \lambda \gg 1 \end{aligned}$$

where $d(x_t)$ is signed distance to the boundary – negative means outside.

MLMC for reflected diffusions

3 different numerical treatments:

- projection: predictor step:

$$\widehat{X}^{(p)} = \widehat{X}_{t_n} + a(\widehat{X}_{t_n}, t_n)\, h_n + b\, \Delta W_n,$$

  followed by correction step

$$\widehat{X}_{t_{n+1}} = \widehat{X}^{(p)} + \nu(\widehat{X}^{(p)})\ \Delta \widehat{L}_n,$$

  with $\Delta \widehat{L}_n > 0$ if needed to put $\widehat{X}_{t_{n+1}}$ on boundary

- reflection: similar but with double the value for $\Delta \widehat{L}_n$ – can give improved weak convergence

- penalised: Euler-Maruyama approximation of penalised SDE

MLMC for reflected diffusions

Concern:

- because b is uniform, Euler-Maruyama method corresponds to first order Milstein scheme, suggesting an $O(h)$ strong error
- however, all treatments of boundary reflection lead to a strong error which is $O(h^{1/2})$ – this is based primarily on empirical evidence, with only limited supporting theory

Idea:

- use adaptive timesteps, with level ℓ timestep given by

$$\max\left(2^{-2\ell}h_0, \min\left(2^{-\ell}h_0, (d/((\ell+3)\,\|b\|_2)^2\right)\right).$$

  based on distance d to boundary.

MLMC for reflected diffusions

This max-min definition leads to 3 zones:

- a boundary zone where $h = 2^{-2\ell} h_0$
- an interior zone where $h = 2^{-\ell} h_0$
- an intermediate zone where $(\ell{+}3)\sqrt{h}\|b\|_2 = d$

As $\ell \to \infty$, there is a very high probability that no reflections take place from the interior or intermediate zones.

- boundary error is $O(\sqrt{2^{-2\ell} h_0}\ ) = O(2^{-\ell})$
- interior error is $O(2^{-\ell} h_0) = O(2^{-\ell})$
- overall, strong error is $O(2^{-\ell}) \implies$ MLMC variance $= O(2^{-2\ell})$.

MLMC for reflected diffusions

Current theoretical analysis:

- if strong error is $O(\sqrt{h}\,)$ for uniform timestep then the MLMC variance is $O(2^{-2\ell})$ for Lipschitz functionals.
- the expected cost is $o(2^{(1+\delta)\ell})$ for any $0 < \delta \ll 1$
- regarding MLMC theory, this gives $\beta = 2, \gamma \approx 1$, so the complexity is $O(\varepsilon^{-2})$ for ε r.m.s. error

Numerical analysis challenge:

- prove that the strong error is $O(\sqrt{h}\,)$ for uniform timestep with oblique reflections, preferably for generalised penalisation method for polygonal boundaries

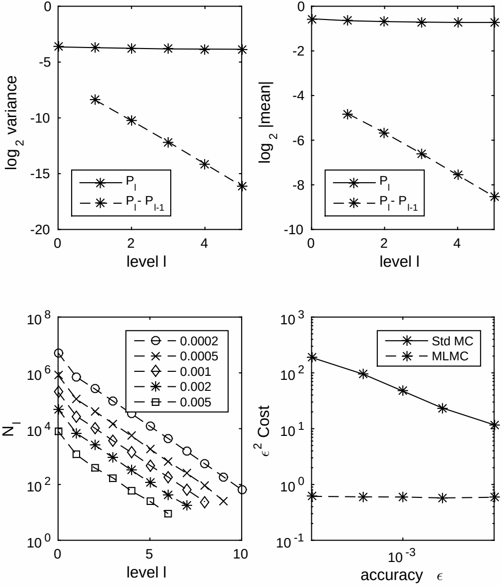
MLMC for reflected diffusions

Simple test case:

- 3D Brownian motion in a unit ball
- normal reflection at the boundary
- $x_0 = 0$
- aim is to estimate $\mathbb{E}[\|x\|_2^2]$ at time $t = 1$.
- implemented with both projection and penalisation schemes

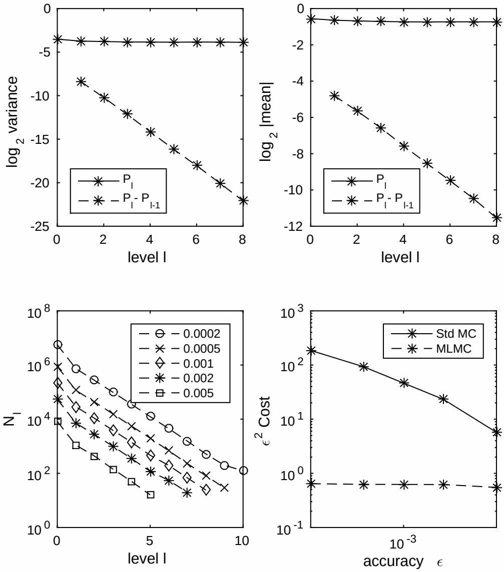
MLMC for reflected diffusions

Projection method:



MLMC for reflected diffusions

Penalisation method:



Conclusions

- simple reflection "trick" improves the MLMC variance for 1D reflected diffusions, for particles with or without mass
- the extension to multiple dimensions should work in simple cases, but not in more general cases
- more difficult cases can use adaptive timestepping, and we're making progress on the numerical analysis
- very keen to hear about new financial applications