MLMC for reflected diffusions

Mike Giles [1] Eike Müller [2] Rob Scheichl [2] Tony Shardlow [2]

[1]Mathematical Institute, University of Oxford

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

SciCADE 2015

Sept 16, 2015

Outline

- multilevel Monte Carlo
- 1D particles with mass
  - standard treatment
  - expanded domain
  - new treatment
  - results
- 1D massless particles
  - new treatment
  - results
- 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 using on level ℓ.

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

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

$\widehat{Y}_\ell$ 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)

1D particles with mass

Position x_t, and velocity u_t subject to steady 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

Key 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 $\displaystyle\sum_{n=1}^{2^\ell} h_\ell \, \widehat{x}_n$.

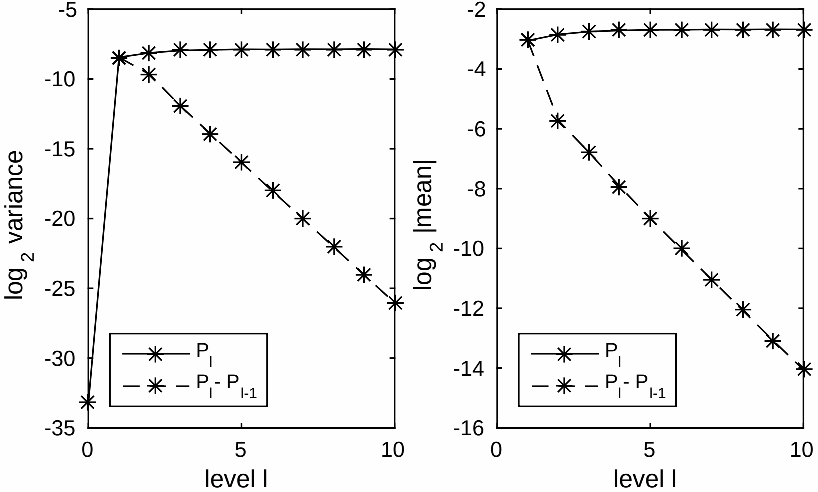
Test case 2: changes drift, volatility to

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

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

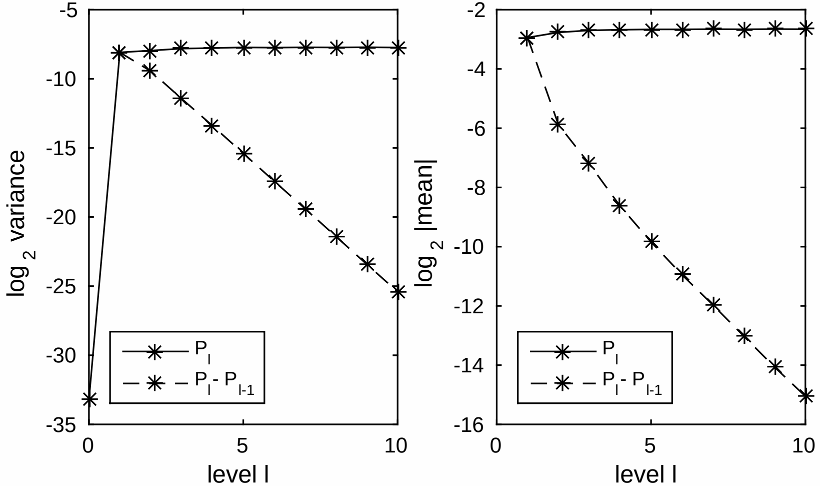
1D particles with mass

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 $\displaystyle\sum_{n=1}^{2^\ell} h_\ell \, \widehat{x}_n$.

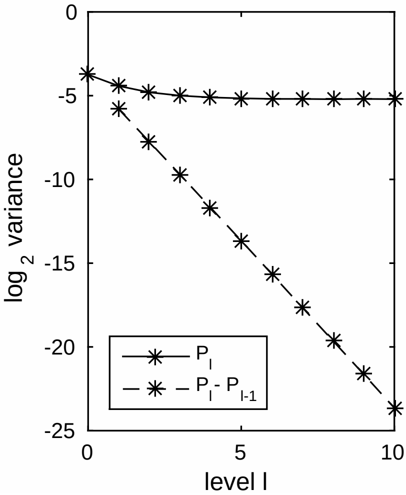
Test case 2: changes drift, volatility to

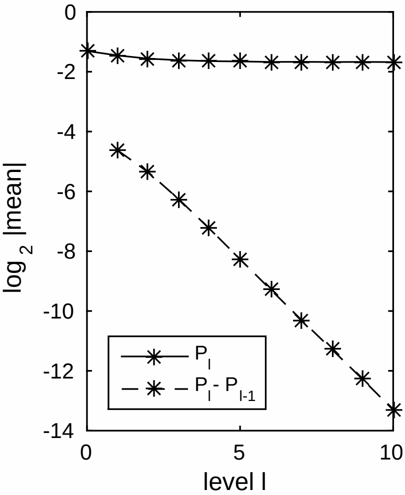
$$a(x,t) = -0.2, \quad b(x,t) = 0.5 + 0.5\, x.$$

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

1D massless particles

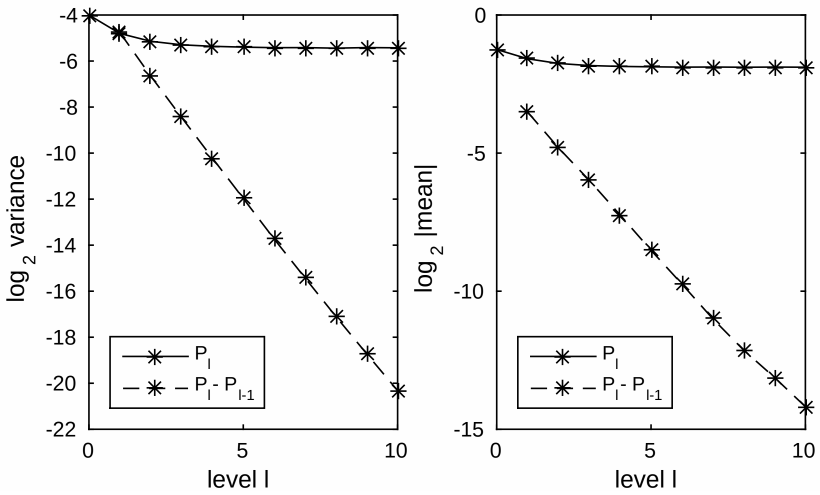
Test case 1: $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2$ $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell$





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 $O(1) \implies V_\ell = O(h^2)$.

Multi-dimensional extensions

Start with particles with mass, and 2D domain with boundary at $x\!=\!0$.

If

$$a = \begin{pmatrix} a_x(x,y) \\ a_y(x,y) \end{pmatrix}, \qquad b = I,$$

then for $x\!<\!0$ extended domain SDE naturally has

$$A = \begin{pmatrix} -a_x(-x,y) \\ a_y(-x,y) \end{pmatrix}, \qquad B = I,$$

This leads to an SDE in the regular half-plane in which the x-component of $\mathrm{d}W$ is "flipped" each time the boundary is hit.

Multi-dimensional extensions

However, if

$$a = \begin{pmatrix} a_x(x,y) \\ a_y(x,y) \end{pmatrix}, \qquad b = \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix},$$

then when $x < 0$

$$A = \begin{pmatrix} -a_x(-x,y) \\ a_y(-x,y) \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ -\beta & 1 \end{pmatrix},$$

because if $X = -x$ then

$$\mathbb{E}[\,\mathrm{d}X\,\mathrm{d}y\,] = -\,\mathbb{E}[\,\mathrm{d}x\,\mathrm{d}y\,].$$

This gives a discontinuity in B, and there seems no way to get $O(h)$ strong convergence.

Multi-dimensional extensions

Hence, multi-dimensional extension for particles with mass will only work in simple cases.

For massless particles, there is extra complication of oblique reflections.

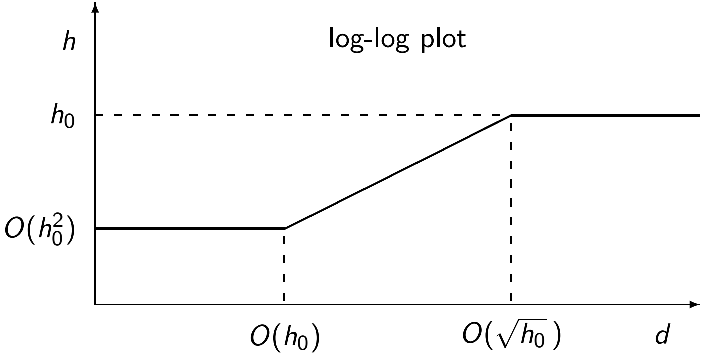
If diffusion is isotropic (i.e. b is a multiple of the identity matrix) and reflections are normal, then can probably get $O(h)$ strong convergence.

Otherwise, probably not.

Multi-dimensional extensions

What else can we do for massless case?

Go back to starting point with standard reflection treatment, and use adaptive timesteps based on distance d to boundary.



$O(h_0^2)$ timestep near boundary $\Longrightarrow$ $O(h_0)$ strong error

Multi-dimensional extensions

In MLMC, on level ℓ would use

$$h = \min\left(2^{-\ell} h_0,\ \max\left(2^{-2\ell} h_0, (\tfrac{1}{3} d/\|b\|_2)^2\right)\right)$$

where d is the distance to the boundary

- $2^{-\ell} h_0$ in interior
- $2^{-2\ell} h_0$ next to boundary
- $(\frac{1}{3} d/\|b\|_2)^2$ in layer in-between

The factor $\frac{1}{3}$ implies $3\sqrt{h}\ \|b\|_2 < d$ so that boundary crossings from the intermediate zone (or interior zone) are unlikely.

Most will come from boundary zone, with resultant $O(2^{-\ell})$ strong error.

Multi-dimensional extensions

Key observation: the computational cost is proportional to $\int_D h^{-1}\,\mathrm{d}x$ and this is $O(2^\ell)$.

Hence, in the usual MLMC theorem, we should get

$$\alpha = 1, \quad \beta = 2, \quad \gamma = 1,$$

and hence obtain $O(\varepsilon^{-2})$ complexity.

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

Webpages: `http://people.maths.ox.ac.uk/gilesm/`