Numerical analysis of MLMC for multi-dimensional SDEs using antithetic Milstein discretisation

Lukasz Szpruch
Mike Giles

Oxford University Mathematical Institute
Oxford-Man Institute of Quantitative Finance

MCQMC 2012
Outline

▷ A new MLMC
▷ Clark-Cameron example
▷ General Case
▷ Pricing with antithetic MLMC
To estimate $\mathbb{E}[P]$ where the payoff $P = f(x_T)$ can be approximated by $\hat{P}_l$ using $\Delta t_l = 2^{-l} T$ uniform timesteps, we use

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

$\mathbb{E}[\hat{P}_l - \hat{P}_{l-1}]$ is estimated using $N_l$ simulations with the same Brownian path $w(t)$ for both $\hat{P}_l$ and $\hat{P}_{l-1}$,

$$
\hat{Y}_l = N_l^{-1} \sum_{i=1}^{N_l} \left( \hat{P}_l^{(i)} - \hat{P}_{l-1}^{(i)} \right)
$$

Because of the strong convergence, on finer levels $\mathbb{V}[\hat{P}_l - \hat{P}_{l-1}]$ is small and so few paths are required.
Problems

- Can we avoid simulating the Lévy areas for Milstein scheme in multidimensional setting?
- Can we extend the strong convergence theorems to the non-global Lipschitz case?
- How to deal with non-smooth, non-differentiable payoffs?
Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ be a complete probability space.

Consider an $n$-dimensional SDE

$$dx(t) = f(x(t))dt + g(x(t))dw(t)$$

on $t \geq 0$ with the initial value $x(0) = x_0 \in \mathbb{R}^n$, where $w(t)$ is an $m$-dimensional Brownian motion, while

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \text{and} \quad g : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}.$$
Milstein Scheme

\[ \hat{X}_{i,n+1} = \hat{X}_{i,n} + f_i(\hat{X}_n) \Delta t + \sum_{j=1}^{D} g_{ij}(\hat{X}_n) \Delta w_{j,n} \]

\[ + \sum_{j,k=1}^{D} h_{ijk}(\hat{X}_n) (\Delta w_{j,n} \Delta w_{k,n} - \Omega_{jk} \Delta t - A_{jk,n}) \]

where \( h_{ijk}(x) = \frac{1}{2} \sum_{l=1}^{d} g_{lk}(x) \frac{\partial g_{ij}}{\partial x_l}(x) \), \( \Omega \) is the correlation matrix and \( A_{jk,n} \) is the Lévy area defined as

\[ A_{jk,n} = \int_{t_n}^{t_{n+1}} (w_j(t) - w_j(t_n)) \, dw_k(t) - \int_{t_n}^{t_{n+1}} (w_k(t) - w_k(t_n)) \, dw_j(t). \]

If \( g(x) \) has a \textit{commutative} property then the Lévy areas term is canceled.
Modified Multilevel approach

Sometimes it is better to use a different approximation for $\hat{P}_l$ and $\hat{P}_{l-1}$ in $\mathbb{E}[\hat{P}_l - \hat{P}_{l-1}]$. A new MLMC

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

is still a valid telescoping sum provided $\mathbb{E}[\hat{P}_l^f] = \mathbb{E}[\hat{P}_l^c]$.

In this work, we use $\hat{P}_l^c = P(X_l^c)$ and

$$\hat{P}_l^f = \frac{1}{2} \left( P(X_l^f) + P(X_l^a) \right)$$

where $X_l^f$ is the fine path, and $X_l^a$ is an “antithetic twin”.
Do we really need a good strong convergence?

Lemma

If $P \in C^2(\mathbb{R}^n, \mathbb{R})$ and there exist constants $L_1, L_2$ such that for all $x \in \mathbb{R}^n$

$$\left\| \frac{\partial P}{\partial x} \right\| \leq L_1, \quad \left\| \frac{\partial^2 P}{\partial x^2} \right\| \leq L_2.$$ 

then for $p \geq 2$,

$$E \left[ \left\| \frac{1}{2} (P(X^f) + P(X^a)) - P(X^c) \right\|^p \right] \leq 2^{p-1} L_1^p E \left[ \left\| \frac{1}{2} (X^f + X^a) - X^c \right\|^p \right] + 2^{-(p+1)} L_2^p E \left[ \left\| X^f - X^a \right\|^{2p} \right].$$
Clark & Cameron problem

In their 1980 paper, Clark & Cameron considered the model problem:

\[
\begin{aligned}
    dx_1(t) &= dw_1(t) \\
    dx_2(t) &= x_1(t)dw_2(t),
\end{aligned}
\]

for independent \( w_1, w_2 \), with \( x_1(0) = x_2(0) = 0 \). We have

\[
x_2(t_{k+1}) = x_2(t_k) + x_1(t_k)\Delta w_{2,k+1} + \frac{1}{2}\Delta w_{1,k+1}\Delta w_{2,k} + \frac{1}{2}[A_{1,2}]_{t_k}^{t_{k+1}}.
\]

For \textit{any} numerical approximation \( X(T) \) based solely on the set of discrete Brownian increments \( \Delta w \),

\[
\mathbb{E}[\|x_2(T) - X_2(T)\|^2] \geq \frac{1}{4} T \Delta t.
\]
Clark & Cameron problem

$X^c$ with timestep $\Delta t$ by neglecting the Lévy area terms to give

\[
\begin{align*}
X_{1,n+1}^c &= X_{1,n}^c + \Delta w_{1,n}, \\
X_{2,n+1}^c &= X_{2,n}^c + X_{1,n}^c \Delta w_{2,n} + \frac{1}{2} \Delta w_{1,n} \Delta w_{2,n}
\end{align*}
\]
Clark & Cameron problem

\[ X^f_{1,n+\frac{1}{2}} = X^f_{1,n} + \delta w_{1,n} \quad \text{and} \quad X^a_{1,n+\frac{1}{2}} = X^a_{1,n} + \delta w_{1,n+\frac{1}{2}} \]

\[ X^f_{2,n+\frac{1}{2}} = X^f_{2,n} + X^f_{1,n} \delta w_{2,n} + \frac{1}{2} \delta w_{1,n} \delta w_{2,n} \]
\[ X^a_{2,n+\frac{1}{2}} = X^a_{2,n} + X^a_{1,n} \delta w_{2,n+\frac{1}{2}} + \frac{1}{2} \delta w_{1,n+\frac{1}{2}} \delta w_{2,n+\frac{1}{2}} \]

\[ X^f_{1,n+1} = X^f_{1,n+\frac{1}{2}} + \delta w_{1,n+\frac{1}{2}} \quad \text{and} \quad X^a_{1,n+1} = X^a_{1,n} + \delta w_{1,n} \]

\[ X^f_{2,n+1} = X^f_{2,n+\frac{1}{2}} + X^f_{1,n+\frac{1}{2}} \delta w_{2,n+\frac{1}{2}} + \frac{1}{2} \delta w_{1,n+\frac{1}{2}} \delta w_{2,n+\frac{1}{2}} \]
\[ X^a_{2,n+1} = X^a_{2,n} + X^a_{1,n+\frac{1}{2}} \delta w_{2,n} + \frac{1}{2} \delta w_{1,n} \delta w_{2,n} \]

in which \( \delta w_n \equiv w(t_{n+\frac{1}{2}}) - w(t_n), \quad \delta w_{n+\frac{1}{2}} \equiv w(t_{n+1}) - w(t_{n+\frac{1}{2}}) \)
Clark & Cameron problem

\[ X_{1,n+1}^f = X_{1,n}^f + \Delta w_{1,n} \]
\[ X_{2,n+1}^f = X_{2,n}^f + X_{1,n}^f \Delta w_{2,n} + \frac{1}{2} \Delta w_{1,n} \Delta w_{2,n} \]
\[ + \frac{1}{2} \left( \delta w_{1,n} \delta w_{2,n+\frac{1}{2}} - \delta w_{2,n} \delta w_{1,n+\frac{1}{2}} \right). \]

\[ X_{1,n+1}^a = X_{1,n}^a + \Delta w_{1,n}, \]
\[ X_{2,n+1}^a = X_{2,n}^a + X_{1,n}^a \Delta w_{2,n} + \frac{1}{2} \Delta w_{1,n} \Delta w_{2,n} \]
\[ - \frac{1}{2} \left( \delta w_{1,n} \delta w_{2,n+\frac{1}{2}} - \delta w_{2,n} \delta w_{1,n+\frac{1}{2}} \right). \]
Clark & Cameron problem

Lemma

If $X^f_n$, $X^a_n$ and $X^c_n$ are as defined above, then

\[
X^f_{1,n} = X^a_{1,n} = X^c_{1,n}, \quad \frac{1}{2} \left( X^f_{2,n} + X^a_{2,n} \right) = X^c_{2,n}, \quad \forall n \leq N
\]

and

\[
\mathbb{E} \left[ \left( X^f_{2,N} - X^a_{2,N} \right)^4 \right] = \frac{3}{4} T \left( T + \Delta t \right) \Delta t^2.
\]
Time-reversed Brownian Motion

\( (w_t, 0 \leq t \leq 1) \) - Brownian motion on the time interval \([0, 1]\)

\( (z(t), 0 \leq t \leq 1) = (w(1) - w(1 - t), 0 \leq t \leq 1) \) in distribution.

We can re-write it in the following form

\( (z(t_k + t) - w(t_k), 0 \leq t \leq t_{k+1} - t_k) \)

\( = (w(t_{k+1}) - w(t_{k+1} - t), 0 \leq t \leq t_{k+1} - t_k) \) in distribution.

Lemma

\[ [A_{i,j}]_0^1 = -[\hat{A}_{i,j}]_0^1 \] in distribution,

where \( \hat{A}_{i,j} \) is a Lévy area generated by time-reversed Brownian motion.
Assumption

Let \( f \in C^2(\mathbb{R}^d, \mathbb{R}^d) \) and \( g \in C^2(\mathbb{R}^d, \mathbb{R}^{d \times m}) \) have bounded first and second derivatives.

Truncated Milstein approximation

\[
X_{i,n+1} = X_{i,n} + f_i(X_n) \Delta t + \sum_{j=1}^{D} g_{ij}(X_n) \Delta w_{j,n} \\
+ \sum_{j,k=1}^{D} h_{ijk}(X_n) (\Delta w_{j,n} \Delta w_{k,n} - \Omega_{jk} \Delta t).
\]

Lemma

For \( p \geq 2 \), there exists a constant \( K_p \), independent of the time step, such that

\[
\mathbb{E} \left[ \max_{0 \leq n \leq N} \| X_n - x(t_n) \|^p \right] \leq K_p \Delta t^{p/2}.
\]
Lemma

Difference equation for $X^f_n$ can be expressed as

$$X^f_{i,n+1} = X^f_{i,n} + f_i(X^f_n) \Delta t + \sum_{j=1}^{D} g_{ij}(X^f_n) \Delta w_{j,n}$$

$$+ \sum_{j,k=1}^{D} h_{ijk}(X^f_n) (\Delta w_{j,n} \Delta w_{k,n} - \Omega_{jk} \Delta t)$$

$$- \sum_{j,k=1}^{D} h_{ijk}(X^f_n) \left( \delta w_{j,n} \delta w_{k,n+\frac{1}{2}} - \delta w_{k,n} \delta w_{j,n+\frac{1}{2}} \right)$$

$$+ M^f_{i,n} + N^f_{i,n},$$

where $\mathbb{E}[M^f_n | \mathcal{F}_n] = 0$, and for any integer $p \geq 2$ there exists a constant $K_p$ such that

$$\max_{0 \leq n \leq N} \mathbb{E} \left[ \| M^f_n \|^p \right] \leq K_p \Delta t^{3p/2}, \quad \max_{0 \leq n \leq N} \mathbb{E} \left[ \| N^f_n \|^p \right] \leq K_p \Delta t^{2p}. $$
Lemma

Difference equation for $X_n^a$ can be expressed as

$$
X_{i,n+1}^a = X_{i,n}^a + f_i(X_n^a) \Delta t + \sum_{j=1}^D g_{ij}(X_n^a) \Delta w_{j,n}
+ \sum_{j,k=1}^D h_{ijk}(X_n^a) (\Delta w_{j,n} \Delta w_{k,n} - \Omega_{jk} \Delta t)
+ \sum_{j,k=1}^D h_{ijk}(X_n^a) \left( \delta w_{j,n} \delta w_{k,n+\frac{1}{2}} - \delta w_{k,n} \delta w_{j,n+\frac{1}{2}} \right)
+ M_{i,n}^a + N_{i,n}^a,
$$

where $\mathbb{E}[M_n^a \mid \mathcal{F}_n] = 0$, and for any integer $p \geq 2$ there exists a constant $K_p$ such that

$$
\max_{0 \leq n \leq N} \mathbb{E} \left[ \| M_n^a \|^p \right] \leq K_p \Delta t^{3p/2}, \quad \max_{0 \leq n \leq N} \mathbb{E} \left[ \| N_n^a \|^p \right] \leq K_p \Delta t^{2p}.
$$
Lemma

The difference equation for \( X_n^f \equiv \frac{1}{2}(X_n^f + X_n^a) \) can be expressed as

\[
X_{i,n+1}^f = X_{i,n}^f + f_i(X_n^f) \Delta t + \sum_{j=1}^{D} g_{ij}(X_n^f) \Delta w_{j,n} \\
+ \sum_{j,k=1}^{D} h_{ijk}(X_n^f) (\Delta w_{j,n} \Delta w_{k,n} - \Omega_{jk} \Delta t) \\
+ M_{i,n} + N_{i,n},
\]

where \( \mathbb{E}[M_n | \mathcal{F}_n] = 0 \), and for any integer \( p \geq 2 \) there exists a constant \( K_p \) such that

\[
\max_{0 \leq n \leq N} \mathbb{E} \left[ \| M_n \|^p \right] \leq K_p \Delta t^{3p/2}, \quad \max_{0 \leq n \leq N} \mathbb{E} \left[ \| N_n \|^p \right] \leq K_p \Delta t^{2p}.
\]
Main Results

**Theorem**

For all $p \geq 2$, there exists a constant $K_p$ such that

$$
\mathbb{E} \left[ \max_{0 \leq n \leq N} \| X_f^n - X_c^n \|^p \right] \leq K_p \Delta t^p.
$$

**Lemma**

For all integers $p \geq 2$, there exists a constant $K_p$ such that

$$
\mathbb{E} \left[ \max_{0 \leq n \leq N} \| X_f^n - X_a^n \|^p \right] \leq K_p \Delta t^{p/2}
$$
Vanilla options

Example of payoff in 2D,

\[ P(x(T)) \equiv \max(0, \min(x_1(T), x_2(T)) - K), \]

Assumption

The payoff \( P \in C(\mathbb{R}^d, \mathbb{R}) \) has a uniform Lipschitz bound, so that there exists a constant \( L \) such that

\[ |P(x) - P(y)| \leq L \|x - y\|, \quad \forall x, y \in \mathbb{R}^d, \]

and the first and second derivatives exist, are continuous and have uniform bound \( L \) at all points \( x \notin K \), where \( K \) is a set of zero measure, and there exists a constant \( c \) such that the probability of the SDE solution \( x(T) \) being within a neighborhood of the set \( K \) has the bound

\[ \mathbb{P} \left( \min_{y \in K} \|x(T) - y\| \leq \varepsilon \right) \leq c \varepsilon, \quad \forall \varepsilon > 0. \]
Vanilla options

**Theorem**

*Under above assumptions*

\[
\mathbb{E} \left[ \left( \frac{1}{2} (P(X^f_N) + P(X^a_N)) - P(X^c_N) \right)^2 \right] = o(\Delta t^{3/2-\delta})
\]

*for any* \( \delta > 0 \).
Asian options

For an Asian option, the payoff depends on the average

\[ x_{\text{ave}} \equiv T^{-1} \int_0^T x(t) \, dt. \]

This can be approximated by integrating the appropriate piecewise linear interpolant which gives

\[ X_{\text{ave}}^c \equiv T^{-1} \int_0^T X^c(t) \, dt = N^{-1} \sum_{n=0}^{N-1} \frac{1}{2} (X_n^c + X_{n+1}^c), \]

\[ X_{\text{ave}}^f \equiv T^{-1} \int_0^T X^f(t) \, dt = N^{-1} \sum_{n=0}^{N-1} \frac{1}{4} (X_n^f + 2X_{n+\frac{1}{2}}^f + X_{n+1}^f), \]

\[ X_{\text{ave}}^a \equiv T^{-1} \int_0^T X^a(t) \, dt = N^{-1} \sum_{n=0}^{N-1} \frac{1}{4} (X_n^a + 2X_{n+\frac{1}{2}}^a + X_{n+1}^a). \]
Analysis on the fine grid

Lemma

For all integers $p \geq 2$, there exists a constant $K_p$ such that

$$\max_{0 \leq n < N} \mathbb{E} \left[ \|X_{n+\frac{1}{2}}^f - X_{n+\frac{1}{2}}^a\|^p \right] \leq K_p \Delta t^{p/2}.$$ 

Lemma

For all $p \geq 2$, there exists a constant $K_p$ such that

$$\max_{0 \leq n < N} \mathbb{E} \left[ \left\| X_{n+\frac{1}{2}}^f - X^c(t_{n+\frac{1}{2}}) \right\|^p \right] \leq K_p \Delta t^p,$$

where $X^c(t_{n+\frac{1}{2}}) = \frac{1}{2}(X_n^c + X_{n+1}^c)$ is the midpoint value of the coarse path interpolant.
Piecewise linear interpolant

The piecewise linear interpolant over \([t_k, t_{k+1}]\) is

\[ X(t) \equiv (1 - \lambda) X_k + \lambda X_{k+1}, \quad \lambda \equiv \frac{t - t_k}{t_{k+1} - t_k}. \]

Theorem

For all \(p \geq 2\), there exists a constant \(K_p\) such that

\[
\sup_{0 \leq t \leq T} \mathbb{E} \left[ \|X^f(t) - X^a(t)\|^p \right] \leq K_p \Delta t^{p/2},
\]

\[
\sup_{0 \leq t \leq T} \mathbb{E} \left[ \left\| \overline{X}^f(t) - X^c(t) \right\|^p \right] \leq K_p \Delta t^p,
\]

where \(\overline{X}^f(t)\) is the average of the piecewise linear interpolants \(X^f(t)\) and \(X^a(t)\).
Asian options

Hence

\[ \mathbb{E} \left[ \left\| X^f_{ave} - X^a_{ave} \right\|^p \right] \leq \sup_{[0,T]} \mathbb{E} \left[ \left\| X^f(t) - X^a(t) \right\|^p \right] , \]

and similarly

\[ \mathbb{E} \left[ \left\| \frac{1}{2} (X^f_{ave} + X^a_{ave}) - X^c_{ave} \right\|^p \right] \leq \sup_{[0,T]} \mathbb{E} \left[ \left\| X^f(t) - X^c(t) \right\|^p \right] , \]
Current Research

- Discontinuous payoffs
- Sub-sampling of Lévy areas
- Non-linear SDEs
- Stability analysis
Message to take home

- A new MLMC may not require simulations of Lévy areas.
- Classical Strong Convergence does not determine the variance of MLMC.