ETH zürich

D-MATH
Seminar for Applied Mathematics

Antithetic Milstein scheme for SPDEs

Andreas Stein

Joint work with Abdul-Lateef Haji-Ali

MCM Paris, 28th June 2023

Model problem

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

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

Model problem

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

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

- $A : D(A) \subset H \to H$ 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.

Model problem

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

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

- $A : D(A) \subset H \to H$ 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.
- $F : H \to H$ is a (Lipschitz) non-linearity.

Model problem

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

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

- $A : D(A) \subset H \to H$ 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.
- $F : H \to H$ 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

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

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

- $A : D(A) \subset H \to H$ 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.
- $F : H \to H$ 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 W.

Model problem

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

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

- $A : D(A) \subset H \to H$ 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.
- $F : H \to H$ 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 W.
- $X_0 \in L^2(\Omega; H)$.

Model problem

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

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

- $A : D(A) \subset H \to H$ 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.
- $F : H \to H$ 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 W.
- $X_0 \in L^2(\Omega; H)$.

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

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)G(X(s))dW(s), \quad t \in [0,T].$$

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

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

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

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

- Time stepping: Use $M \in \mathbb{N}$ time steps and a **rational approximation** $r(\Delta t A_N) \approx S_N(\Delta t)$ for $\Delta t = T/M$. For the non-linear part:

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

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

- Time stepping: Use $M \in \mathbb{N}$ time steps and a **rational approximation** $r(\Delta t A_N) \approx S_N(\Delta t)$ for $\Delta t = T/M$. For the non-linear part:
  - Euler-Maruyama scheme: Slow variance decay of order $\mathcal{O}(M^{-1}) \ldots$

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

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

- Time stepping: Use $M \in \mathbb{N}$ time steps and a **rational approximation** $r(\Delta t A_N) \approx S_N(\Delta t)$ for $\Delta t = T/M$. For the non-linear part:
  - Euler-Maruyama scheme: Slow variance decay of order $\mathcal{O}(M^{-1})$ ...
  - Milstein scheme: Requires to simulate infinite-dim. **iterated integrals**, which is expensive/infeasible ...

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

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

- Time stepping: Use $M \in \mathbb{N}$ time steps and a **rational approximation** $r(\Delta t A_N) \approx S_N(\Delta t)$ for $\Delta t = T/M$. For the non-linear part:
  - Euler-Maruyama scheme: Slow variance decay of order $\mathcal{O}(M^{-1})$ ...
  - Milstein scheme: Requires to simulate infinite-dim. **iterated integrals**, which is expensive/infeasible ...

$\Rightarrow$ Extend the **antithetic MLMC-Milstein** scheme of Giles and Szpruch (2014) for SDEs to infinite dimensions.

Truncated Milstein scheme

From now on, assume $\boldsymbol{F} \equiv \mathbf{0}$. For fixed $\boldsymbol{M}, \boldsymbol{N}, \boldsymbol{K}$, the **truncated Milstein** iteration reads

Truncated Milstein scheme

From now on, assume $\boldsymbol{F \equiv 0}$. For fixed $\boldsymbol{M, N, K}$, the **truncated Milstein** iteration reads

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

Truncated Milstein scheme

From now on, assume $F \equiv 0$. For fixed M, N, K, the **truncated Milstein** iteration reads

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

We have used $w_k(\cdot) := (W(\cdot), e_k)_H$, where $(e_k, k \in \mathbb{N})$ denote the eigenfunctions of Q with corresponding eigenvalues $(\eta_k, k \in \mathbb{N}) \subset \mathbb{R}_{\geq 0}$ in decaying order.

Truncated Milstein scheme

From now on, assume $F \equiv 0$. For fixed M, N, K, the **truncated Milstein** iteration reads

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

We have used $w_k(\cdot) := (W(\cdot), e_k)_H$, where $(e_k, k \in \mathbb{N})$ denote the eigenfunctions of Q with corresponding eigenvalues $(\eta_k, k \in \mathbb{N}) \subset \mathbb{R}_{\geq 0}$ in decaying order.

We introduce $\mathcal{G} : H \to \mathcal{L}_{HS}(\mathcal{L}_{HS}(\mathcal{H}); H)$ and the $\mathcal{L}_1(H)$-valued increment

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

Truncated Milstein scheme

From now on, assume $F \equiv 0$. For fixed M, N, K, the **truncated Milstein** iteration reads

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

We have used $w_k(\cdot) := (W(\cdot), e_k)_H$, where $(e_k, k \in \mathbb{N})$ denote the eigenfunctions of Q with corresponding eigenvalues $(\eta_k, k \in \mathbb{N}) \subset \mathbb{R}_{\geq 0}$ in decaying order.

We introduce $\mathcal{G} : H \to \mathcal{L}_{HS}(\mathcal{L}_{HS}(\mathcal{H}); H)$ and the $\mathcal{L}_1(H)$-valued increment

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

$$\overset{\text{trunc. Milstein}}{\Longrightarrow} Y_{m+1}^{N,K} = r(\Delta t A_N) P_N \left(Y_m^{N,K} + G(Y_m^{N,K}) \Delta_m W_K + \mathcal{G}(Y_m^{N,K}) \Delta_m \mathcal{W}_{m,K} \right).$$

Antithetic coupling I

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

$$Y^c_{m+1} = r(\Delta t A_N) P_N \left(Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K} \right), \quad m = 0, \ldots, M-1.$$

Antithetic coupling I

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

$$Y^c_{m+1} = r(\Delta t A_N) P_N \left(Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K} \right), \quad m = 0, \ldots, M-1.$$

Fine scale: Let $\delta t := {}^{\Delta t}/_2$ and denote for $m = 0, {}^1/_2, 1, \ldots, M - {}^1/_2, M$, the corresonding "fine increments" $\delta_m W_K$ and $\delta_m \mathcal{W}_{m,K}$, so that $\Delta_m W_K = \delta_{m+1/2} W_K + \delta_m W_K$.

Antithetic coupling I

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

$$Y^c_{m+1} = r(\Delta t A_N) P_N \left(Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K}\right), \quad m = 0, \ldots, M-1.$$

Fine scale: Let $\delta t := {}^{\Delta t}/_2$ and denote for $m = 0, {}^1/_2, 1, \ldots, M - {}^1/_2, M$, the corresonding "fine increments" $\delta_m W_K$ and $\delta_m \mathcal{W}_{m,K}$, so that $\Delta_m W_K = \delta_{m+1/2} W_K + \delta_m W_K$.

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

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

Antithetic coupling I

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

$$Y^c_{m+1} = r(\Delta t A_N) P_N \left(Y^c_m + G(Y^c_m)\Delta_m W_K + \mathcal{G}(Y^c_m)\Delta_m \mathcal{W}_{m,K} \right), \quad m = 0, \dots, M-1.$$

Fine scale: Let $\delta t := \Delta t/2$ and denote for $m = 0, 1/2, 1, \dots, M - 1/2, M$, the corresonding "fine increments" $\delta_m W_K$ and $\delta_m \mathcal{W}_{m,K}$, so that $\Delta_m W_K = \delta_{m+1/2} W_K + \delta_m W_K$.

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

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

The **antithetic** counter part of the fine discretization is

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

Antithetic coupling II

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

Antithetic coupling II

We aim to estimate $\mathbb{E}\left(\boldsymbol{\Psi}(\boldsymbol{X}(\boldsymbol{T}))\right)$ for $\boldsymbol{\Psi} \in \boldsymbol{C}_{\boldsymbol{b}}^{\boldsymbol{2}}(\boldsymbol{H}; \mathbb{R})$ with MLMC.

Rather than using $\boldsymbol{\Psi}(\boldsymbol{Y}_{\boldsymbol{M}}^{\boldsymbol{f}})$ on the fine levels of the MLMC estimator, we use the **antithetic average**

$$\overline{\boldsymbol{\Psi}}_{\boldsymbol{M}} := \frac{\boldsymbol{\Psi}(\boldsymbol{Y}_{\boldsymbol{M}}^{\boldsymbol{f}}) + \boldsymbol{\Psi}(\boldsymbol{Y}_{\boldsymbol{M}}^{\boldsymbol{a}})}{\boldsymbol{2}}.$$

Antithetic coupling II

We aim to estimate $\mathbb{E}\left(\boldsymbol{\Psi}(\boldsymbol{X}(\boldsymbol{T}))\right)$ for $\boldsymbol{\Psi} \in C_b^2(\boldsymbol{H};\mathbb{R})$ with MLMC.

Rather than using $\boldsymbol{\Psi}(\boldsymbol{Y}_M^f)$ on the fine levels of the MLMC estimator, we use the **antithetic average**

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

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

- $\mathbb{E}(\overline{\boldsymbol{\Psi}}_M) = \mathbb{E}(\boldsymbol{\Psi}(\boldsymbol{Y}_M^f)),$ (no additional bias)

Antithetic coupling II

We aim to estimate $\mathbb{E}\left(\Psi(X(T))\right)$ 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

- $\mathbb{E}(\overline{\Psi}_M) = \mathbb{E}(\Psi(Y_M^f)),$ (no additional bias)
- $\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(X(T))\right|^2\right) \leq C\left(M^{-1} + N^{-2\alpha_0} + K^{-2\beta}\right),$ (strong error preserved)

Antithetic coupling II

We aim to estimate $\mathbb{E}\left(\Psi(X(T))\right)$ 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

- $\mathbb{E}(\overline{\Psi}_M) = \mathbb{E}(\Psi(Y_M^f)),$ (no additional bias)
- $\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(X(T))\right|^2\right) \leq C\left(M^{-1} + N^{-2\alpha_0} + K^{-2\beta}\right),$ (strong error preserved)
- For $\overline{Y}_M := \frac{Y_M^f + Y_M^a}{2}$ there holds $\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(Y_M^c)\right|^2\right) \leq C\mathbb{E}\left(\left\|\overline{Y}_M - Y_M^c\right\|_H^2\right).$

Antithetic coupling II

We aim to estimate $\mathbb{E}\left(\Psi(X(T))\right)$ 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

- $\mathbb{E}(\overline{\Psi}_M) = \mathbb{E}(\Psi(Y_M^f))$, (no additional bias)
- $\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(X(T))\right|^2\right) \le C\left(M^{-1} + N^{-2\alpha_0} + K^{-2\beta}\right)$, (strong error preserved)
- For $\overline{Y}_M := \frac{Y_M^f + Y_M^a}{2}$ there holds $\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(Y_M^c)\right|^2\right) \le C\mathbb{E}\left(\left\|\overline{Y}_M - Y_M^c\right\|_H^2\right)$.

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

Improved variance decay for antithetic coupling in SPDEs

ML corrections: $$\mathbb{E}\left(\left|\overline{\boldsymbol{\Psi}}_M - \boldsymbol{\Psi}(\boldsymbol{Y}_M^c)\right|^2\right) \leq C\mathbb{E}\left(\left\|\overline{\boldsymbol{Y}}_M - \boldsymbol{Y}_M^c\right\|_H^2\right).$$

Improved variance decay for antithetic coupling in SPDEs

ML corrections: $$\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(Y_M^c)\right|^2\right) \leq C\mathbb{E}\left(\left\|\overline{Y}_M - Y_M^c\right\|_H^2\right).$$

Theorem (A.L. Haji-Ali and A.S., 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, 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*

$$\mathbb{E}\left(\left\|\overline{Y}_M - Y_M^c\right\|_H^2\right) \leq C\left(M^{-\min(\alpha,2)} + N^{-2\alpha_0} + K^{-2\beta}\right).$$

Improved variance decay for antithetic coupling in SPDEs

ML corrections: $\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(Y_M^c)\right|^2\right) \le C\mathbb{E}\left(\left\|\overline{Y}_M - Y_M^c\right\|_H^2\right).$

Theorem (A.L. Haji-Ali and A.S., 2023)

*Let $\sup_{t\in[0,T]} X(t) \in L^8(\Omega; \dot{H}^\alpha)$ hold for some $\alpha \ge 1$, and let $M, N_f, N, K_f, K \in \mathbb{N}$ be such that $N_f \ge N$ and $K_f \ge K$. Under suitable assumptions on F, G, 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*

$$\mathbb{E}\left(\left\|\overline{Y}_M - Y_M^c\right\|_H^2\right) \le C\left(M^{-\min(\alpha,2)} + N^{-2\alpha_0} + K^{-2\beta}\right).$$

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

$$\mathbb{E}\left(\left\|Y_M^f - Y_M^c\right\|_H^2\right) \le C\left(M^{-1} + N^{-2\alpha_0} + K^{-2\beta}\right).$$

Improved variance decay for antithetic coupling in SPDEs

ML corrections: $$\mathbb{E}\left(\left|\overline{\Psi}_M - \Psi(Y^c_M)\right|^2\right) \leq C\mathbb{E}\left(\left\|\overline{Y}_M - Y^c_M\right\|^2_H\right).$$

Theorem (A.L. Haji-Ali and A.S., 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, 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*

$$\mathbb{E}\left(\left\|\overline{Y}_M - Y^c_M\right\|^2_H\right) \leq C\left(M^{-\min(\alpha,2)} + N^{-2\alpha_0} + K^{-2\beta}\right).$$

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

$$\mathbb{E}\left(\left\|Y^f_M - Y^c_M\right\|^2_H\right) \leq C\left(M^{-1} + N^{-2\alpha_0} + K^{-2\beta}\right).$$

- Error balancing via $N \approx M^{\min(\alpha,2)/2\alpha_0}$ and $K \approx M^{\min(\alpha,2)/2\beta}$ on all levels.

Theorem (A.L. Haji-Ali and A.S., 2023)

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

Theorem (A.L. Haji-Ali and A.S., 2023)

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*

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

Theorem (A.L. Haji-Ali and A.S., 2023)

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*

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

Theorem (A.L. Haji-Ali and A.S., 2023)

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*

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

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

$$\mathbb{E}\left(\left|E_L^{\text{anti}}(\overline{\Psi}_M) - \mathbb{E}(\Psi(X(T)))\right|^2\right) \leq \varepsilon^2.$$

Theorem (A.L. Haji-Ali and A.S., 2023)

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*

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

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

$$\mathbb{E}\left(\left| E_L^{\text{anti}}(\overline{\Psi}_M) - \mathbb{E}(\Psi(X(T))) \right|^2 \right) \leq \varepsilon^2.$$

The computational complexity $\mathcal{C}_{\mathrm{ML}}$ *to compute a realization of* $E_L^{\text{anti}}(\overline{\Psi}_M)$ *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 $\mathcal{D} = [0,1]^d$, $d \in \{1,2\}$, $H := L^2(\mathcal{D})$ 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.

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2\}$, $H := L^2(\mathcal{D})$ 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.
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > 0$.

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2\}$, $H := L^2(\mathcal{D})$ 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.
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > 0$.
- We consider the stochastic heat equation given by

$$dX(t) = \triangle X(t)dt + G(X(t))dW(t), \quad X(0) = X_0, \tag{1}$$

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

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

Numerical example: Stochastic heat equation

- Let $\mathcal{D} = [0,1]^d$, $d \in \{1,2\}$, $H := L^2(\mathcal{D})$ 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.
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > 0$.
- We consider the stochastic heat equation given by

$$dX(t) = \triangle X(t)dt + G(X(t))dW(t), \quad X(0) = X_0, \tag{1}$$

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

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

- 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 $\mathcal{D} = [0,1]^d$, $d \in \{1,2\}$, $H := L^2(\mathcal{D})$ 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.
- W is a Q-Wiener process with operator $Q = ((-\triangle)^{-s})$ for a smoothness parameter $s > 0$.
- We consider the stochastic heat equation given by

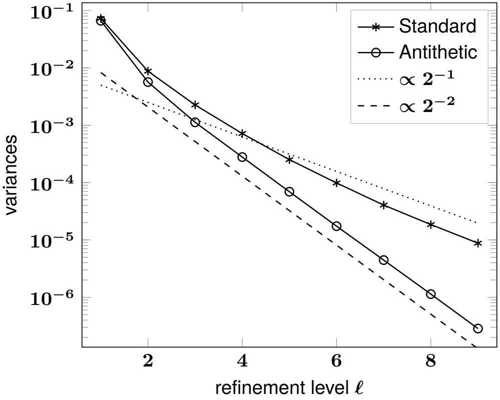
$$dX(t) = \triangle X(t)dt + G(X(t))dW(t), \quad X(0) = X_0, \tag{1}$$

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

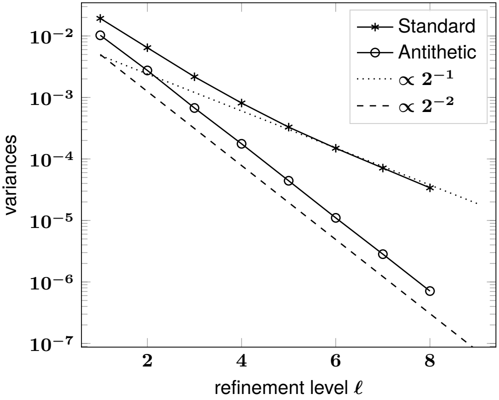
$$G(v)u := \sum_{j=1}^{\infty} (v, e_j)_H e_{j+1}(u, \sqrt{\eta_{j+1}} e_{j+1})_{\mathcal{H}} + (1, e_j)_H e_j (u, \sqrt{\eta_j} e_j)_{\mathcal{H}}, \quad v \in H, u \in \mathcal{H}.$$

- 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 $\alpha_0 = \alpha$ and $\beta = {}^s/_d - {}^1/_2$.

$d = 1, s = 1$



$d = 2, s = 2$



Log-plot of the estimated antithetic difference $\mathbb{E}(\|\overline{Y}^{\ell}_{M_\ell} - Y^{\ell-1,c}_{M_\ell}\|^2_H)$, and the standard multilevel difference $\mathbb{E}(\|Y^{\ell,f}_{M_\ell} - Y^{\ell-1,c}_{M_{\ell-1}}\|^2_H)$ against the refinement level. The "antithetic variances" decrease proportional to $M_\ell^{-\min(1+s,2)}$, whereas the standard ML difference decreases proportional to M_ℓ^{-1}.

Conclusions and outlook

Summary:

- First infinite-dimensional antithetic Milstein scheme for (parabolic) SPDEs
- Avoid simulation of iterated integrals
- Significantly improved complexity (under certain conditions)
- Increase in efficiency depends on smoothness of the mild solution
- Performance depends on the cost of evaluating G and the decay of the EVs of Q

Conclusions and outlook

Summary:

- First infinite-dimensional antithetic Milstein scheme for (parabolic) SPDEs
- Avoid simulation of iterated integrals
- Significantly improved complexity (under certain conditions)
- Increase in efficiency depends on smoothness of the mild solution
- Performance depends on the cost of evaluating G and the decay of the EVs of Q

Extensions:

- Include an (antithetic/improved) noise approximation
- SPDEs with Lévy noise ($\Rightarrow$ BDG inequalities)
- First-order hyperbolic SPDEs (exploit weak formulation)
- Tamed schemes for non-Lipschitz drift coefficients

References

A.-L. Haji-Ali and A. Stein: **An Antithetic Multilevel Monte Carlo-Milstein Scheme for Stochastic Partial Differential Equations**, *manuscript in preparation*, 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.