A multilevel Monte-Carlo theorem for stable numerical methods

Raphael Kruse
ETH Zürich

Joint work with M. Giles, A. Jentzen, L. Szpruch

Bielefeld University
6th Workshop on Random Dynamical Systems
November 2, 2013

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Abstract problem

Let H be a Hilbert space and $Y \in L_2(\Omega, \mathcal{F}, \mathbf{P}; H)$.

Aim: Compute $\mathbf{E}\big[Y\big]$.

Abstract problem

Let H be a Hilbert space and $Y \in L_2(\Omega, \mathcal{F}, \mathbf{P}; H)$.

Aim: Compute $\mathbf{E}\big[Y\big]$.

Examples:

- Parametric integration (Heinrich 2001): Let $\mathcal{D} \subset \mathbb{R}^d$ bounded domain, $H = L_2([0,1])$, $g\colon [0,1] \times \mathcal{D} \to \mathbb{R}$,

$$Y \equiv u(\lambda) = \int_{\mathcal{D}} g(\lambda, x)\,\mathrm{d}x.$$

- Option pricing (Giles 2008): $H = \mathbb{R}$, payoff function $\varphi\colon [0,\infty) \to \mathbb{R}$ and

$$Y = \varphi(X(T)),$$

where $X\colon [0,T] \times \Omega \to \mathbb{R}$ solves some SODE.

Standard Monte Carlo approach

Generate independent and identically distributed copies $(Y^m)_{m=1}^M$, $M \in \mathbb{N}$, of Y

$$\mathcal{MC}(M) := \frac{1}{M} \sum_{m=1}^{M} Y^m.$$

Standard Monte Carlo approach

Generate independent and identically distributed copies $(Y^m)_{m=1}^M$, $M \in \mathbb{N}$, of Y

$$\mathcal{MC}(M) := \frac{1}{M} \sum_{m=1}^{M} Y^m.$$

Error estimates:

$$\begin{aligned} \|\mathbf{E}[Y] - \mathcal{MC}(M)\|^2_{L_2(\Omega;H)} &= \frac{1}{M^2} \Big\| \sum_{m=1}^{M} \big(\mathbf{E}[Y] - Y^m\big) \Big\|^2_{L_2(\Omega;H)} \\ &= \frac{1}{M} \Big\| \mathbf{E}[Y] - Y^1 \Big\|^2_{L_2(\Omega;H)} = \frac{1}{M} \mathrm{Var}(Y). \end{aligned}$$

Standard Monte Carlo approach

Generate independent and identically distributed copies $(Y^m)_{m=1}^M$, $M \in \mathbb{N}$, of Y

$$\mathcal{MC}(M) := \frac{1}{M}\sum_{m=1}^{M} Y^m.$$

Error estimates:

$$\begin{aligned}\|\mathbf{E}[Y] - \mathcal{MC}(M)\|^2_{L_2(\Omega;H)} &= \frac{1}{M^2}\Big\|\sum_{m=1}^{M}\big(\mathbf{E}[Y] - Y^m\big)\Big\|^2_{L_2(\Omega;H)} \\ &= \frac{1}{M}\Big\|\mathbf{E}[Y] - Y^1\Big\|^2_{L_2(\Omega;H)} = \frac{1}{M}\mathrm{Var}(Y).\end{aligned}$$

Hence

$$\|\mathbf{E}[Y] - \mathcal{MC}(M)\|_{L_2(\Omega;H)} = \mathcal{O}(M^{-\frac{1}{2}}).$$

Trouble ahead

Direct generation of copies of Y often not possible:

- distribution of Y is unknown,
- H is of infinite dimensions,
- just too expensive…

Trouble ahead

Direct generation of copies of Y often not possible:

- distribution of Y is unknown,
- H is of infinite dimensions,
- just too expensive...

For this: Assume existence of a sequence $(Y_\ell)_{\ell\in\mathbb{N}} \subset L_2(\Omega; H)$ such that $\exists C_1, C_2, p_1, p_2 > 0$ with $p_1 \leq p_2$ and

$$\begin{aligned} \|Y_\ell - Y\|_{L_2(\Omega;H)} &\leq C_1 2^{-p_1\ell} \quad \text{(Strong conv)},\\ \big|\mathbf{E}[Y_\ell] - \mathbf{E}[Y]\big| &\leq C_2 2^{-p_2\ell} \quad \text{(Weak conv)} \end{aligned}$$

for all $\ell \in \mathbb{N}$.

Single level Monte Carlo

Instead of Y we use Y_L for some $L \in \mathbb{N}$:

$$\mathcal{MC}_1(M, L) := \frac{1}{M} \sum_{m=1}^{M} Y_L^m.$$

Single level Monte Carlo

Instead of Y we use Y_L for some $L \in \mathbb{N}$:

$$\mathcal{MC}_1(M, L) := \frac{1}{M} \sum_{m=1}^{M} Y_L^m.$$

Error representation:

$$\begin{aligned}
&\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|^2_{L_2(\Omega;H)} \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \|\mathbf{E}[Y_L] - \mathcal{MC}_1(M, L)\|^2_{L_2(\Omega;H)} \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \frac{1}{M}\mathrm{Var}(Y_L).
\end{aligned}$$

Single level Monte Carlo

Instead of Y we use Y_L for some $L \in \mathbb{N}$:

$$\mathcal{MC}_1(M, L) := \frac{1}{M} \sum_{m=1}^{M} Y_L^m.$$

Error representation:

$$\begin{aligned}
&\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)}^2 \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \|\mathbf{E}[Y_L] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)}^2 \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \frac{1}{M}\mathrm{Var}(Y_L).
\end{aligned}$$

Hence, by weak convergence

$$\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)} = \mathcal{O}\left(\sqrt{2^{-2p_2 L} + M^{-1}}\right).$$

Computational cost

For a given precision $\epsilon > 0$:

Level $L \in \mathbb{N}$ is determined by the weak error:

$$L := \left\lceil \frac{\log(\epsilon^{-1})}{\log(2) p_2} \right\rceil.$$

If L is large enough we may assume $\mathrm{Var}(Y_L) \approx \mathrm{Var}(Y)$, thus

$$M \geq \epsilon^{-2}$$

Monte Carlo samples are needed. Then

$$\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)} = \mathcal{O}\big(\epsilon\big).$$

Computational cost

For a given precision $\epsilon > 0$:

Level $L \in \mathbb{N}$ is determined by the weak error:

$$L := \left\lceil \frac{\log(\epsilon^{-1})}{\log(2) p_2} \right\rceil.$$

If L is large enough we may assume $\mathrm{Var}(Y_L) \approx \mathrm{Var}(Y)$, thus

$$M \geq \epsilon^{-2}$$

Monte Carlo samples are needed. Then

$$\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)} = \mathcal{O}\big(\epsilon\big).$$

Let $\tau(Z)$ be the computational cost necessary to simulate Z.

Computational cost

For a given precision $\epsilon > 0$:

Level $L \in \mathbb{N}$ is determined by the weak error:

$$L := \left\lceil \frac{\log(\epsilon^{-1})}{\log(2)p_2} \right\rceil.$$

If L is large enough we may assume $\mathrm{Var}(Y_L) \approx \mathrm{Var}(Y)$, thus

$$M \geq \epsilon^{-2}$$

Monte Carlo samples are needed. Then

$$\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)} = \mathcal{O}\big(\epsilon\big).$$

Let $\tau(Z)$ be the computational cost necessary to simulate Z. Assume that

$$\tau(Y_\ell) = 2^{(\ell-1)}, \quad \ell \in \mathbb{N}.$$

Computational cost

For a given precision $\epsilon > 0$:

Level $L \in \mathbb{N}$ is determined by the weak error:

$$L := \left\lceil \frac{\log(\epsilon^{-1})}{\log(2) p_2} \right\rceil.$$

If L is large enough we may assume $\mathrm{Var}(Y_L) \approx \mathrm{Var}(Y)$, thus

$$M \geq \epsilon^{-2}$$

Monte Carlo samples are needed. Then

$$\|\mathbf{E}[Y] - \mathcal{MC}_1(M, L)\|_{L_2(\Omega;H)} = \mathcal{O}\big(\epsilon\big).$$

Let $\tau(Z)$ be the computational cost necessary to simulate Z. Assume that

$$\tau(Y_\ell) = 2^{(\ell-1)}, \quad \ell \in \mathbb{N}.$$

Then

$$\tau(\mathcal{MC}_1(M, L)) = M 2^{L-1} \geq \frac{1}{2} \epsilon^{-(2+\frac{1}{p_2})}.$$

Multilevel Monte Carlo sampler

Idea: Use the telescopic sum

$$\mathbf{E}\left[Y_L\right] = \sum_{\ell=1}^{L} \mathbf{E}\left[Y_\ell - Y_{\ell-1}\right]$$

with $Y_0 = 0$.

Multilevel Monte Carlo sampler

Idea: Use the telescopic sum

$$\mathbf{E}\big[Y_L\big] = \sum_{\ell=1}^{L} \mathbf{E}\big[Y_\ell - Y_{\ell-1}\big]$$

with $Y_0 = 0$. Define

$$\Delta_\ell := Y_\ell - Y_{\ell-1}.$$

Multilevel Monte Carlo sampler

$$\mathcal{MLMC}(M_1, \dots, M_L, L) := \sum_{\ell=1}^{L} \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \Delta_\ell^m,$$

where Δ_ℓ^m, $m \in \mathbb{N}$, is an i.i.d family of copies of Δ_ℓ for all $\ell \in \mathbb{N}$.

Multilevel Monte Carlo – error representation

Error representation:

$$\begin{aligned}
&\|\mathbf{E}[Y] - \mathcal{MLMC}(M_1, \dots, M_L, L)\|^2_{L_2(\Omega;H)} \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \|\mathbf{E}[Y_L] - \mathcal{MLMC}(M_1, \dots, M_L, L)\|^2_{L_2(\Omega;H)} \\
&= \dots + \Big\| \sum_{\ell=1}^{L} \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \left(\mathbf{E}[\Delta_\ell] - \Delta_\ell^m\right) \Big\|^2_{L_2(\Omega;H)} \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \sum_{\ell=1}^{L} \frac{1}{M_\ell} \mathrm{Var}(\Delta_\ell).
\end{aligned}$$

Multilevel Monte Carlo – error representation

Error representation:

$$\begin{aligned}
&\|\mathbf{E}[Y] - \mathcal{MLMC}(M_1, \ldots, M_L, L)\|_{L_2(\Omega;H)}^2 \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \|\mathbf{E}[Y_L] - \mathcal{MLMC}(M_1, \ldots, M_L, L)\|_{L_2(\Omega;H)}^2 \\
&= \ldots + \left\| \sum_{\ell=1}^{L} \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \left(\mathbf{E}[\Delta_\ell] - \Delta_\ell^m\right) \right\|_{L_2(\Omega;H)}^2 \\
&= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \sum_{\ell=1}^{L} \frac{1}{M_\ell} \mathrm{Var}(\Delta_\ell).
\end{aligned}$$

By strong convergence

$$\begin{aligned}
\|Y_\ell - Y_{\ell-1}\|_{L_2(\Omega;H)} &\leq \|Y_\ell - Y\|_{L_2(\Omega;H)} + \|Y - Y_{\ell-1}\|_{L_2(\Omega;H)} \\
&\leq C_1(1 + 2^{-p_1})2^{-p_1\ell}
\end{aligned}$$

for all $\ell \in \mathbb{N}$. Thus,

$$\mathrm{Var}(Y_\ell - Y_{\ell-1}) \leq C_3 2^{-2p_1\ell}.$$

Multilevel Monte Carlo – parameter choice

$L \in \mathbb{N}$ is again determined by the weak error:

$$L \geq \frac{\log(\epsilon^{-1})}{p_2 \log(2)}.$$

Multilevel Monte Carlo – parameter choice

$L \in \mathbb{N}$ is again determined by the weak error:

$$L \geq \frac{\log(\epsilon^{-1})}{p_2 \log(2)}.$$

$M_1, \ldots, M_L \in \mathbb{N}$ are given by

$$M_\ell \geq \left\lceil \epsilon^{-2} 2^{-\frac{(2p_1+1)\ell}{2}} \sum_{k=1}^{L} 2^{\frac{(1-2p_1)k}{2}} \right\rceil$$

as a solution to the optimization problem

$$\min_{(M_1,\ldots,M_L)\in\mathbb{N}^L} \sum_{\ell=1}^{L} M_\ell \tau(\Delta_\ell) \Big(= \frac{3}{2} \sum_{\ell=1}^{L} M_\ell 2^{\ell-1} \Big),$$

$$\text{s/t} \qquad C_3 \sum_{\ell=1}^{L} \frac{1}{M_\ell} 2^{-2p_1\ell} \leq \epsilon^2.$$

Multilevel Monte Carlo – total computational cost

If $p_1 = \frac{1}{2}$:

$$C \sum_{\ell=1}^{L} M_\ell 2^{\ell-1} \geq C\epsilon^2 L^2 \geq C\epsilon^2 \log(\epsilon^{-1})^2.$$

Multilevel Monte Carlo – total computational cost

If $p_1 = \frac{1}{2}$:

$$C \sum_{\ell=1}^{L} M_\ell 2^{\ell-1} \geq C\epsilon^2 L^2 \geq C\epsilon^2 \log(\epsilon^{-1})^2.$$

If $p_1 > \frac{1}{2}$:

$$C \sum_{\ell=1}^{L} M_\ell 2^{\ell-1} \geq C\epsilon^2.$$

Reference: [Giles 2008].

Applications to SODEs

Let $X\colon [0,T]\times\Omega\to\mathbb{R}^d$ be solution to

$$\begin{aligned} \mathrm{d}X(t) &= b^0(X(t))\,\mathrm{d}t + \sum_{r=1}^{m} b^r(X(t))\,\mathrm{d}W^r(t), \\ X(0) &= X_0, \end{aligned} \tag{SODE}$$

where $b^r\colon \mathbb{R}^d\to\mathbb{R}^d$, $r\in\{0,1,\dots,m\}$, are sufficiently smooth.

Applications to SODEs

Let $X\colon [0,T]\times\Omega\to\mathbb{R}^d$ be solution to

$$\begin{aligned} \mathrm{d}X(t) &= b^0(X(t))\,\mathrm{d}t + \sum_{r=1}^{m} b^r(X(t))\,\mathrm{d}W^r(t), \\ X(0) &= X_0, \end{aligned} \tag{SODE}$$

where $b^r\colon\mathbb{R}^d\to\mathbb{R}^d$, $r\in\{0,1,\ldots,m\}$, are sufficiently smooth. Then, for some smooth function $\varphi\colon\mathbb{R}^d\to\mathbb{R}$ our aim is to approximate $\mathbf{E}[Y]$ with

$$Y := \varphi(X(T)).$$

Approximation of $X(T)$ by numerical schemes $X_\ell(T)$, $\ell\in\mathbb{N}$,

$$Y_\ell := \varphi(X_\ell(T)).$$

Numerical schemes

Temporal step sizes $h_\ell = 2^{-(\ell-1)}T$ and grid points $t_n^\ell := nh$,

Numerical schemes

Temporal step sizes $h_\ell = 2^{-(\ell-1)}T$ and grid points $t_n^\ell := nh$,
Euler-Maruyama method: $p_1 = \frac{1}{2}$, $p_2 = 1$

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + h_\ell b^0(X_\ell(t_{n-1}^\ell)) + \sum_{r=1}^{m} b^r(X_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell},$$

$$X_\ell(t_0) = X_0.$$

Numerical schemes

Temporal step sizes $h_\ell = 2^{-(\ell-1)}T$ and grid points $t_n^\ell := nh$,
Euler-Maruyama method: $p_1 = \frac{1}{2}$, $p_2 = 1$

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + h_\ell b^0(X_\ell(t_{n-1}^\ell)) + \sum_{r=1}^{m} b^r(X_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell},$$

$$X_\ell(t_0) = X_0.$$

Milstein scheme: $p_1 = 1$, $p_2 = 1$

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + h_\ell b^0(X_\ell(t_{n-1}^\ell)) + \sum_{r=1}^{m} b^r(X_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$

$$+ \sum_{r_1, r_2 = 1}^{m} \mathcal{L}^{r_1} b^{r_2}(X_\ell(t_{n-1}^\ell)) I_{(r_1, r_2)}^{t_n^\ell, t_{n-1}^\ell}$$

for $n \in \{1, \ldots, 2^{\ell-1}\}$.

Iterated stochastic integrals

Here

$$I_{(r)}^{t_n^\ell, t_{n-1}^\ell} := W^r(t_n^\ell) - W^r(t_{n-1}^\ell)$$

and

$$I_{(r_1,r_2)}^{t_n^\ell, t_{n-1}^\ell} := \int_{t_{n-1}^\ell}^{t_n^\ell} \int_{t_{n-1}^\ell}^{\sigma} \mathrm{d}W^{r_1}(\tau)\,\mathrm{d}W^{r_2}(\sigma)$$

for $r, r_1, r_2 \in \{1, \ldots, m\}$.

Iterated stochastic integrals

Here

$$I_{(r)}^{t_n^\ell, t_{n-1}^\ell} := W^r(t_n^\ell) - W^r(t_{n-1}^\ell)$$

and

$$I_{(r_1,r_2)}^{t_n^\ell, t_{n-1}^\ell} := \int_{t_{n-1}^\ell}^{t_n^\ell} \int_{t_{n-1}^\ell}^{\sigma} \mathrm{d}W^{r_1}(\tau)\,\mathrm{d}W^{r_2}(\sigma)$$

for $r, r_1, r_2 \in \{1, \ldots, m\}$.

New problem: $I_{(r_1,r_2)}^{t_n^\ell, t_{n-1}^\ell}$ are not easily simulated.

Truncated Milstein scheme

Giles, Szpruch 2012: Consider truncated Milstein:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + h_\ell b^0(X_\ell(t_{n-1}^\ell)) + \sum_{r=1}^{m} b^r(X_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$
$$+ \sum_{r_1,r_2=1}^{m} \mathcal{L}^{r_1} b^{r_2}(X_\ell(t_{n-1}^\ell)) \frac{1}{2}\left(I_{(r_1)}^{t_n^\ell, t_{n-1}^\ell} I_{(r_2)}^{t_n^\ell, t_{n-1}^\ell} - \delta_{r_1,r_2} h_\ell \right)$$

for $n \in \{1, \ldots, 2^{\ell-1}\}$.

Truncated Milstein scheme

Giles, Szpruch 2012: Consider truncated Milstein:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + h_\ell b^0(X_\ell(t_{n-1}^\ell)) + \sum_{r=1}^{m} b^r(X_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$
$$+ \sum_{r_1,r_2=1}^{m} \mathcal{L}^{r_1} b^{r_2}(X_\ell(t_{n-1}^\ell)) \frac{1}{2}\left(I_{(r_1)}^{t_n^\ell, t_{n-1}^\ell} I_{(r_2)}^{t_n^\ell, t_{n-1}^\ell} - \delta_{r_1,r_2} h_\ell\right)$$

for $n \in \{1, \ldots, 2^{\ell-1}\}$.

The truncated Milstein has strong order $p_1 = \frac{1}{2}$.

Truncated Milstein scheme

Giles, Szpruch 2012: Consider truncated Milstein:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + h_\ell b^0(X_\ell(t_{n-1}^\ell)) + \sum_{r=1}^{m} b^r(X_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$

$$+ \sum_{r_1, r_2 = 1}^{m} \mathcal{L}^{r_1} b^{r_2}(X_\ell(t_{n-1}^\ell)) \tfrac{1}{2}\left(I_{(r_1)}^{t_n^\ell, t_{n-1}^\ell} I_{(r_2)}^{t_n^\ell, t_{n-1}^\ell} - \delta_{r_1, r_2} h_\ell\right)$$

for $n \in \{1, \ldots, 2^{\ell-1}\}$.

The truncated Milstein has strong order $p_1 = \frac{1}{2}$.

Motivation: Recall the relationship

$$I_{(r_1, r_2)}^{t_n^\ell, t_{n-1}^\ell} + I_{(r_2, r_1)}^{t_n^\ell, t_{n-1}^\ell} = I_{(r_1)}^{t_n^\ell, t_{n-1}^\ell} I_{(r_2)}^{t_n^\ell, t_{n-1}^\ell} - \delta_{r_1, r_2} h_\ell.$$

Idea of antithetic MLMC

Use again the telescopic sum

$$\mathbf{E}\big[Y_L\big] = \sum_{\ell=1}^{L} \mathbf{E}\big[Y_\ell - Y_{\ell-1}\big]$$
$$= \sum_{\ell=1}^{L} \big(\mathbf{E}\big[Y_\ell - \overline{Y}_\ell\big] + \mathbf{E}\big[\overline{Y}_\ell - Y_{\ell-1}\big]\big)$$

with $Y_0 = 0$ and

$$\overline{Y}_\ell := \frac{1}{2}\big(\varphi(X_\ell) + \varphi(X_\ell^a)\big)$$

where X_ℓ is generated by the truncated Milstein scheme and X_ℓ^a is its antithetic twin.

Definition of the antithetic twin

Write truncated Milstein scheme in terms of an increment function:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + \Phi_1\big(X_\ell(t_{n-1}^\ell), h_\ell, (I_{(r)}^{t_n, t_{n-1}})_{r=1}^m\big).$$

Similarly, this can be done for two consecutive steps:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-2}^\ell) + \Phi_2\big(X_\ell(t_{n-2}^\ell), h_\ell, (I_{(r)}^{t_{n-1}^\ell, t_{n-2}^\ell})_{r=1}^m, (I_{(r)}^{t_n^\ell, t_{n-1}^\ell})_{r=1}^m\big).$$

Definition of the antithetic twin

Write truncated Milstein scheme in terms of an increment function:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-1}^\ell) + \Phi_1\big(X_\ell(t_{n-1}^\ell), h_\ell, (I_{(r)}^{t_n,t_{n-1}})_{r=1}^m\big).$$

Similarly, this can be done for two consecutive steps:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-2}^\ell) + \Phi_2\big(X_\ell(t_{n-2}^\ell), h_\ell, (I_{(r)}^{t_{n-1}^\ell,t_{n-2}^\ell})_{r=1}^m, (I_{(r)}^{t_n^\ell,t_{n-1}^\ell})_{r=1}^m\big).$$

Antithetic twin is given by interchanging the role of the stochastic increments:

$$X_\ell^a(t_n^\ell) = X_\ell^a(t_{n-2}^\ell) + \Phi_2\big(X_\ell^a(t_{n-2}^\ell), h_\ell, (I_{(r)}^{t_n^\ell,t_{n-1}^\ell})_{r=1}^m, (I_{(r)}^{t_{n-1}^\ell,t_{n-2}^\ell})_{r=1}^m\big).$$

Antithetic MLMC sampler

Recall the telescopic sum

$$\mathbf{E}\left[Y_L\right] = \sum_{\ell=1}^{L} \Big(\underbrace{\mathbf{E}\left[Y_\ell - \overline{Y}_\ell\right]}_{=0} + \mathbf{E}\left[\overline{Y}_\ell - Y_{\ell-1}\right] \Big).$$

Define the antithetic MLMC sampler by

$$\mathcal{MLMC}_a(L, M_1, \ldots, M_L) := \sum_{\ell=1}^{L} \frac{1}{M_\ell} \sum_{m=1}^{M_\ell} \overline{\Delta}_\ell^m,$$

where

$$\overline{\Delta}_\ell := \overline{Y}_\ell - Y_{\ell-1}.$$

Antithetic MLMC sampler – error representation

Error representation:

$$\|\mathbf{E}[Y] - \mathcal{MLMC}_1(M_1, \dots, M_L, L)\|^2_{L_2(\Omega;H)}$$

$$= \left|\mathbf{E}[Y] - \mathbf{E}[Y_L]\right|^2 + \sum_{\ell=1}^{L} \frac{1}{M_\ell} \mathrm{Var}(\overline{\Delta}_\ell).$$

Theorem (Giles, Szpruch 2012)

Let $\varphi \in C^2(\mathbb{R}^d, \mathbb{R})$ satisfy some polynomial growth conditions, let $b^r : \mathbb{R}^d \to \mathbb{R}^d$, $r \in \{1, \dots, m\}$, be sufficiently smooth. Then

$$\mathrm{Var}(\overline{\Delta}_\ell) \leq C 2^{2\ell}$$

for all $\ell \in \mathbb{N}$.

Thus, antithetic MLMC with truncated Milstein behaves in the same way as MLMC with full Milstein.

First step of the proof

Lemma (Giles, Szpruch 2012)

Let $\varphi \in C^2(\mathbb{R}^d, \mathbb{R})$ satisfy some polynomial growth conditions. Then

$$\mathrm{Var}(\overline{\Delta}_\ell) \leq C\Big(\Big\|\frac{1}{2}(X_\ell(T) + X_\ell^a(T)) - X_{\ell-1}(T)\Big\|^2_{L_p(\Omega,\mathbb{R})} + \Big\|X_\ell(T) - X_\ell^a(T)\Big\|^4_{L_p(\Omega,\mathbb{R})}\Big)$$

for all $\ell \in \mathbb{N}$.

We concentrate on first summand.

A stability concept for numerical schemes

Let $\mathcal{G}_\ell$ be space of grid functions $Z_\ell \colon \{t_0^\ell, \ldots, t_{2^\ell - 1}^\ell\} \to L_2(\Omega; \mathbb{R}^d)$.

A stability concept for numerical schemes

Let $\mathcal{G}_\ell$ be space of grid functions $Z_\ell \colon \{t_0^\ell, \dots, t_{2^\ell-1}^\ell\} \to L_2(\Omega; \mathbb{R}^d)$.

Write numerical scheme in terms of a residual operator $\mathcal{R}_\ell \colon \mathcal{G}_\ell \to \mathcal{G}_\ell$, such that $\mathcal{R}_\ell[X_\ell] = 0 \in \mathcal{G}_\ell$:

$$\mathcal{R}_\ell[Z_\ell](t_0^\ell) = Z_\ell(t_0^\ell) - X_0,$$

$$\mathcal{R}_\ell[Z_\ell](t_n^\ell) = Z_\ell(t_n^\ell) - Z_\ell(t_{n-1}^\ell) - h_\ell b^0(Z_\ell(t_{n-1}^\ell)) - \sum_{r=1}^{m} b^r(Z_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$

A stability concept for numerical schemes

Let $\mathcal{G}_\ell$ be space of grid functions $Z_\ell \colon \{t_0^\ell, \dots, t_{2^\ell - 1}^\ell\} \to L_2(\Omega; \mathbb{R}^d)$.

Write numerical scheme in terms of a residual operator $\mathcal{R}_\ell \colon \mathcal{G}_\ell \to \mathcal{G}_\ell$, such that $\mathcal{R}_\ell[X_\ell] = 0 \in \mathcal{G}_\ell$:

$$\mathcal{R}_\ell[Z_\ell](t_0^\ell) = Z_\ell(t_0^\ell) - X_0,$$

$$\mathcal{R}_\ell[Z_\ell](t_n^\ell) = Z_\ell(t_n^\ell) - Z_\ell(t_{n-1}^\ell) - h_\ell b^0(Z_\ell(t_{n-1}^\ell)) - \sum_{r=1}^{m} b^r(Z_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$

Endow $\mathcal{G}_\ell$ with the two norms

$$\|Z_\ell\|_{0,\ell} = \max_{0 \le i \le 2^\ell - 1} \|Z_\ell(t_i^\ell)\|_{L_2(\Omega; \mathbb{R}^d)}.$$

A stability concept for numerical schemes

Let $\mathcal{G}_\ell$ be space of grid functions $Z_\ell \colon \{t_0^\ell, \dots, t_{2^\ell-1}^\ell\} \to L_2(\Omega;\mathbb{R}^d)$.

Write numerical scheme in terms of a residual operator $\mathcal{R}_\ell \colon \mathcal{G}_\ell \to \mathcal{G}_\ell$, such that $\mathcal{R}_\ell[X_\ell] = 0 \in \mathcal{G}_\ell$:

$$\mathcal{R}_\ell[Z_\ell](t_0^\ell) = Z_\ell(t_0^\ell) - X_0,$$

$$\mathcal{R}_\ell[Z_\ell](t_n^\ell) = Z_\ell(t_n^\ell) - Z_\ell(t_{n-1}^\ell) - h_\ell b^0(Z_\ell(t_{n-1}^\ell)) - \sum_{r=1}^{m} b^r(Z_\ell(t_{n-1}^\ell)) I_{(r)}^{t_n^\ell, t_{n-1}^\ell}$$

Endow $\mathcal{G}_\ell$ with the two norms

$$\|Z_\ell\|_{0,\ell} = \max_{0 \le i \le 2^\ell - 1} \|Z_\ell(t_i^\ell)\|_{L_2(\Omega;\mathbb{R}^d)}.$$

and with the stochastic Spijker-norm

$$\|Z_\ell\|_{-1,\ell} = \|Y_\ell(t_0^\ell)\|_{L_2(\Omega;\mathbb{R}^d)} + \max_{1 \le n \le 2^\ell - 1} \left\| \sum_{i=1}^{n} Z_\ell(t_i^\ell) \right\|_{L_2(\Omega;\mathbb{R}^d)}.$$

Bistability

Definition

A numerical scheme is called bistable, if there exist constants C_1, C_2 such that

$$\begin{aligned} C_1 \|\mathcal{R}_\ell[Z_\ell] - \mathcal{R}_\ell[\tilde{Z}_\ell]\|_{-1,\ell} \\ \leq \|Z_\ell - \tilde{Z}_\ell\|_{0,\ell} \\ \leq C_2 \|\mathcal{R}_\ell[Z_\ell] - \mathcal{R}_\ell[\tilde{Z}_\ell]\|_{-1,\ell} \end{aligned}$$

holds for all $\ell \in \mathbb{N}$ and $Z_\ell, \tilde{Z}_\ell \in \mathcal{G}_\ell$.

Consistency

Definition

A numerical scheme is called consistent of order $\gamma_1 > 0$, if there exists a constant C_1 such that

$$\left\| \mathcal{R}_\ell[X|_{\mathcal{G}_\ell}] \right\|_{-1,\ell} \leq C_1 2^{-\gamma_1 \ell}$$

for all $\ell \in \mathbb{N}$.

Consistency

Definition

A numerical scheme is called consistent of order $\gamma_1 > 0$, if there exists a constant C_1 such that

$$\left\|\mathcal{R}_\ell[X|_{\mathcal{G}_\ell}]\right\|_{-1,\ell} \leq C_1 2^{-\gamma_1 \ell}$$

for all $\ell \in \mathbb{N}$.

A sequence of grid functions $(Z_\ell)_{\ell\in\mathbb{N}}$ is called level-consistent of order $\gamma_2 > 0$, if there exists a constant C_2 such that

$$\left\|\mathcal{R}_\ell[Z_\ell]\right\|_{-1,\ell} \leq C_2 2^{-\gamma_2 \ell}$$

for all $\ell \in \mathbb{N}$.

Apply stability concept to antithetic sampler

$$\left\| \frac{1}{2}(X_\ell(T) + X_\ell^a(T)) - X_{\ell-1}(T) \right\|_{L_p(\Omega,\mathbb{R})}$$

$$\leq C_2 \left\| \mathcal{R}_{\ell-1}\left[\frac{1}{2}(X_\ell + X_\ell^a)\right] \right\|_{-1,\ell-1} \leq \dots \leq C2^{-\ell}.$$

Apply stability concept to antithetic sampler

$$
\begin{aligned}
&\left\| \frac{1}{2}(X_\ell(T) + X_\ell^a(T)) - X_{\ell-1}(T) \right\|_{L_p(\Omega,\mathbb{R})} \\
&\quad \leq C_2 \left\| \mathcal{R}_{\ell-1}\left[\frac{1}{2}(X_\ell + X_\ell^a)\right] \right\|_{-1,\ell-1} \leq \ldots \leq C 2^{-\ell}.
\end{aligned}
$$

Clark-Cameron Example

$$
\begin{aligned}
\mathrm{d}X(t) &= \begin{pmatrix} 1 & 0 \\ 0 & X_1(t) \end{pmatrix} \mathrm{d}\begin{pmatrix} W^1(t) \\ W^2(t) \end{pmatrix} \\
X(0) &= 0 \in \mathbb{R}^2.
\end{aligned}
$$

Truncated Milstein for Clark-Cameron

One step of truncated Milstein:

$$X_{\ell-1}(t_n^{\ell-1}) = X_{\ell-1}(t_{n-1}^{\ell-1}) + \begin{pmatrix} I_{(1)}^{t_n^{\ell-1}, t_{n-1}^{\ell-1}} \\ X_{\ell-1}^1(t_{n-1}^{\ell-1}) I_{(2)}^{t_n^{\ell-1}, t_{n-1}^{\ell-1}} + \frac{1}{2} I_{(1)}^{t_n^{\ell-1}, t_{n-1}^{\ell-1}} I_{(2)}^{t_n^{\ell-1}, t_{n-1}^{\ell-1}} \end{pmatrix}.$$

Two steps of truncated Milstein:

$$X_\ell(t_n^\ell) = X_\ell(t_{n-2}^\ell) + \begin{pmatrix} I_{(1)}^{t_{n-1}^\ell, t_{n-2}^\ell} + I_{(1)}^{t_n^\ell, t_{n-1}^\ell} \\ X_\ell^1(t_{n-2}^\ell)\big(I_{(2)}^{t_{n-1}^\ell, t_{n-2}^\ell} + I_{(2)}^{t_n^\ell, t_{n-1}^\ell}\big) \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ I_{(1)}^{t_{n-1}^\ell, t_{n-2}^\ell} I_{(2)}^{t_n^\ell, t_{n-1}^\ell} + \frac{1}{2}\big(I_{(1)}^{t_{n-1}^\ell, t_{n-2}^\ell} I_{(2)}^{t_{n-1}^\ell, t_{n-2}^\ell} + I_{(1)}^{t_{n-1}^\ell, t_{n-2}^\ell} I_{(2)}^{t_{n-1}^\ell, t_{n-2}^\ell}\big) \end{pmatrix}.$$

Truncated Milstein for Clark-Cameron

One step of truncated Milstein:

$$X_{\ell-1}(t_n^{\ell-1}) = X_{\ell-1}(t_{n-1}^{\ell-1}) + \begin{pmatrix} I_{(1)}^{t_n^{\ell-1},t_{n-1}^{\ell-1}} \\ X_{\ell-1}^1(t_{n-1}^{\ell-1}) I_{(2)}^{t_n^{\ell-1},t_{n-1}^{\ell-1}} + \frac{1}{2} I_{(1)}^{t_n^{\ell-1},t_{n-1}^{\ell-1}} I_{(2)}^{t_n^{\ell-1},t_{n-1}^{\ell-1}} \end{pmatrix}.$$

Two steps of truncated Milstein:

$$\begin{aligned} X_\ell(t_n^\ell) = X_\ell(t_{n-2}^\ell) &+ \begin{pmatrix} I_{(1)}^{t_{n-1}^\ell,t_{n-2}^\ell} + I_{(1)}^{t_n^\ell,t_{n-1}^\ell} \\ X_\ell^1(t_{n-2}^\ell)\big(I_{(2)}^{t_{n-1}^\ell,t_{n-2}^\ell} + I_{(2)}^{t_n^\ell,t_{n-1}^\ell}\big) \end{pmatrix} \\ &+ \begin{pmatrix} 0 \\ I_{(1)}^{t_{n-1}^\ell,t_{n-2}^\ell} I_{(2)}^{t_n^\ell,t_{n-1}^\ell} + \frac{1}{2}\big(I_{(1)}^{t_{n-1}^\ell,t_{n-2}^\ell} I_{(2)}^{t_{n-1}^\ell,t_{n-2}^\ell} + I_{(1)}^{t_{n-1}^\ell,t_{n-2}^\ell} I_{(2)}^{t_{n-1}^\ell,t_{n-2}^\ell}\big) \end{pmatrix}. \end{aligned}$$

It directly follows

$$\left\| \mathcal{R}_{\ell-1}\left[\frac{1}{2}(X_\ell + X_\ell^a)\right] \right\|_{-1,\ell-1} = 0.$$

References

Multilevel Monte Carlo

- Heinrich, LSSC Proceeding, 2001.
- Giles, Operations Research, 2008.

Numerical Schemes for SODEs

- Kloeden, Platen, Springer, 1992.

Stability concept

- Beyn, RK, DCDS B, 2010.
- RK, BIT, 2012.

Antithetic MLMC approach

- Giles, Szpruch, MC & QMC Methods, 2012, and arXiv:1204.1647.