Antithetic Milstein scheme for SPDEs

AI Haji-Ali
Joint work with Andreas Stein (ETH)

Heriot-Watt University

Stochastic Numerics and Statistical Learning: Theory and Applications Workshop
28 May, 2024

Monte Carlo

Our goal is to compute

$$\mathrm{E}[\,\Psi(X)\,]$$

for some stochastic model $X \approx \hat{X}^{(\ell)}$, for $\ell \in \mathbb{N}$, which can only be sampled approximately.

Monte Carlo

Our goal is to compute

$$\mathrm{E}[\,\Psi(X)\,]$$

for some stochastic model $X \approx \hat{X}^{(\ell)}$, for $\ell \in \mathbb{N}$, which can only be sampled approximately. Assume for $w, \gamma > 0$ that

$$\begin{aligned} |\mathrm{E}[\,\Psi(X) - \Psi(\hat{X}^{(\ell)})\,]| &= \mathcal{O}(2^{-w\ell}) \\ \mathrm{Cost}(\hat{X}^{(\ell)}) &= \mathcal{O}(2^{\gamma\ell}) \end{aligned}$$

Then a Monte Carlo estimator for a fixed $L \in \mathbb{N}$ with $\mathfrak{N}$ samples has cost $\mathcal{O}(\mathfrak{N}\, 2^{\gamma L})$, bias $\mathcal{O}(2^{-wL})$, statistical error $\mathcal{O}(\mathfrak{N}^{-1/2})$. Complexity is $\mathcal{O}(\varepsilon^{-2-\gamma/w})$ for an RMSE ε.

Multilevel Monte Carlo

Assume we have an estimator $\Delta^{(\ell)}\hat{Y}$ such that

$$\mathrm{E}[\,\Delta^{(\ell)}\hat{Y}\,] = \begin{cases} \mathrm{E}[\,\Psi(X^{(0)})\,] & \ell = 0 \\ \mathrm{E}[\,\Psi(\hat{X}^{(\ell)}) - \Psi(\hat{X}^{(\ell-1)})\,] & \text{otherwise} \end{cases}$$

with $\mathrm{Cost}(\Delta^{(\ell)}\hat{Y}) = \mathcal{O}(2^{\gamma\ell})$ and

$$\mathrm{Var}[\,\Delta^{(\ell)}\hat{Y}\,] = \mathcal{O}(2^{-2s\ell})$$

Then, write

$$\mathrm{E}[\,\Psi(X)\,] = \sum_{\ell=0}^{\infty} \mathrm{E}[\,\Delta^{(\ell)}\hat{Y}\,] \approx \sum_{\ell=0}^{L} \frac{1}{\mathfrak{N}_\ell} \sum_{n=1}^{\mathfrak{N}_\ell} \Delta^{(\ell,n)}\hat{Y}$$

Then for same choice of L, and optimal choices of $\{\mathfrak{N}_\ell\}_{\ell=0}^{L}$, the complexity can be shown to be

$$\begin{cases} \varepsilon^{-2} & 2s > \gamma \\ \varepsilon^{-2}|\log \varepsilon|^2 & 2s = \gamma \\ \varepsilon^{-2-\frac{2s-\gamma}{w}} & 2s > \gamma \end{cases}$$

SDEs and Euler-Maruyama

Assume

$$dX_t = a(X_t)\,\mathrm{d}t + \sum_{i=1}^{d'} b_i(X_t)\,\mathrm{d}W_t^i$$

for $a, b_i : \mathbb{R}^d \to \mathbb{R}^d$ and d' independent Wiener processes $(W^i)_{i=1}^{d'}$, and use Euler-Maruyama with Δt_ℓ^{-1} time-steps for the approximation

$$\hat{X}^{(\ell)}_{(m+1)\Delta t_\ell} = \hat{X}^{(\ell)}_{m\Delta t_\ell} + a\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta t_\ell + \sum_{i=1}^{d'} b_i\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta_m W^i$$

where $\Delta_m W^i = (W^i_{(m+1)\Delta t_\ell} - W^i_{m\Delta t_\ell})$ and set $\Delta^{(\ell)}\hat{Y} = \Psi(\hat{X}_1^{(\ell)}) - \Psi(\hat{X}_1^{(\ell-1)})$, then for Lipschitz Ψ,

$$|\mathrm{E}[\,\Psi(X_1) - \Psi(\hat{X}_1^{(\ell)})\,]| = \mathcal{O}(\Delta t_\ell)$$

$$\mathrm{Cost}(\Psi(\hat{X}_1^{(\ell)})) = \mathcal{O}(\Delta t_\ell^{-1})$$

$$\mathrm{Var}[\,\Delta^{(\ell)}\hat{Y}\,] = \mathcal{O}(\Delta t_\ell)$$

Hence complexity is $\mathcal{O}(\varepsilon^{-2}|\log\varepsilon|^2)$.

SDE and Milstein

Use Milstein scheme instead with Δt_ℓ^{-1} time-steps for the approximation

$$\hat{X}^{(\ell)}_{(m+1)\Delta t_\ell} = \hat{X}^{(\ell)}_{m\Delta t_\ell} + a\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta t_\ell + \sum_{i=1}^{d'} b_i\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta_m W^i$$
$$+ \frac{1}{2}\sum_{i=1}^{d'}\sum_{j=1}^{d'} J_{b_j}\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right) b_i(\hat{X}^{(\ell)}_{m\Delta t_\ell})\left(\Delta_m W^i \Delta_m W^j - \delta_{i,j}\Delta t_\ell - A^{ij}_m\right)$$

where the Lévy area is defined as

$$A^{ij}_m = \int_{m\Delta t_\ell}^{(m+1)\Delta t_\ell} \int_{m\Delta t_\ell}^{s_1} \mathrm{d}W^j_{s_2}\,\mathrm{d}W^i_{s_1} - \mathrm{d}W^i_{s_2}\,\mathrm{d}W^j_{s_1}$$

SDE and Milstein

Again, set $\Delta^{(\ell)}\hat{Y} = \Psi(\hat{X}_1^{(\ell)}) - \Psi(\hat{X}_1^{(\ell-1)})$, then for Lipschitz Ψ,

$$\begin{aligned} |\mathrm{E}[\,\Psi(X_1) - \Psi(\hat{X}_1^{(\ell)})\,]| &= \mathcal{O}(\Delta t_\ell) \\ \mathrm{Cost}(\Psi(\hat{X}_1^{(\ell)})) &= \mathcal{O}(\Delta t_\ell^{-1} + (d'-1)(d'-2)\Delta t_\ell^{-2}) \\ \mathrm{Var}[\,\Delta^{(\ell)}\hat{Y}\,] &= \mathcal{O}(\Delta t_\ell^2) \end{aligned}$$

Hence complexity is $\mathcal{O}(\varepsilon^{-2})$ when $d' \in \{1, 2\}$ otherwise $\mathcal{O}(\varepsilon^{-2}|\log \varepsilon|^2)$.
We also get $\mathcal{O}(\varepsilon^{-2})$ when the noise is commutative

$$J_{b_j}(x)b_i(x) = J_{b_i}(x)b_j(x)$$

since $A_m^{ij} = -A_m^{ji}$.

Antithetic Milstein

(Giles & Szpruch, 2014) proposed the following scheme which depends on Brownian increments only

$$\hat{X}^{(\ell)}_{(m+1)\Delta t_\ell} = \hat{X}^{(\ell)}_{m\Delta t_\ell} + a\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta t_\ell + \sum_{i=1}^{d'} b_i\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta_m W^i$$

$$+ \frac{1}{2}\sum_{i=1}^{d'}\sum_{j=1}^{d'} J_{b_j}\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right) b_i(\hat{X}^{(\ell)}_{m\Delta t_\ell})\left(\Delta_m W^i \Delta_m W^j - \delta_{i,j}\Delta t_\ell\right)$$

and a corresponding antithetic scheme, $\hat{X}^{(\mathrm{a},\ell)}$, that swaps the Brownian increments between each two successive time steps.

Antithetic Milstein

(Giles & Szpruch, 2014) proposed the following scheme which depends on Brownian increments only

$$\hat{X}^{(\ell)}_{(m+1)\Delta t_\ell} = \hat{X}^{(\ell)}_{m\Delta t_\ell} + a\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta t_\ell + \sum_{i=1}^{d'} b_i\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right)\Delta_m W^i$$
$$+ \frac{1}{2}\sum_{i=1}^{d'}\sum_{j=1}^{d'} J_{b_j}\left(\hat{X}^{(\ell)}_{m\Delta t_\ell}\right) b_i(\hat{X}^{(\ell)}_{m\Delta t_\ell})\big(\Delta_m W^i \Delta_m W^j - \delta_{i,j}\Delta t_\ell\big)$$

and a corresponding antithetic scheme, $\hat{X}^{(\mathrm{a},\ell)}$, that swaps the Brownian increments between each two successive time steps.
Then they noted that for $\Delta^{(\ell)}\hat{Y} \equiv \frac{1}{2}\Big(\Psi(\hat{X}^{(\ell)}_1) + \Psi(\hat{X}^{(\mathrm{a},\ell)}_1)\Big) - \Psi(\hat{X}^{(\ell-1)}_1)$, and a (relaxable) twice-differentiable Ψ, the variance convergence is improved to

$$\mathrm{Var}[\,\Delta^{(\ell)}\hat{Y}\,] = \mathcal{O}(\Delta t_\ell^2).$$

Stochastic Partial Differential Equations

- Let H be a separable Hilbert space (e.g. $H = L^2(\mathcal{D})$), $T > 0$ and consider the H-valued Itô-SDE

$$dX(t) = [AX(t) + F(X(t))]dt + G(X(t))dW(t), \quad t \in [0, T],$$

(Financial modeling, stochastic epidemic models, stochastic forcing in heat transfer, filtering etc.)

Stochastic Partial Differential Equations

- Let H be a separable Hilbert space (e.g. $H = L^2(\mathcal{D})$), $T > 0$ and consider the H-valued Itô-SDE

$$dX(t) = [AX(t) + F(X(t))]dt + G(X(t))dW(t), \quad t \in [0, T],$$

  (Financial modeling, stochastic epidemic models, stochastic forcing in heat transfer, filtering etc.)
- Aim: Estimate $\mathrm{E}[\,\Psi(X(T))\,]$ by MLMC methods, where Ψ is a given (deterministic) functional.

Stochastic Partial Differential Equations

- Let H be a separable Hilbert space (e.g. $H = L^2(\mathcal{D})$), $T > 0$ and consider the H-valued Itô-SDE

$$dX(t) = [AX(t) + F(X(t))]dt + G(X(t))dW(t), \quad t \in [0, T],$$

(Financial modeling, stochastic epidemic models, stochastic forcing in heat transfer, filtering etc.)
- Aim: Estimate $\mathrm{E}[\,\Psi(X(T))\,]$ by MLMC methods, where Ψ is a given (deterministic) functional.
- Potential issues:
    - low temporal and spatial regularity
    - variance of the corrections in a MLMC Euler scheme decays slowly with order $\mathcal{O}(\Delta t)$.
    - Milstein schemes yield faster decay, but require the simulation of iterated integrals.

Stochastic Partial Differential Equations

- Let H be a separable Hilbert space (e.g. $H = L^2(\mathcal{D})$), $T > 0$ and consider the H-valued Itô-SDE

$$dX(t) = [AX(t) + F(X(t))]dt + G(X(t))dW(t), \quad t \in [0, T],$$

  (Financial modeling, stochastic epidemic models, stochastic forcing in heat transfer, filtering etc.)
- Aim: Estimate $\mathrm{E}[\,\Psi(X(T))\,]$ by MLMC methods, where Ψ is a given (deterministic) functional.
- Potential issues:
  - low temporal and spatial regularity
  - variance of the corrections in a MLMC Euler scheme decays slowly with order $\mathcal{O}(\Delta t)$.
  - Milstein schemes yield faster decay, but require the simulation of iterated integrals.
- Extend the antithetic coupling within a truncated Milstein scheme, as proposed by Giles and Szpruch for finite-dim, to the infinite dimensional setting.

Model problem & Assumptions

$$dX(t) = [AX(t)+F(X(t))]dt+G(X(t))dW(t), \quad t \in [0, T], \quad X(0) = X_0. \tag{SPDE}$$

Model problem & Assumptions

$$dX(t) = [AX(t)+F(X(t))]dt+G(X(t))dW(t), \quad t \in [0, T], \quad X(0) = X_0. \tag{SPDE}$$

- $A : D(A) \subset H \to H$ is a densely defined, self-adjoint, linear operator. Further, A generates an analytic semigroup $(S(t) = e^{At}, t \geq 0) \subset \mathcal{L}(H)$ and is boundedly invertible. (e.g. $A := \triangle$)

Model problem & Assumptions

$$dX(t) = [AX(t)+F(X(t))]dt+G(X(t))dW(t), \quad t \in [0, T], \quad X(0) = X_0. \tag{SPDE}$$

- $A : D(A) \subset H \to H$ is a densely defined, self-adjoint, linear operator. Further, A generates an analytic semigroup $(S(t) = e^{At}, t \geq 0) \subset \mathcal{L}(H)$ and is boundedly invertible. (e.g. $A := \triangle$)
- $F : H \to H$ is a (Lipschitz) non-linearity.

Model problem & Assumptions

$$dX(t) = [AX(t)+F(X(t))]dt+G(X(t))dW(t), \quad t \in [0, T], \quad X(0) = X_0. \tag{SPDE}$$

- $A : D(A) \subset H \to H$ is a densely defined, self-adjoint, linear operator. Further, A generates an analytic semigroup $(S(t) = e^{At}, t \geq 0) \subset \mathcal{L}(H)$ and is boundedly invertible. (e.g. $A := \triangle$)
- $F : H \to H$ is a (Lipschitz) non-linearity.
- $W : \Omega \times [0, T] \to H$ is a Q-Wiener process with trace class covariance operator $Q \in \mathcal{L}_1^+(H)$.

Model problem & Assumptions

$$dX(t) = [AX(t)+F(X(t))]dt+G(X(t))dW(t), \quad t \in [0,T], \quad X(0) = X_0. \tag{SPDE}$$

- $A : D(A) \subset H \to H$ is a densely defined, self-adjoint, linear operator. Further, A generates an analytic semigroup $(S(t) = e^{At}, t \geq 0) \subset \mathcal{L}(H)$ and is boundedly invertible. (e.g. $A := \triangle$)
- $F : H \to H$ is a (Lipschitz) non-linearity.
- $W : \Omega \times [0,T] \to H$ is a Q-Wiener process with trace class covariance operator $Q \in \mathcal{L}_1^+(H)$.
- $G : H \to \mathcal{L}_{HS}(\mathcal{H}; H)$ is Lipschitz, where $\mathcal{H} := Q^{1/2}H$ is the RKHS associated to Q.

Model problem & Assumptions

$$dX(t) = [AX(t)+F(X(t))]dt+G(X(t))dW(t), \quad t \in [0,T], \quad X(0) = X_0. \tag{SPDE}$$

- $A : D(A) \subset H \to H$ is a densely defined, self-adjoint, linear operator. Further, A generates an analytic semigroup $(S(t) = e^{At}, t \geq 0) \subset \mathcal{L}(H)$ and is boundedly invertible. (e.g. $A := \triangle$)
- $F : H \to H$ is a (Lipschitz) non-linearity.
- $W : \Omega \times [0,T] \to H$ is a Q-Wiener process with trace class covariance operator $Q \in \mathcal{L}_1^+(H)$.
- $G : H \to \mathcal{L}_{HS}(\mathcal{H}; H)$ is Lipschitz, where $\mathcal{H} := Q^{1/2}H$ is the RKHS associated to Q.
- $X_0 \in L^2(\Omega; H)$.

Regularity of mild solutions

There is a unique mild solution $X : \Omega \times [0, T] \to H$ to (SPDE), given by

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)G(X(s))dW(s),$$

for $t \in [0, T]$. Under mild assumptions: $X(t) \in L^p(\Omega; \dot{H}^\alpha)$ for some $\alpha > 0$, $t \in [0, T]$ and $\dot{H}^\alpha := D((-A)^{\alpha/2})$.

Pathwise approximations

- Spatial approximation: Replace H by a discrete subspace V_N with $\dim(V_N) = N \in \mathbb{N}$ and let $P_N : H \to V_N$ be the ONP onto V_N. The discrete operator $A_N : V_N \to V_N$ generates a semigroup $S_N = (S_N(t), t \geq 0)$ on V_N.

Pathwise approximations

- Spatial approximation: Replace H by a discrete subspace V_N with $\dim(V_N) = N \in \mathbb{N}$ and let $P_N : H \to V_N$ be the ONP onto V_N. The discrete operator $A_N : V_N \to V_N$ generates a semigroup $S_N = (S_N(t), t \geq 0)$ on V_N.
- Noise approximation: Let $(e_k, k \in \mathbb{N})$ denote the (orthonormal) eigenbasis of Q. We use a truncated Karhunen-Loève expansion to approximate W via

$$W(t) \approx W_K(t) := \sum_{k=1}^{K} (W(t), e_k)_H e_k, \quad K \in \mathbb{N}.$$

where $\{(W(t), e_k)_H\}_k$ is a sequence of real-valued and independent Brownian motions with variance $\eta_k = (Qe_k, e_k)_H$ (the k'th eigenvalue of Q).

Pathwise approximations (cont.)

- Time stepping: Use $M \in \mathbb{N}$ time steps and a rational approximation $r(\Delta t A_N) \approx S_N(\Delta t)$ for $\Delta t = {}^T\!/_M$.

$$r(\Delta t A_N)v = \sum_{n=1}^{N} r(\Delta t \widetilde{\lambda}_n)(v, \widetilde{f}_n)_H \widetilde{f}_n, \quad v \in H.$$

given the H-orthonormal eigenbasis $(\widetilde{f}_1, \dots, \widetilde{f}_N) \subset V_N$ of eigenfunctions of $(-A_N)$, with corresponding non-decreasing eigenvalues $(\widetilde{\lambda}_1, \dots, \widetilde{\lambda}_N)$.

Assumptions on pathwise approximations

1. The rational approximation r of S_N is of order $q \in \mathbb{N}$ and stable. That is, $r(z) = e^{-z} + \mathcal{O}(z^{q+1})$ as $z \to 0$, $|r(z)| < 1$ for $z > 0$ and $\lim_{z\to\infty} r(z) = 0$.
2. Subspace approximation property: Fix $\alpha > 0$ and let $(V_N, N \in \mathbb{N})$ be a sequence of subspaces $V_N \subset V$ such that $\dim(V_N) = N$. There are constants $C, \widetilde{\alpha} > 0$, depending on α and d, such that for any $N \in \mathbb{N}$ and any $v \in \dot{H}^\alpha$ there holds
$$\|v - P_N v\|_H \le C N^{-\widetilde{\alpha}} \|v\|_{\dot{H}^\alpha}, \quad \text{and} \quad \|A_N^{\min(\alpha,2)/2} P_N v\|_H \le C \|v\|_{\dot{H}^{\min(\alpha,2)}}.$$
3. Strong convergence: There are constants $C, \widetilde{\alpha}, \beta > 0$ such that for $p \in (0, 8]$ and all discretization parameters $M, N, K \in \mathbb{N}$ there holds the strong error estimate
$$\max_{m=0,\ldots,M} \|X(m\Delta t) - Y_m^{N,K}\|_{L^p(\Omega;H)} \le C \left(M^{-1/2} + N^{-\widetilde{\alpha}} + K^{-\beta} \right).$$

Truncated Milstein scheme

$$X_N(t) = S_N(t)P_N X_0 + \int_0^t S_N(t-s)P_N F(X_N(s))ds + \int_0^t S_N(t-s)P_N G(X_N(s))dW(s).$$

From now on, assume $F \equiv 0$. For fixed M, N, K, the truncated Milstein iteration is obtained using first order Taylor expansion of G and reads

Truncated Milstein scheme

$$X_N(t) = S_N(t)P_N X_0 + \int_0^t S_N(t-s)P_N F(X_N(s))ds + \int_0^t S_N(t-s)P_N G(X_N(s))dW(s).$$

From now on, assume $F \equiv 0$. For fixed M, N, K, the truncated Milstein iteration is obtained using first order Taylor expansion of G and reads

$$Y_{m+1}^{N,K} = r(\Delta t A_N)P_N Y_m^{N,K} + r(\Delta t A_N)P_N G(Y_m^{N,K})\Delta_m W_K$$

Truncated Milstein scheme

$$X_N(t) = S_N(t)P_N X_0 + \int_0^t S_N(t-s)P_N F(X_N(s))ds$$
$$+ \int_0^t S_N(t-s)P_N G(X_N(s))dW(s).$$

From now on, assume $F \equiv 0$. For fixed M, N, K, the truncated Milstein iteration is obtained using first order Taylor expansion of G and reads

$$Y_{m+1}^{N,K} = r(\Delta t A_N)P_N Y_m^{N,K} + r(\Delta t A_N)P_N G(Y_m^{N,K})\Delta_m W_K$$
$$+ \frac{r(\Delta t A_N)P_N}{2}\sum_{k,l=1}^{K} G'(Y_m^{N,K})\left(P_N G(Y_m^{N,K})\sqrt{\eta_l}e_l\right)$$
$$\sqrt{\eta_k}e_k(\Delta_m w_k \Delta_m w_l - \delta_{k,l}\Delta t).$$

where w_k are standard Brownian processes and $(e_k, k \in \mathbb{N})$ denote the eigenfunctions of Q with corresponding eigenvalues $(\eta_k, k \in \mathbb{N}) \subset \mathbb{R}_{\geq 0}$ in decaying order.

We introduce a continuous, square-integrable, $\mathcal{L}_1(H)$-valued, martingale on $[t_m, T]$

$$\mathcal{W}_{m,K}(s) := (W_K(s) - W_K(t_m)) \otimes (W_K(s) - W_K(t_m)) - (s - t_m) \sum_{k=1}^{K} \eta_k \, e_k \otimes e_k,$$

with a corresponding $\mathcal{L}_1(H)$-valued increment

$$\Delta_m \mathcal{W}_{m,K} := \Delta_m W_K \otimes \Delta_m W_K - \Delta t \sum_{k=1}^{K} \eta_k \, e_k \otimes e_k.$$

We introduce a continuous, square-integrable, $\mathcal{L}_1(H)$-valued, martingale on $[t_m, T]$

$$\mathcal{W}_{m,K}(s) := (W_K(s) - W_K(t_m)) \otimes (W_K(s) - W_K(t_m)) - (s - t_m) \sum_{k=1}^{K} \eta_k \, e_k \otimes e_k,$$

with a corresponding $\mathcal{L}_1(H)$-valued increment

$$\Delta_m \mathcal{W}_{m,K} := \Delta_m W_K \otimes \Delta_m W_K - \Delta t \sum_{k=1}^{K} \eta_k \, e_k \otimes e_k.$$

And defined $\mathcal{G} : H \to \mathcal{L}_{HS}(\mathcal{L}_{HS}(\mathcal{H}); H)$ such that the truncated Milstein is written as

$$Y_{m+1}^{N,K} = r(\Delta t A_N) P_N \Big(Y_m^{N,K} + G(Y_m^{N,K}) \Delta_m W_K + \mathcal{G}(Y_m^{N,K}) \Delta_m \mathcal{W}_{m,K} \Big).$$

Antithetic coupling

Fix $M, N, K \in \mathbb{N}$ and let the coarse scale discretization be given by

$$Y^c_{m+1} = r(\Delta t A_N) P_N (Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K})$$

for $m = 0, \ldots, M-1$.

Antithetic coupling

Fix $M, N, K \in \mathbb{N}$ and let the coarse scale discretization be given by

$$Y^c_{m+1} = r(\Delta t A_N) P_N(Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K})$$

for $m = 0, \dots, M-1$.

Fine scale: Let $\delta t := {}^{\Delta t}/_2$ and denote for $m = 0, {}^1/_2, 1, \dots, M - {}^1/_2, M$, the corresponding "fine increments" $\delta_m W_K$ and $\delta_m \mathcal{W}_{m,K}$, so that

$$\begin{aligned} \Delta_m W_K &= \delta_{m+1/2} W_K + \delta_m W_K \\ \Delta_m \mathcal{W}_{m,K_f} &= \delta_{m+1/2} \mathcal{W}_{m,K_f} + \delta_m \mathcal{W}_{m,K_f} \\ &+ \end{aligned}$$

Antithetic coupling

Fix $M, N, K \in \mathbb{N}$ and let the coarse scale discretization be given by

$$Y^c_{m+1} = r(\Delta t A_N) P_N(Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K})$$

for $m = 0, \ldots, M-1$.

Fine scale: Let $\delta t := {}^{\Delta t}/_2$ and denote for $m = 0, {}^1/_2, 1, \ldots, M - {}^1/_2, M$, the corresponding "fine increments" $\delta_m W_K$ and $\delta_m \mathcal{W}_{m,K}$, so that

$$\begin{aligned} \Delta_m W_K &= \delta_{m+1/2} W_K + \delta_m W_K \\ \Delta_m \mathcal{W}_{m,K_f} &= \delta_{m+1/2}\mathcal{W}_{m,K_f} + \delta_m \mathcal{W}_{m,K_f} \\ &\quad + \delta_{m+1/2} W_{K_f} \otimes \delta_m W_{K_f} + \delta_m W_{K_f} \otimes \delta_{m+1/2} W_{K_f} \end{aligned} .$$

Antithetic coupling

The fine discretization with $2M$ time steps and $N_f \geq N, K_f \geq K$ is then given by

$$\begin{aligned} Y^f_{m+1/2} &= r(\delta t A_{N_f}) P_{N_f} \big(Y^f_m + G(Y^f_m) \delta_m W_{K_f} + \mathcal{G}(Y^f_m) \delta_m \mathcal{W}_{m,K_f} \big), \\ Y^f_{m+1} &= r(\delta t A_{N_f}) P_{N_f} \big(Y^f_{m+1/2} + G(Y^f_{m+1/2}) \delta_{m+1/2} W_{K_f} \\ &\qquad\qquad + \mathcal{G}(Y^f_{m+1/2}) \delta_{m+1/2} \mathcal{W}_{m,K_f} \big). \end{aligned}$$

Antithetic coupling

The fine discretization with $2M$ time steps and $N_f \geq N, K_f \geq K$ is then given by

$$\begin{aligned} Y^f_{m+1/2} &= r(\delta t A_{N_f}) P_{N_f} \big(Y^f_m + G(Y^f_m) \delta_m W_{K_f} + \mathcal{G}(Y^f_m) \delta_m \mathcal{W}_{m,K_f} \big), \\ Y^f_{m+1} &= r(\delta t A_{N_f}) P_{N_f} \big(Y^f_{m+1/2} + G(Y^f_{m+1/2}) \delta_{m+1/2} W_{K_f} \\ &\qquad\qquad + \mathcal{G}(Y^f_{m+1/2}) \delta_{m+1/2} \mathcal{W}_{m,K_f} \big). \end{aligned}$$

The antithetic counter part of the fine discretization is

$$\begin{aligned} Y^a_{m+1/2} &= r(\delta t A_{N_f}) P_{N_f} \big(Y^a_m + G(Y^a_m) \delta_{m+1/2} W_{K_f} + \mathcal{G}(Y^a_m) \delta_{m+1/2} \mathcal{W}_{m,K_f} \big), \\ Y^a_{m+1} &= r(\delta t A_{N_f}) P_{N_f} \big(Y^a_{m+1/2} + G(Y^a_{m+1/2}) \delta_m W_{K_f} \\ &\qquad\qquad + \mathcal{G}(Y^a_{m+1/2}) \delta_m \mathcal{W}_{m,K_f} \big). \end{aligned}$$

Antithetic coupling II

We aim to estimate $\mathrm{E}[\,\Psi(X(T))\,]$ for $\Psi \in C_b^2(H;\mathbb{R})$ with MLMC.

Antithetic coupling II

We aim to estimate $\mathrm{E}[\,\Psi(X(T))\,]$ for $\Psi \in C_b^2(H;\mathbb{R})$ with MLMC.

Rather than using $\Psi(Y_M^f)$ on the fine levels of the MLMC estimator, we use the antithetic average

$$\overline{\Psi}_M := \frac{\Psi(Y_M^f) + \Psi(Y_M^a)}{2}.$$

Antithetic coupling II

We aim to estimate $\mathrm{E}[\,\Psi(X(T))\,]$ for $\Psi \in C_b^2(H;\mathbb{R})$ with MLMC.

Rather than using $\Psi(Y_M^f)$ on the fine levels of the MLMC estimator, we use the antithetic average

$$\overline{\Psi}_M := \frac{\Psi(Y_M^f) + \Psi(Y_M^a)}{2}.$$

For any $\Psi \in C_b^2(H;\mathbb{R})$ it holds that that

- $\mathrm{E}[\,\overline{\Psi}_M\,] = \mathrm{E}[\,\Psi(Y_M^f)\,]$, (no additional bias)

Antithetic coupling II

We aim to estimate $\mathrm{E}[\,\Psi(X(T))\,]$ for $\Psi \in C_b^2(H;\mathbb{R})$ with MLMC.

Rather than using $\Psi(Y_M^f)$ on the fine levels of the MLMC estimator, we use the antithetic average

$$\overline{\Psi}_M := \frac{\Psi(Y_M^f) + \Psi(Y_M^a)}{2}.$$

For any $\Psi \in C_b^2(H;\mathbb{R})$ it holds that that

- $\mathrm{E}[\,\overline{\Psi}_M\,] = \mathrm{E}[\,\Psi(Y_M^f)\,]$, (no additional bias)
- $\mathrm{E}[\,|\overline{\Psi}_M - \Psi(X(T))|^2\,] \le C\left(M^{-1} + N^{-2\tilde{\alpha}} + K^{-2\beta}\right)$, (strong error preserved)

Antithetic coupling II

We aim to estimate $\mathrm{E}[\,\Psi(X(T))\,]$ for $\Psi \in C_b^2(H;\mathbb{R})$ with MLMC.

Rather than using $\Psi(Y_M^f)$ on the fine levels of the MLMC estimator, we use the antithetic average

$$\overline{\Psi}_M := \frac{\Psi(Y_M^f) + \Psi(Y_M^a)}{2}.$$

For any $\Psi \in C_b^2(H;\mathbb{R})$ it holds that that

- $\mathrm{E}[\,\overline{\Psi}_M\,] = \mathrm{E}[\,\Psi(Y_M^f)\,]$, (no additional bias)
- $\mathrm{E}[\,|\overline{\Psi}_M - \Psi(X(T))|^2\,] \leq C\left(M^{-1} + N^{-2\tilde{\alpha}} + K^{-2\beta}\right)$, (strong error preserved)
- For $\overline{Y}_M := \frac{Y_M^f + Y_M^a}{2}$ there holds $\mathrm{E}[\,|\overline{\Psi}_M - \Psi(Y_M^c)|^2\,] \leq C\mathrm{E}[\,\|\overline{Y}_M - Y_M^c\|_H^2\,]$.

Antithetic coupling II

We aim to estimate $\mathrm{E}[\,\Psi(X(T))\,]$ for $\Psi \in C_b^2(H;\mathbb{R})$ with MLMC.

Rather than using $\Psi(Y_M^f)$ on the fine levels of the MLMC estimator, we use the antithetic average

$$\overline{\Psi}_M := \frac{\Psi(Y_M^f) + \Psi(Y_M^a)}{2}.$$

For any $\Psi \in C_b^2(H;\mathbb{R})$ it holds that that

- $\mathrm{E}[\,\overline{\Psi}_M\,] = \mathrm{E}[\,\Psi(Y_M^f)\,]$, (no additional bias)
- $\mathrm{E}[\,|\overline{\Psi}_M - \Psi(X(T))|^2\,] \leq C\left(M^{-1} + N^{-2\tilde{\alpha}} + K^{-2\beta}\right)$, (strong error preserved)
- For $\overline{Y}_M := \frac{Y_M^f + Y_M^a}{2}$ there holds $\mathrm{E}[\,|\overline{\Psi}_M - \Psi(Y_M^c)|^2\,] \leq C\mathrm{E}[\,\|\overline{Y}_M - Y_M^c\|_H^2\,]$.

$\Rightarrow$ "Antithetic variances" decay faster than $\mathcal{O}(M^{-1})$.

Improved variance decay for antithetic coupling in SPDEs

Theorem (H.-A. and A. Stein, 2023)

Let $\sup_{t\in[0,T]} X(t) \in L^8(\Omega; \dot{H}^\alpha)$ *hold for some* $\alpha \geq 1$, *and let* $M, N_f, N, K_f, K \in \mathbb{N}$ *be such that* $N_f \geq N$ *and* $K_f \geq K$. *Under suitable assumptions on* F, G *(twice Fréchet differentiable with bounded derivatives, linear growth, ...),* X_0 *and* Q, *there is a constant* $C > 0$, *independent of* M, N, *and* K, *such that the corrections in the antithetic Milstein scheme satisfy*

$$\mathrm{E}[\,\|\overline{Y}_M - Y_M^c\|_H^2\,] \leq C\left(M^{-\min(\alpha,2)} + N^{-2\tilde{\alpha}} + K^{-2\beta}\right).$$

Recall that for the Euler/truncated Milstein scheme without antithetic correction, we have

$$\mathrm{E}[\,\|Y_M^f - Y_M^c\|_H^2\,] \leq C\left(M^{-1} + N^{-2\tilde{\alpha}} + K^{-2\beta}\right).$$

Proof ideas

Prove that $\mathcal{G}$ is Fréchet differentiable, then using a Taylor expansion of G and $\mathcal{G}$, write

$$\begin{aligned}\overline{Y}_{m+1} = r(\delta t A_{N_f})^2 P_{N_f}\big(\overline{Y}_m + G(\overline{Y}_m)\Delta_m W_{K_f} + \mathcal{G}(\overline{Y}_m)\Delta_m \mathcal{W}_{m,K_f}\big) \\ + \overline{\Xi}_m + \overline{O}_m,\end{aligned}$$

where $\overline{\Xi}_m, \overline{O}_m : \Omega \to H$ are random variables such that

$$\mathrm{E}[\,\|\overline{\Xi}_m\|_H^2\,] \le C\Delta t^2\Big(M^{-\min(\alpha,2)} + N_f^{-2\alpha_0} + K_f^{-4\beta}\Big),$$

$$\mathrm{E}[\,\overline{O}_m\,\big|\mathcal{F}_{t_m}\,] = 0 \quad \text{and} \quad \mathrm{E}[\,\|\overline{O}_m\|_H^2\,] \le C\Delta t\Big(M^{-\min(\alpha,2)} + N_f^{-2\alpha_0} + K_f^{-4\beta}\Big).$$

Then expand the difference $\overline{Y}_M - Y_M^c$ and use Grönwall's inequality.

Proof ideas: Regularity limit

During expansion, we will have terms of the form

$$
\begin{aligned}
&F(r(\delta t A_{N_f}) Y_m^f) - r(\delta t A_{N_f}) P_{N_f} F(Y_m^f) \\
&= F(Y_m^f) + F'(\xi_m^2)\Big[r(\delta t A_{N_f}) Y_m^f - Y_m^f\Big] - r(\delta t A_{N_f}) P_{N_f} F(Y_m^f) \qquad (1)\\
&= \Big[I - r(\delta t A_{N_f}) P_{N_f}\Big] F(Y_m^f) + F'(\xi_m^2)\Big[r(\delta t A_{N_f}) - I\Big] Y_m^f
\end{aligned}
$$

Wherein, for example,

$$
\mathrm{E}\left[\left\| [r(\delta t A_{N_f}) - I] Y_m^f \right\|_H^2\right] \leq \| r(\delta t A_{N_f}) - I \|^2_{\mathcal{L}(\dot{H}^\alpha, H)} \mathrm{E}[\, \| Y_m^f \|^2_{\dot{H}^\alpha}\,]
$$

Finally, this is bounded using

$$
\begin{aligned}
\| r(\delta t A) - I \|_{\mathcal{L}(\dot{H}^\alpha, H)} &\leq \| r(\delta t A) - S(\delta t) \|_{\mathcal{L}(\dot{H}^\alpha, H)} + \| S(\delta t) - I \|_{\mathcal{L}(\dot{H}^\alpha, H)} \\
&\leq C \delta t^{\alpha/2}.
\end{aligned}
$$

From (Thomée, 2007) and (Pazy, 1983). Similar bounds hold when considering G and $\mathcal{G}$ after using Burkholder-Davis-Gundy inequality.

We can balance the errors via $N \approx M^{\min(\alpha,2)/2\tilde{\alpha}}$ and $K \approx M^{\min(\alpha,2)/2\beta}$ on all levels.

We can balance the errors via $N \approx M^{\min(\alpha,2)/2\tilde{\alpha}}$ and $K \approx M^{\min(\alpha,2)/2\beta}$ on all levels.

Let $\Psi \in C_b^2(H;\mathbb{R})$, $M_0 \in \mathbb{N}$, and let $M_\ell := M_0 2^\ell$ for $\ell \in \mathbb{N}_0$.

We can balance the errors via $N \approx M^{\min(\alpha,2)/2\tilde{\alpha}}$ and $K \approx M^{\min(\alpha,2)/2\beta}$ on all levels.

Let $\Psi \in C_b^2(H;\mathbb{R})$, $M_0 \in \mathbb{N}$, and let $M_\ell := M_0 2^\ell$ for $\ell \in \mathbb{N}_0$. Assume that

- there are $C > 0$ and $\gamma = \gamma(G) > 0$ such that

$$\text{"Cost of sampling } \overline{\Psi}_M \text{ on level } \ell\text{"} \leq C M_\ell^{1+\gamma}, \quad \forall \ell \in \mathbb{N}_0.$$

We can balance the errors via $N \approx M^{\min(\alpha,2)/2\tilde{\alpha}}$ and $K \approx M^{\min(\alpha,2)/2\beta}$ on all levels.

Let $\Psi \in C_b^2(H;\mathbb{R})$, $M_0 \in \mathbb{N}$, and let $M_\ell := M_0 2^\ell$ for $\ell \in \mathbb{N}_0$. Assume that

- there are $C > 0$ and $\gamma = \gamma(G) > 0$ such that

$$\text{"Cost of sampling } \overline{\Psi}_M \text{ on level } \ell\text{"} \leq C M_\ell^{1+\gamma}, \quad \forall \ell \in \mathbb{N}_0.$$

- for any $\delta \in (0,1)$ there is a constant $C = C(\Psi,\delta) > 0$ such that

$$|\mathrm{E}[\Psi(X(T))] - \mathrm{E}[\Psi(Y_{M_\ell}^{N_\ell,K_\ell})]| \leq C M_\ell^{-(1-\delta)}, \quad \forall \ell \in \mathbb{N}_0.$$

Theorem (H.-A. and A. Stein, 2023)

Under the previous conditions, there exists for any $\varepsilon \in (0, e^{-1})$ *an* *antithetic MLMC-Milstein* *estimator* $E_L^{anti}(\overline{\Psi}_M)$ *such that*

$$\mathrm{E}\big[\,|E_L^{anti}(\overline{\Psi}_M) - \mathrm{E}[\,\Psi(X(T))\,]|^2\,\big] \leq \varepsilon^2.$$

Theorem (H.-A. and A. Stein, 2023)

Under the previous conditions, there exists for any $\varepsilon \in (0, e^{-1})$ *an antithetic MLMC-Milstein estimator* $E_L^{anti}(\overline{\Psi}_M)$ *such that*

$$\mathrm{E}\left[\, |E_L^{anti}(\overline{\Psi}_M) - \mathrm{E}[\,\Psi(X(T))\,]|^2 \,\right] \leq \varepsilon^2.$$

The computational complexity $\mathcal{C}_{\mathrm{ML}}$ *to compute a realization of* $E_L^{anti}(\overline{\Psi}_M)$ *is bounded by*

$$\mathcal{C}_{\mathrm{ML}} \leq \begin{cases} C\varepsilon^{-2}, & \min(\alpha, 2) > 1 + \gamma, \\ C\varepsilon^{-2} |\log(\varepsilon)|^2, & \min(\alpha, 2) = 1 + \gamma, \\ C\varepsilon^{-2-\frac{1+\gamma-\min(\alpha,2)}{1-\delta}}, & \min(\alpha, 2) < 1 + \gamma. \end{cases}$$

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2,3\}$, $H := L^2(\mathcal{D})$ and let $A := \triangle$ be the Laplace-operator with hom. Dirichlet BCs. The eigenpairs $((\lambda_n, f_n), k \in \mathbb{N})$ of $(-A)$ are given in closed form $(\lambda_n \propto n^{2/d})$

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2,3\}$, $H := L^2(\mathcal{D})$ and let $A := \triangle$ be the Laplace-operator with hom. Dirichlet BCs. The eigenpairs $((\lambda_n, f_n), k \in \mathbb{N})$ of $(-A)$ are given in closed form ($\lambda_n \propto n^{2/d}$)
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > d/2$.

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2,3\}$, $H := L^2(\mathcal{D})$ and let $A := \triangle$ be the Laplace-operator with hom. Dirichlet BCs. The eigenpairs $((\lambda_n, f_n), k \in \mathbb{N})$ of $(-A)$ are given in closed form $(\lambda_n \propto n^{2/d})$
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > d/2$.
- We consider the stochastic heat equation given by

$$dX(t) = \triangle X(t)dt + G(X(t))dW(t), \quad X(0) = X_0, \tag{2}$$

for a random $X_0 \in L^8(\Omega; \dot{H}^2)$ and with diffusion coefficient $G : H \mapsto \mathcal{L}_{HS}(\mathcal{H}; H)$ given by (for $v \in H, u \in \mathcal{H}$)

$$G(v)u := \sum_{j=1}^{\infty} (v, e_j)_H e_{j+1}(u, \sqrt{\eta_{j+1}} e_{j+1})_{\mathcal{H}} + j^{-1/2-\varepsilon} e_j(u, \sqrt{\eta_j} e_j)_{\mathcal{H}}.$$

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2,3\}$, $H := L^2(\mathcal{D})$ and let $A := \triangle$ be the Laplace-operator with hom. Dirichlet BCs. The eigenpairs $((\lambda_n, f_n), k \in \mathbb{N})$ of $(-A)$ are given in closed form ($\lambda_n \propto n^{2/d}$)
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > d/2$.
- We consider the stochastic heat equation given by

$$dX(t) = \triangle X(t)dt + G(X(t))dW(t), \quad X(0) = X_0, \tag{2}$$

for a random $X_0 \in L^8(\Omega; \dot{H}^2)$ and with diffusion coefficient $G : H \mapsto \mathcal{L}_{HS}(\mathcal{H}; H)$ given by (for $v \in H, u \in \mathcal{H}$)

$$G(v)u := \sum_{j=1}^{\infty} (v, e_j)_H e_{j+1}(u, \sqrt{\eta_{j+1}} e_{j+1})_{\mathcal{H}} + j^{-1/2-\varepsilon} e_j(u, \sqrt{\eta_j} e_j)_{\mathcal{H}}.$$

- It holds that $X(t) \in L^8(\Omega; \dot{H}^\alpha)$ for $\alpha \in [1, \min(1+s, 2))$.

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2,3\}$, $H := L^2(\mathcal{D})$ and let $A := \triangle$ be the Laplace-operator with hom. Dirichlet BCs. The eigenpairs $((\lambda_n, f_n), k \in \mathbb{N})$ of $(-A)$ are given in closed form ($\lambda_n \propto n^{2/d}$)
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > d/2$.
- We consider the stochastic heat equation given by

$$dX(t) = \triangle X(t)dt + G(X(t))dW(t), \quad X(0) = X_0, \tag{2}$$

  for a random $X_0 \in L^8(\Omega; \dot{H}^2)$ and with diffusion coefficient $G : H \mapsto \mathcal{L}_{HS}(\mathcal{H}; H)$ given by (for $v \in H, u \in \mathcal{H}$)

$$G(v)u := \sum_{j=1}^{\infty} (v, e_j)_H e_{j+1}(u, \sqrt{\eta_{j+1}} e_{j+1})_{\mathcal{H}} + j^{-1/2-\varepsilon} e_j(u, \sqrt{\eta_j} e_j)_{\mathcal{H}}.$$

- It holds that $X(t) \in L^8(\Omega; \dot{H}^\alpha)$ for $\alpha \in [1, \min(1+s, 2))$.
- We combine the antithetic Milstein scheme with a spectral Galerkin approach and truncated Karhunen-Loève expansions for W. All errors are balanced via $\tilde{\alpha} = \alpha$ and $\beta = {}^{s}/_{d} - {}^{1}/_{2}$.

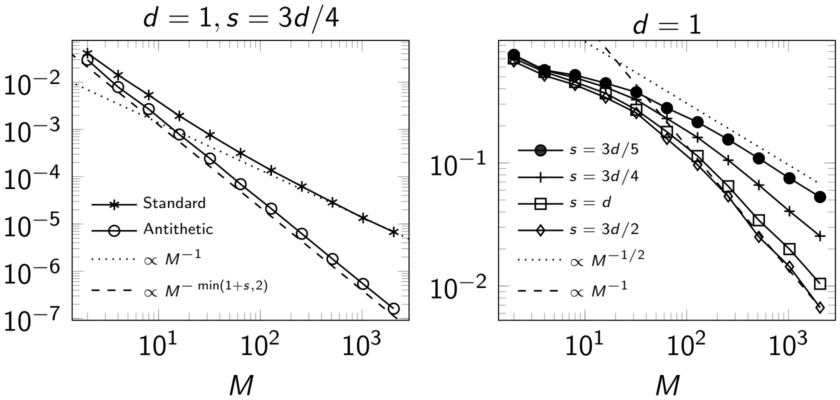

Figure: *(left)* Shows the the variance for the antithetic estimator and the variance for the "Standard" truncated Milstein estimator without the antithetic correction, for the smoothness parameter $s = 3d/4$. *(right)* Shows the relative variance decay between the two estimators, for different smoothness parameters s. The variance estimates were obtained using Monte Carlo sampling with at least 4000 samples. Recall $\alpha < \min(1 + s, 2)$.

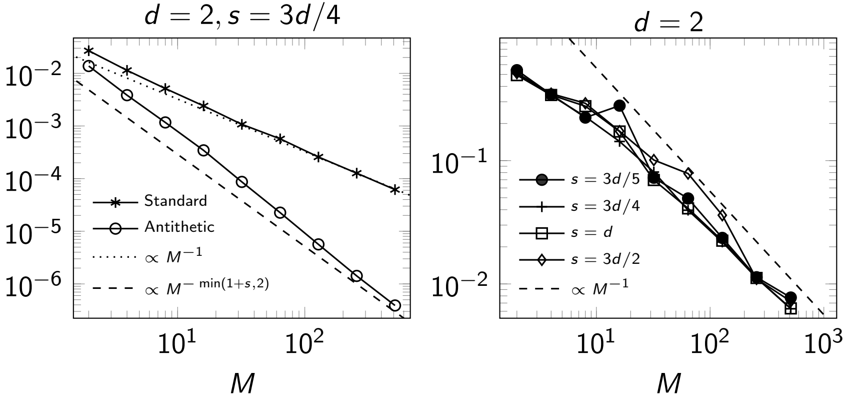

Figure: *(left)* Shows the the variance for the antithetic estimator and the variance for the "Standard" truncated Milstein estimator without the antithetic correction, for the smoothness parameter $s = 3d/4$. *(right)* Shows the relative variance decay between the two estimators, for different smoothness parameters s. The variance estimates were obtained using Monte Carlo sampling with at least 4000 samples. Recall $\alpha < \min(1 + s, 2)$.

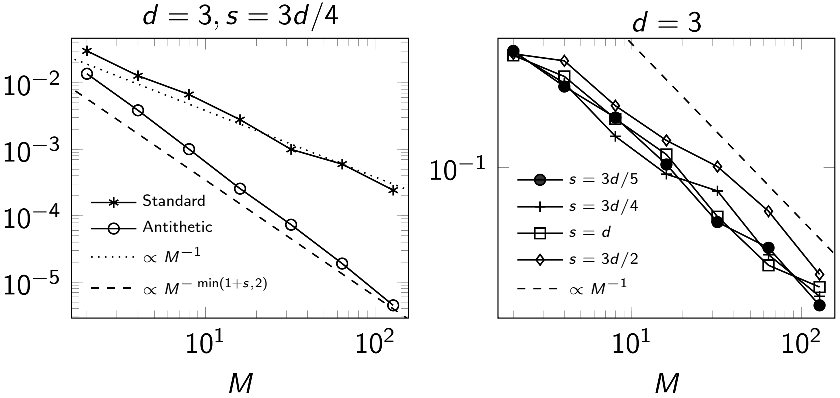

Figure: *(left)* Shows the the variance for the antithetic estimator and the variance for the "Standard" truncated Milstein estimator without the antithetic correction, for the smoothness parameter $s = 3d/4$. *(right)* Shows the relative variance decay between the two estimators, for different smoothness parameters s. The variance estimates were obtained using Monte Carlo sampling with at least 4000 samples. Recall $\alpha < \min(1 + s, 2)$.

Conclusions and outlook

Summary:

- Infinite-dimensional antithetic Milstein scheme for (parabolic) SPDEs.
- Avoids simulation of iterated integrals.
- Improved complexity (under certain conditions).
- Increase in efficiency depends on smoothness of the mild solution.

Conclusions and outlook

Summary:

- Infinite-dimensional antithetic Milstein scheme for (parabolic) SPDEs.
- Avoids simulation of iterated integrals.
- Improved complexity (under certain conditions).
- Increase in efficiency depends on smoothness of the mild solution.

Next steps:

- Treatment of noise approximation (antithetic/improved).
- Milstein (and truncated Milstein) has a quadratic cost for dense operators!
- Develop an antithetic Lévy area approximation for path dependent estimates.
- SPDEs with Lévy noise ($\Rightarrow$ BDG inequalities).
- Relax assumptions on F and G (Lipschitz and piece-wise twice-differentiable).
- First-order hyperbolic SPDEs (exploit weak formulation).
- Tamed schemes for non-Lipschitz drift coefficients.

References

A.-L. H.-A. and A. Stein: **An Antithetic Multilevel Monte Carlo-Milstein Scheme for Stochastic Partial Differential Equations**, preprint `arXiv:2307.14169`, 2023.

M.B. Giles and L. Szpruch **Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation**, *The Annals of Applied Probability*, 4:24, 1585–1620, 2014

R. Kruse: **Strong and Weak Approximation of Semilinear Stochastic Evolution Equations**, *Springer*, 2014.