A multi-index Monte Carlo method for semilinear parabolic SPDEs

Abdul-Lateef Haji-Ali.
Work in progress with Håkon Hoel (University of Oslo) and Andreas Petersson (Linnaeus University).

Heriot-Watt University

Numerical Analysis seminar, Mathematical Institute, Oxford

Problem description

Given an of H-valued SDE

$$\mathrm{d}X(t) + \big(AX(t) - F(X(t))\big)\,\mathrm{d}t = \big(I + GX(t)\big)\,\mathrm{d}W(t), \quad t \in (0, T]$$
$$X(0) = X_0 \in H$$

Goal: Approximate $\mathbb{E}[\Psi(X(T))]$ for smooth QoI $\Psi : H \to U$.

Contribution: A multi-index Monte Carlo method

$$\mu_{\mathrm{MI}} := \sum_{\boldsymbol{\ell} \in \mathcal{I} \subset \mathbb{N}_0^2} \sum_{i_{\boldsymbol{\ell}}=1}^{m_{\boldsymbol{\ell}}} \frac{\Psi(X_{N_{\ell_2}}^{M_{\ell_1},\boldsymbol{\ell},i}) - \Psi(X_{N_{\ell_2}}^{M_{\ell_1-1},\boldsymbol{\ell},i}) - \Psi(X_{N_{\ell_2-1}}^{M_{\ell_1},\boldsymbol{\ell},i}) + \Psi(X_{N_{\ell_2-1}}^{M_{\ell_1-1},\boldsymbol{\ell},i})}{m_{\boldsymbol{\ell}}}$$

that achieves

$$\mathbb{E}[\|\mu_{\mathrm{MI}} - \mathbb{E}[\Psi(X(T))]\|_U^2] \lesssim \varepsilon^2$$

at a computational cost of almost $\mathcal{O}(\varepsilon^{-2})$ [1].

[1]under favourable conditions.

Overview

1. The SPDE and mild solutions
2. Monte Carlo methods for approximating $\mathbb{E}[\Psi(X(T))]$
3. Accelerated exponential integrator method
4. Numerical experiments and conclusion

Canonical example: Stochastic heat equation

$$\mathrm{d}X(t) + \big(-\Delta X(t) - F(X(t))\big)\,\mathrm{d}t = \big(I + GX(t)\big)\,\mathrm{d}W(t)$$

Here $X \in L^2(\mathcal{D})$ for a bounded domain $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$, convex or with $\mathcal{C}^2$ boundary $\partial\mathcal{D}$. $-\Delta : \mathrm{Dom}(-\Delta) = W^{2,2}(\mathcal{D}) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is a densely defined, positive definite linear operator with orthonormal eigenbasis

$$\big((e_j, \lambda_j)\big)_{j=1}^{\infty}, \qquad \text{where } \lambda_j \eqsim j^{2/d}$$

by Weyl's law. We consider $\mathcal{D}$ for which (e_j, λ_j) are known. E.g., $\mathcal{D} = (0, 1)$ with

$$e_j(x) = \sqrt{2}\sin(j\pi x), \qquad \text{and} \quad \lambda_j = \pi^2 j^2.$$

A natural example of F is a composition (Nemytskii) mapping:

$$(F(u))(x) = f(u(x))$$

where $f : \mathbb{R} \to \mathbb{R}$ is sufficiently smooth with bounded derivatives.

The General SPDE

We consider

$$\mathrm{d}X(t) + \big(AX(t) - F(X(t))\big)\,\mathrm{d}t = \big(I + GX(t)\big)\,\mathrm{d}W(t)$$

on Hilbert space $H = (H, \langle\cdot,\cdot\rangle, \|\cdot\|)$ and where: $A : \mathrm{Dom}(A) \subset H \to H$ is a

densely defined, positive definite linear operator with orthonormal eigenbasis

$$\big((e_j, \lambda_j)\big)_{j=1}^{\infty}, \qquad \text{where } \lambda_j \eqsim j^{\nu} \quad \text{for some } \nu > 0$$

Fractional operators $\qquad A^r v := \sum_{j=1}^{\infty} \lambda_j^r \langle e_j, v\rangle e_j \quad \text{for} \quad r \in \mathbb{R}$

Extension of norm: $\|\cdot\|_{\dot{H}^r} := \|A^{r/2}\cdot\|$ and associated Hilbert space

$$\dot{H}^r = \begin{cases} \mathrm{Dom}(A^{r/2}) & r \geq 0 \\ \overline{H}^{\|\cdot\|_{\dot{H}^r}} & r < 0, \end{cases}$$

Q-Wiener process on Hilbert Space

W is a $Q-$Wiener process, meaning it has independent increments

$$W(t+\Delta t) - W(t) \sim \mathcal{N}(0, \Delta t Q).$$

Here, for any $\phi, \psi \in H, t \geq 0$,

$$\mathbb{E}[\langle W(t), \phi \rangle \langle W(t), \psi \rangle] = t \langle Q\phi, \psi \rangle,$$

for $Q \in \mathcal{L}_1^+(H)$ being a covariance operator (non-negative, trace-class, self-adjoint) with eigenpairs (e_k, μ_k) (same eigenbasis as operator A!).

Representation: $$W(t) = \sum_{j=1}^{\infty} \sqrt{\mu_j} B_j(t) e_j$$

with $B_j(t)$ independent, scalar-valued Brownian motions, defined on a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P})$.

The noise operator $G \in \mathcal{L}\big(H, \mathcal{L}_2(Q^{1/2}(H), H)\big)$ is specified later.

Q-Wiener process

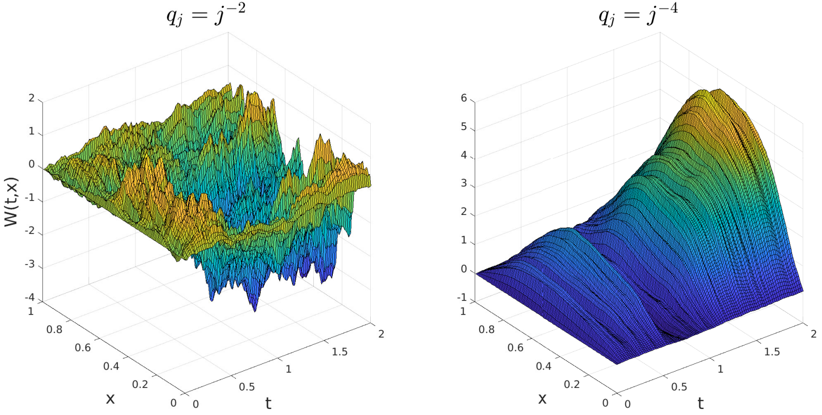


Figure 1: Regularity of W depends on decay rate of $(\mu_j = q_j)$

Mild solution[2]

If X_0 is $\mathcal{F}_0-$measurable and $X_0 \in L^p(\Omega, H)$, then there exists a unique mild solution in $C([0,T], L^p(\Omega, H))$ to the SPDE

$$\mathrm{d}X(t) + \big(AX(t) - F(X(t))\big)\,\mathrm{d}t = \big(I + GX(t)\big)\,\mathrm{d}W(t), \quad t \in (0,T]$$
$$X(0) = X_0.$$

Definition: a mild solution is an H-valued predictable process $\{X(t)\}_{t\in[0,T]}$ satisfying

$$X(t) = e^{-At}X_0 + \int_0^t e^{-A(t-s)}F(X(s))\,\mathrm{d}s + \int_0^t e^{-A(t-s)}\big(I + GX(s)\big)\,\mathrm{d}W(s)$$

for each $t \in [0,T]$. Here, $e^{-At}e_j = e^{-\lambda_j t}e_j$.

[2] Da Prato, G. & Zabczyk, J. *Stochastic equations in infinite dimensions.* Second, xviii+493 (Cambridge University Press, Cambridge, 2014).

Overview

1. The SPDE and mild solutions
2. Monte Carlo methods for approximating $\mathbb{E}[\Psi(X(T))]$
3. Accelerated exponential integrator method
4. Numerical experiments and conclusion

Multilevel Monte Carlo (MLMC)[3]

Let X_N^M be a *general* numerical method with respect to integers M (e.g. timesteps) and N (e.g. eigenfunctions), and let $\Psi : H \to U$, for a real separable Hilbert space space $U = (U, \langle \cdot, \cdot \rangle_U, \| \cdot \|_U)$.
For integer sequences $(M_\ell)_{\ell=0}^\infty$ and $(N_\ell)_{\ell=0}^\infty$, consider the telescoping sum

$$\mathbb{E}[\Psi(X_{N_L}^{M_L}(T))] = \sum_{\ell=1}^{L} \mathbb{E}[\Psi(X_{N_\ell}^{M_\ell}(T)) - \Psi(X_{N_{\ell-1}}^{M_{\ell-1}}(T))] + \mathbb{E}[\Psi(X_{N_0}^{M_0}(T))].$$

This motivates the MLMC estimator based on sample averages E_{m_ℓ},

$$\mu_{\mathrm{ML}} := \sum_{\ell=1}^{L} E_{m_\ell}[\Psi(X_{N_\ell}^{M_\ell}(T)) - \Psi(X_{N_{\ell-1}}^{M_{\ell-1}}(T))] + E_{m_0}[\Psi(X_{N_0}^{M_0}(T))].$$

We assume that (here and below we ignore logarithmic terms!):

$$\mathrm{Cost}(\Psi(X_N^M)) \eqsim MN.$$

[3] Giles, M. B. Multilevel monte carlo path simulation. *Operations research* **56,** 607–617 (2008).

Multilevel Monte Carlo (MLMC)

Central point: given that

$$\|\Psi(X(T)) - \Psi(X_N^M(T))\|^2_{L^2(\Omega,H)} \lesssim M^{-\beta_1} + N^{-\beta_2}$$

for some $\beta_1, \beta_2 \geq 0$, then with $M_\ell = 2^\ell$ and $N_\ell \eqsim 2^{\beta_1 \ell/\beta_2}$

$$\|\Psi(X_{N_\ell}^{M_\ell}(T)) - \Psi(X_{N_{\ell-1}}^{M_{\ell-1}}(T))\|^2_{L^2(\Omega,U)} \lesssim 2^{-\ell\beta_1}$$

so very few samples m_ℓ needed when $\ell \gg 1$.
Performance[4]**:** For any $\varepsilon > 0$, there exists $(m_\ell)_\ell \subset \mathbb{N}$ s.t.

$$\mathbb{E}[\|\mu_{\mathrm{ML}} - \mathbb{E}[\Psi(X(T))]\|_U^2] \lesssim \varepsilon^2$$

with, given *weak convergence rates* α_1, α_2,
$\mathrm{Cost}(\mu_{\mathrm{ML}}) \lesssim \varepsilon^{-2-\max\left(0, \frac{1+\beta_1/\beta_2-\beta_1}{\min(\alpha_1, \alpha_2\beta_1/\beta_2)}\right)} \lesssim \varepsilon^{-2-2\max\left(0, \frac{1}{\beta_1}+\frac{1}{\beta_2}-1\right)}$

Monte Carlo cost is $\varepsilon^{-2-\frac{1}{\alpha_1}-\frac{1}{\alpha_2}} \leq \varepsilon^{-2-2\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right)}$

[4] Chada, N. K., Hoel, H., Jasra, A. & Zouraris, G. E. Improved efficiency of multilevel Monte Carlo for stochastic PDE through strong pairwise coupling. *J. Sci. Comput.* **93,** Paper No. 62, 29 (2022).

Multi-index Monte Carlo (MIMC)[5]

Consider the telescoping double-sum

$$\mathbb{E}[\Psi(X_{N_L}^{M_L}(T))] = \sum_{\ell_1=0}^{L}\sum_{\ell_2=0}^{L}\mathbb{E}[\underbrace{\Psi(X_{N_{\ell_2}}^{M_{\ell_1}}) - \Psi(X_{N_{\ell_2}}^{M_{\ell_1-1}}) - \Psi(X_{N_{\ell_2-1}}^{M_{\ell_1}}) + \Psi(X_{N_{\ell_2-1}}^{M_{\ell_1-1}})}_{=:\Delta_{\boldsymbol{\ell}}\Psi(X)}]$$

with $\Psi(X_{N_{\ell_2}}^{M_{\ell_1}}(T)) := 0$ whenever $\min(\ell_1, \ell_2) < 0$.
This motivates the MIMC estimator

$$\mu_{\mathrm{MI}} := \sum_{\boldsymbol{\ell}\in\mathcal{I}} E_{m_{\boldsymbol{\ell}}}[\Delta_{\boldsymbol{\ell}}\Psi(X)]$$

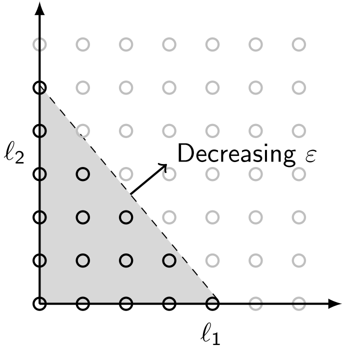
where $\mathcal{I} = \{\boldsymbol{\ell} = (\ell_1, \ell_2) \in \mathbb{N}_0^2 \mid \max(\ell_1, \ell_2) \leq L\}$ and $(m_{\boldsymbol{\ell}})_{\boldsymbol{\ell}\in\mathbb{N}_0^2} \subset \mathbb{N}$.

[5] Haji-Ali, A.-L., Nobile, F. & Tempone, R. Multi-index Monte Carlo: when sparsity meets sampling. *Numerische Mathematik* **132**, 767–806 (2016).

The index set

Triangular index sets are more efficent than rectangular ones, so for some suitable weights (w_1, w_2) we actually use:

$$\mathcal{I} = \{\boldsymbol{\ell} \in \mathbb{N}_0^2 \mid w_1\ell_1 + w_2\ell_2 \le L(\varepsilon)\}$$



MIMC for SPDE for general numerical method

Theorem

Suppose that $\Psi(X_N^M(T)) \to \Psi(X(T))$ *as* $M, N \to \infty$. *Assume also that we have* **a multiplicative** *bound on the* **second order difference**

$$\|\Delta_{\boldsymbol{\ell}} \Psi(X)\|^2_{L^2(\Omega,U)} \lesssim 2^{-\beta_1 \ell_1 - \beta_2 \ell_2}$$

for $\beta_1, \beta_2 > 0$. *Then, there is* $\alpha_i \geq \beta_i/2$ *such that*

$$\|\mathbb{E}[\Delta_{\boldsymbol{\ell}} \Psi(X)]\|_U \lesssim 2^{-\alpha_1 \ell_1 - \alpha_2 \ell_2}$$

and there exist MIMC parameters $\mathcal{I}$ *and* $(m_{\boldsymbol{\ell}}) \subset \mathbb{N}$ *s.t.*

$$\|\mu_{\mathrm{MI}} - \mathbb{E}[\Psi(X(T))]\|^2_{L^2(\Omega,U)} \lesssim \varepsilon^2$$

with

$$\mathrm{Cost}(\mu_{\mathrm{MI}}) \eqsim \varepsilon^{-2-2u} |\log \varepsilon^{-1}|^r$$

for some $0 < r \leq 4 + 2u$ *and*

$$2u = \max\left(0, \frac{1-\beta_1}{\alpha_1}, \frac{1-\beta_2}{\alpha_2}\right)$$

MIMC for SPDE for general numerical method

Theorem

Suppose that $\Psi(X_N^M(T)) \to \Psi(X(T))$ *as* $M, N \to \infty$*. Assume also that we have* **a multiplicative** *bound on the* **second order difference**

$$\|\Delta_{\boldsymbol{\ell}} \Psi(X)\|^2_{L^2(\Omega,U)} \lesssim 2^{-\beta_1 \ell_1 - \beta_2 \ell_2}$$

for $\beta_1, \beta_2 > 0$*. Then, there is* $\alpha_i \geq \beta_i/2$ *such that*

$$\|\mathbb{E}[\Delta_{\boldsymbol{\ell}} \Psi(X)]\|_U \lesssim 2^{-\alpha_1 \ell_1 - \alpha_2 \ell_2}$$

and there exist MIMC parameters $\mathcal{I}$ *and* $(m_{\boldsymbol{\ell}}) \subset \mathbb{N}$ *s.t.*

with
$$\|\mu_{\mathsf{MI}} - \mathbb{E}[\Psi(X(T))]\|^2_{L^2(\Omega,U)} \lesssim \varepsilon^2$$
$$\mathrm{Cost}(\mu_{\mathsf{MI}}) \eqsim \varepsilon^{-2-2u} |\log \varepsilon^{-1}|^r$$

for some $0 < r \leq 4 + 2u$ *and*

$$2u = \max\left(0, \frac{1}{\beta_1} - 1, \frac{1}{\beta_2} - 1\right)$$

MIMC for SPDE for general numerical method

Theorem

Suppose that $\Psi(X_N^M(T)) \to \Psi(X(T))$ *as* $M, N \to \infty$. *Assume also that we have* **a multiplicative** *bound on the* **second order difference**

$$\|\Delta_{\boldsymbol{\ell}} \Psi(X)\|_{L^2(\Omega,U)}^2 \lesssim 2^{-\beta_1 \ell_1 - \beta_2 \ell_2 - \vartheta \max(\ell_1 - \upsilon \ell_2, 0)}$$

for $\beta_1, \beta_2 > 0$. *Then, there is* $\alpha_i \geq \beta_i/2$ *such that*

$$\|\mathbb{E}[\Delta_{\boldsymbol{\ell}} \Psi(X)]\|_U \lesssim 2^{-\alpha_1 \ell_1 - \alpha_2 \ell_2 - (\vartheta/2) \max(\ell_1 - \upsilon \ell_2, 0)}$$

and there exist MIMC parameters $\mathcal{I}$ *and* $(m_{\boldsymbol{\ell}}) \subset \mathbb{N}$ *s.t.*

$$\|\mu_{\mathrm{MI}} - \mathbb{E}[\Psi(X(T))]\|_{L^2(\Omega,U)}^2 \lesssim \varepsilon^2$$

with

$$\mathrm{Cost}(\mu_{\mathrm{MI}}) \eqsim \varepsilon^{-2-2u} |\log \varepsilon^{-1}|^r$$

for some $0 < r \leq 4 + 2u$ *and*

$$2u = 2 \max\left(0, \frac{1}{\beta_1 + \vartheta} - 1, \frac{1}{\beta_2} - 1, \frac{\upsilon + 1}{\upsilon \beta_1 + \beta_2} - 1\right)$$

Time-discretization

- Applied to the Zakai equation discretized by finite differences and implicit Milstein, MIMC does not outperform MLMC[6].
- Analysis was done using Fourier analysis. Complexity was shown to be improved depending on the used discretisation.
- We also started with applying MIMC for the SPDE's considered here with Backward Euler, and did not see an improvement of MIMC over MLMC (one has to trade spatial regularity for temporal regularity).
- Need to choose a suitable discretization scheme!

[6] Reisinger, C. & Wang, Z. Analysis of multi-index Monte Carlo estimators for a Zakai SPDE. English. *J. Comput. Math.* **36,** 202–236 (2018).

Overview

Accelerated exponential integrator method[7]

A *spectral numerical method* employing a semigroup shift.
Let $T > 0$, $M \in \mathbb{N}$ and $P_N : H \to \mathrm{Span}(e_1, \ldots, e_N) =: H_N$ be orthogonal projection.

$X_N^M : [0, T] \to H_N$ given by

$$X_N^M(t) = e^{-At} P_N X_0 + \int_0^t e^{-A(t-s)} P_N F(X_N^M(\lfloor s \rfloor_{M^{-1}}))\, \mathrm{d}s + \int_0^t e^{-A(t-s)} P_N(I + G X_N^M(\lfloor s \rfloor_{M^{-1}})) dW(s)$$

where $\lfloor t \rfloor_{M^{-1}} = \max\{ jT/M \,:\, jT/M < t, j = 0, 1, 2, \ldots, M \}$.
Note: When Q and A share an eigenbasis and when F is nonlinear, iterations can often be solved using FFT at an additional log-cost.

[7] Jentzen, A. & Kloeden, P. E. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. English. *Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci.* **465,** 649–667 (2009).

Assumptions on X_0, Q, G

For a *regularity parameter* $\kappa \in (0,2)$:

1. $X_0 \in L^{10}(\Omega, \dot{H}^{\kappa})$ is $\mathcal{F}_0$-measurable,
2. $Q^{1/2}(H) \overset{\mathcal{L}_2}{\hookrightarrow} \dot{H}^{\kappa-1}$, i.e., $\sum_j \lambda_j^{\kappa-1}\mu_j < \infty$
3. G is a linear operator on the spectrum of A:

$$(Gu)v = \sum_{j=1}^{\infty} \zeta_j \langle u, e_{j+m}\rangle \langle v, e_j\rangle e_j,$$

for a shift $m \in \mathbb{N}_0$ and a sequence $(\zeta_j)_{j=1}^{\infty} \subset \mathbb{R}$ fulfilling

$$|\zeta_j|\mu_j^{1/2} \leq C\lambda_j^{\frac{1-\kappa-\delta}{2}},$$

for some $1 \gg \delta > 0$, and ensuring that $G \in \mathcal{L}(H, \mathcal{L}_2(Q^{1/2}(H), \dot{H}^{\kappa+\delta-1}))$. When $m \neq 0$, an Euler-Maruyama scheme is *not* equivalent to a Milstein one (otherwise we would obtain improved convergence rate w.r.t. M).

Assumptions on F

Original assumptions[8]: $F\colon H \to H$ is twice Fréchet differentiable and

$$\|F'(u)v\|_{\dot H^{-r}} \le C\|v\|_{\dot H^{-r}}, \quad \text{for } r \in \{0,1,2\} \text{ and all } u,v \in H$$
$$\|F''(u)(v,w)\|_{\dot H^{-2}} \le C\|v\|_{\dot H^{-1}}\|w\|_{\dot H^{-1}}, \text{ for all } u,v \in H$$

for C ind. of u. If F is a composition operator in $L^2(\mathcal{D})$, *this implies linearity*.

Instead: Let $F \in \mathcal{G}^1(H,H) \cap \mathcal{G}^2(H,\dot H^{-\eta})$ for some $\eta \in [0,2)$.

1. $\|F'(u)v\| \le C\|v\|$ for all $u,v \in H$,
2. $\|F(u)-F(v)\|_{\dot H^{\kappa}} \le C(1+\|u\|^2_{\dot H^{\kappa}} + \|v\|^2_{\dot H^{\kappa}})$ for all $u,v \in \dot H^{\kappa}$,
3. $\|F'(u)v\|_{\dot H^{-\eta}} \le C(1+\|u\|^{\lceil\kappa\rceil}_{\dot H^{\kappa}})\|v\|_{\dot H^{-\kappa}}$ for all $u \in \dot H^{\kappa}, v \in H$ and
4. $\|F''(u)(v_1,v_2)\|_{\dot H^{-\eta}} \le C\|v_1\|\|v_2\|$ for all $u,v_1,v_2 \in H$.

[8] Jentzen, A. & Kloeden, P. E. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. English. *Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci.* **465,** 649–667 (2009).

Computation

The exponential integrator approximation X_N^M of the SPDE is now given by $X_N^M(0) := P_N X_0$ and for $j \in \{0, \dots, M-1\}, \Delta t = T/M, t_j = j\Delta t$, by

$$X_N^M(t_{j+1}) := e^{-A\Delta t} X_N^M(t_j) + \int_{t_j}^{t_{j+1}} e^{-A(t_{j+1}-s)} P_N F(X_N^M(t_j)) \, \mathrm{d}s + \int_{t_j}^{t_{j+1}} e^{-A(t_{j+1}-s)} \big(P_N + G X_N^M(t_j)\big) \, \mathrm{d}W(s).$$

Note that

$$\Big\langle \int_{t_j}^{t_{j+1}} e^{-A(t_{j+1}-s)} \big(P_N + G X_N^M(t_j)\big) \, \mathrm{d}W(s), e_k \Big\rangle = \mu_k^{\frac{1}{2}} (1 + \alpha_k \langle X_N^M(t_j), e_{k+m} \rangle) \int_{t_j}^{t_{j+1}} e^{-\lambda_k (t_{j+1}-s)} \, \mathrm{d}B_k(s)$$

if $k \leq N$ and 0 otherwise, meaning the stochastic term can be sampled exactly.

Regularity and error estimates

Recall that $\lambda_k \eqsim k^\nu$ for some $\nu > 0$. Given the assumptions above:

Theorem

For $\bar{N} \geq N$, $\bar{M} \geq M$ and a known $C_N, C > 0$

1. $\sup_t \|X_N^M(t)\|^2_{L^p(\Omega,\dot{H}^\kappa)} < \infty$
2. $\sup_t \|X_{\bar{N}}^M(t) - X_N^M(t))\|^2_{L^p(\Omega,H)} \leq C\lambda_{N+1}^{-\kappa} \eqsim N^{-\nu\kappa}$,
3. $\sup_t \|X_N^{\bar{M}}(t) - X_N^M(t)\|^2_{L^p(\Omega,H)} \leq \min(CM^{-\min(\kappa,1)}, C_N M^{-1})$ *and*
4. $$\sup_t \|X_{\bar{N}}^{\bar{M}}(t) - X_N^{\bar{M}}(t) - X_{\bar{N}}^M(t) + X_N^M(t)\|^2_{L^2(\Omega,H)}$$ $$\leq C\lambda_N^{-\kappa} \begin{cases} M^{-\kappa} & \kappa \in (0,1/2) \\ M^{-\kappa}\min(M^{\kappa-1}\lambda_N^{1-\kappa}, 1) & \kappa \in [1/2,1) \\ M^{-1} & \kappa \in [1,2) \end{cases}$$

yields $$\mathbb{E}[\|\Delta_\ell \Psi(X)\|_U^2] \lesssim \begin{cases} 2^{-\kappa\,\ell_1 - \kappa\nu\,\ell_2} & \kappa \in (0,1/2) \\ 2^{-\kappa\,\ell_1 - \kappa\nu\,\ell_2 - (1-\kappa)\max(\ell_1 - \nu\ell_2, 0)} & \kappa \in [1/2,1) \\ 2^{-\ell_1 - \kappa\nu\,\ell_2} & \kappa \in [1,2) \end{cases}$$

Proof ideas

$$
\begin{aligned}
&X_{\bar{N}}^{\bar{M}}(t) - X_{N}^{\bar{M}}(t) - X_{\bar{N}}^{M}(t) + X_{N}^{M}(t) \\
&\quad = e^{-At}(P_{\bar{N}} - P_N - P_{\bar{N}} + P_N)X_0 \\
&\qquad + \int_0^t e^{-A(t-s)}[\text{second order difference for } PF](s)\,\mathrm{d}s \\
&\qquad + \int_0^t e^{-A(t-s)}[\text{second order difference for } PG](s)\,\mathrm{d}W(s) \\
&\qquad + \int_0^t e^{-A(t-s)}(P_{\bar{N}} - P_N - P_{\bar{N}} + P_N)\,\mathrm{d}W(s).
\end{aligned}
$$

Neither the initial term nor the additive stochastic term directly contribute to either the spatial or the temporal part of the error!

Proof ideas

Focus on PF:

$$
\begin{aligned}
&P_{\bar N}F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - P_N F(X_N^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - P_{\bar N}F(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}})) + P_N F(X_N^{M}(\lfloor s\rfloor_{M^{-1}}))\\
&= P_{\bar N}(I - P_N)(F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - F(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}})))\\
&+ P_N \left(F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{M^{-1}})) - F(X_N^{\bar M}(\lfloor s\rfloor_{M^{-1}})) - F(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}})) + F(X_N^{M}(\lfloor s\rfloor_{M^{-1}}))\right)\\
&+ P_N \left(F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - F(X_N^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{M^{-1}})) + F(X_N^{\bar M}(\lfloor s\rfloor_{M^{-1}}))\right)
\end{aligned}
$$

Then

$$
\begin{aligned}
&\left\| \int_0^t e^{-A(t-s)} P_{\bar N}(I - P_N)(F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - F(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}})))ds \right\|_{L^2(\Omega,H)}\\
&\lesssim \int_0^t \|(I - P_N)A^{-\frac{\kappa}{2}}\|_{\mathcal{L}(H)} \|A^{\frac{\kappa}{2}} e^{-A(t-s)}\|_{\mathcal{L}(H)} \,\|(F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{\bar M^{-1}})) - F(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}})))\|_{L^2(\Omega,H)}
\end{aligned}
$$

Here

$$
\|(I - P_N)A^{-\frac{\kappa}{2}}\|_{\mathcal{L}(H)} = \|(I - P_N)\|_{\mathcal{L}(\dot H^{\kappa},H)} \le \lambda_{N+1}^{-\kappa/2}
$$

and critically

$$
\|A^{\frac{\kappa}{2}} e^{-A(t-s)}\|_{\mathcal{L}(H)} = \|e^{-A(t-s)}\|_{\mathcal{L}(H,\dot H^{\kappa})} \lesssim (t-s)^{-\kappa/2}
$$

Proof ideas

Repeated use of the mean value theorem, e.g.,

$$
\begin{aligned}
&F(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{M^{-1}})) - F(X_{N}^{\bar M}(\lfloor s\rfloor_{M^{-1}})) - F(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}})) + F(X_{N}^{M}(\lfloor s\rfloor_{M^{-1}})) \\
&\quad = \int_0^1 F'(\dots)\left(X_{\bar N}^{\bar M}(\lfloor s\rfloor_{M^{-1}}) - X_{N}^{\bar M}(\lfloor s\rfloor_{M^{-1}}) - X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}}) + X_{N}^{M}(\lfloor s\rfloor_{M^{-1}})\right)\,\mathrm{d}\lambda \\
&\qquad + \int_0^1\int_0^1 F''(\dots)\big(X_{N}^{\bar M}(\lfloor s\rfloor_{M^{-1}}) - X_{N}^{M}(\lfloor s\rfloor_{M^{-1}})\big)\big(X_{\bar N}^{M}(\lfloor s\rfloor_{M^{-1}}) - X_{N}^{M}(\lfloor s\rfloor_{M^{-1}})\big)\,\mathrm{d}\lambda\,\mathrm{d}\tilde\lambda,
\end{aligned}
$$

Using single-difference bounds, and use BDG inequality to deal with G, conclude with Grönwall inequality.

We derive sharper rates for $\mathbb{E}[\|\Delta_\ell\Psi(X)\|_U^2]$ when Ψ is linear.

Summary of Computational Complexities

For different Monte Carlo based methods, the cost is $\mathcal{O}(\varepsilon^{-2-2u}\,|\log(\varepsilon^{-1})|^r)$

- For Monte Carlo, $r = 0$ and

$$2u = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} \leq 2\left(\frac{1}{\min(1,\kappa)} + \frac{1}{\kappa\nu}\right)$$

- For MLMC, $r \leq 2 + 2u$ and

$$2u = \max\left(0, \frac{1 + \min(1,2\kappa)/(\kappa\nu) - \min(1,2\kappa)}{\min(\alpha_1, \alpha_2 \min(1,2\kappa)/(\kappa\nu))}\right)$$
$$\leq 2\max\left(0, \frac{1}{\min(1,\kappa)} + \frac{1}{\kappa\nu} - 1\right)$$

- For MIMC

$$2u = \max\left(0, \frac{1-\min(\kappa,1)}{\alpha_1}, \frac{1-\kappa\nu}{\alpha_2}\right) \leq 2\max\left(0, \frac{1}{\min(1,\kappa)} - 1, \frac{1}{\kappa\nu} - 1\right)$$

and $r \leq 4 + 2u$. Moreover, when $\kappa \in [1,2)$,

$$r = \begin{cases} 2 & \text{if} \quad \kappa\nu > 1 \\ 4 & \text{if} \quad \kappa\nu = 1\,. \end{cases}$$

Summary of Computational Complexities

For different Monte Carlo based methods, the cost is $\mathcal{O}(\varepsilon^{-2-2u}\,|\log(\varepsilon^{-1})|^r)$

- For Monte Carlo, $r = 0$ and

$$2u = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} \leq 2\left(\frac{1}{\min(1,\kappa)} + \frac{1}{\kappa\nu}\right)$$

- For MLMC, $r \leq 2 + 2u$ and

$$2u = \max\left(0, \frac{1 + \min(1,2\kappa)/(\kappa\nu) - \min(1,2\kappa)}{\min(\alpha_1, \alpha_2 \min(1,2\kappa)/(\kappa\nu))}\right)$$

$$\leq 2\max\left(0, \frac{1}{\min(1,\kappa)} + \frac{1}{\kappa\nu} - 1\right)$$

- For MIMC when $\kappa \in [1/2, 1]$

$$2u \leq 2\max\left(0, \frac{1}{\kappa\nu} - 1, \frac{\nu+1}{\nu+\nu\kappa} - 1\right)$$

and $r \leq 4 + 2u$. Moreover, when $\kappa \in [1, 2)$,

$$r = \begin{cases} 2 & \text{if} \quad \kappa\nu > 1 \\ 4 & \text{if} \quad \kappa\nu = 1\,. \end{cases}$$

Overview

Verification of multiplicative convergence rate

Test for QoI $\Psi(x) = x$ and seek to verify multiplicative convergence property

$$\|\Delta_{\boldsymbol{\ell}} X\|_{L^2(\Omega,H)} \approx \sqrt{E_{M=10^4}[\|\Delta_{\boldsymbol{\ell}} X\|^2]} =: e(\ell_1, \ell_2)$$

with $M_\ell \eqsim N_\ell \eqsim 2^\ell$.
When $\kappa \geq 1$, our sharp theoretical rates are:

$$\|\Delta_{\boldsymbol{\ell}} X\|^2_{L^2(\Omega,H)} \approx C N_{\ell_2}^{-\nu\kappa} \min(M_{\ell_1}^{-1}, N_{\ell_2}^{-\nu\kappa})$$

Numerical verification: Find $\beta_1, \beta_2 > 0$ by a least square fit, such that

$$p^2(\ell_1, \ell_2) := C \min(2^{-\beta_1 \ell_1 - \beta_2 \ell_2}, 2^{-2\beta_2 \ell_2})$$

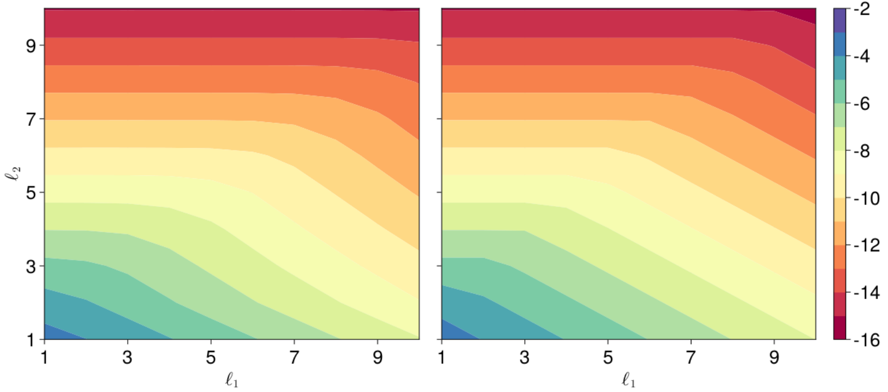
dominates $e(\ell_1, \ell_2)$ and verify that $\beta_1 \approx 1, \beta_2 \approx \nu\kappa$, when $\kappa \approx 1$ and $\nu = 4/3$.

For plotting, $\log_2(p(\ell_1, \ell_2))$ for a product $p(\ell_1, \ell_2)$ would be a plane over (ℓ_1, ℓ_2).

Note: Monte Carlo cost is $\mathcal{O}(\varepsilon^{-5.5})$, MLMC cost is $\mathcal{O}(\varepsilon^{-3.5})$.

Numerical test I

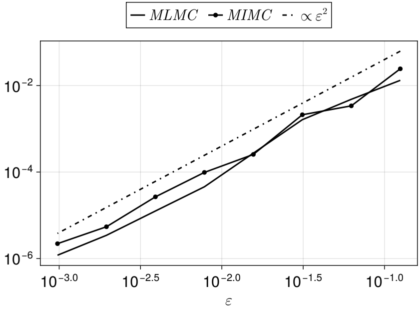
SPDE with $A = 0.2(-\Delta)^{2/3}$ on $\mathcal{D} = (0,1)$. Choose Q such that $\kappa < 1.01$ and let $f(x) = x$. Plot e and p in loglog scale.

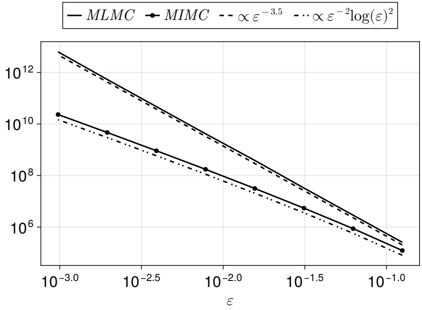


Left: $e(\ell_1, \ell_2)$. **Right:** $p(\ell_1, \ell_2)$ with $\beta_1 = 0.98$, $\beta_2 = 1.62$.
Expected rates: $\beta_1 = 1$ and $\beta_2 = 4/3$.

Numerical test II

Performance comparison MLMC vs MIMC, linear case.

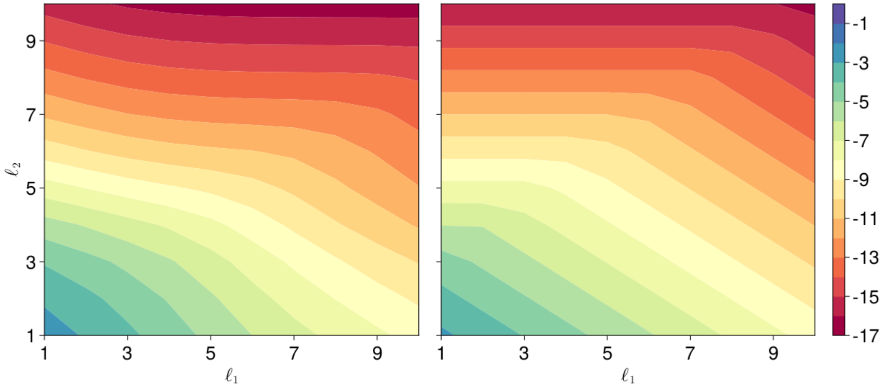




Left: Error $\|\mu_{\mathsf{MI}}(\varepsilon) - \mathbb{E}[X(1)]\|^2$. **Right:** $\mathrm{Cost}(\mu_{\mathsf{MI}}(\varepsilon))$

Numerical test III

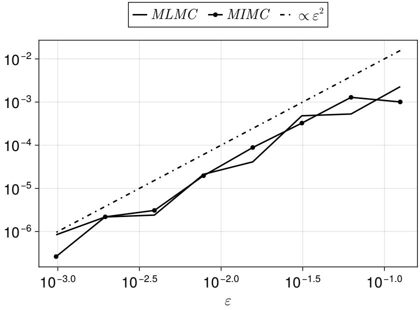
SPDE with $A = (-\Delta)^{2/3}$ on $\mathcal{D} = (0, 1)$. Choose Q such that $\kappa < 1.01$ and let $f(x) = \sin(\pi x)$.

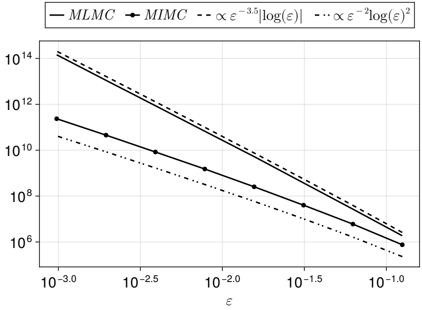


Left: $e(\ell_1, \ell_2)$. **Right:** $p(\ell_1, \ell_2)$ with $\beta_1 = 1.24$, $\beta_2 = 1.74$.
Expected rates: $\beta_1 = 1$ and $\beta_2 = 4/3$.

Numerical test IV

Performance comparison MLMC vs MIMC, nonlinear case.





Left: Error $\|\mu_{\mathsf{MI}}(\varepsilon) - \mathbb{E}[X(1)]\|^2$. **Right:** $\mathrm{Cost}(\mu_{\mathsf{MI}}(\varepsilon))$

Summary and future work

- Developed efficent multi-index Monte Carlo method for approximations of semilinear SPDE
- We obtain high convergence in space and can handle sufficiently differentiable composition mappings F
- Restriction: Operators Q and A have to share eigenbasis (e_k) on which G acts
- Future work: Extension to finite element exponential integrators
- Future work: Nonlinear G acting on the eigenbasis
- Future work: Sharp rates in time for $G = 0$ using stochastic sewing[9]

[9] Djurdjevac, A., Gerencsér, M. & Kremp, H. Higher order approximation of nonlinear SPDEs with additive space-time white noise. *arXiv preprint arXiv:2406.03058.* arXiv: 2406.03058 [math.PR]. https://doi.org/10.48550/arXiv.2406.03058 (June 2024).

References

- Haji-Ali, A.-L., Nobile, F. & Tempone, R. Multi-index Monte Carlo: when sparsity meets sampling. *Numerische Mathematik* **132,** 767–806 (2016)
- Jentzen, A. & Kloeden, P. E. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. English. *Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci.* **465,** 649–667 (2009)
- Chada, N. K., Hoel, H., Jasra, A. & Zouraris, G. E. Improved efficiency of multilevel Monte Carlo for stochastic PDE through strong pairwise coupling. *J. Sci. Comput.* **93,** Paper No. 62, 29 (2022)

- **Soon:** Haji-Ali, A.-L., Hoel, H. & Petersson, A. *The multi-index Monte Carlo method for semilinear stochastic partial differential equations.* 2025. arXiv: ? [math.NA]

Other notions of solutions

An H-valued predictable process $\{X(t)\}_{t\in[0,T]}$ is called:

- a **strong solution** of the SPDE if for all $t \in (0, T]$,

$$X(t) = X_0 + \int_0^t F(X(s)) - AX(s)\,\mathrm{d}s + \int_0^t (I + GX(s))\,\mathrm{d}W(s).$$

  Problem: Need that $X \in \mathrm{Dom}(A)$, and often it is not that smooth.

- a **weak solution** of the SPDE if for all $t \in (0, T]$ and $v \in \mathrm{Dom}(A)$

$$\begin{aligned}\langle X(t), v\rangle =&\langle X_0, v\rangle + \int_0^t \langle F(X(s)), v\rangle - \langle X(s), Av\rangle\,\mathrm{d}s \\ &+ \int_0^t \langle (I + GX(s))\,\mathrm{d}W(s), v\rangle.\end{aligned}$$

Relationship[10]**:** Strong solutions are weak solutions and weak solutions are typically mild solutions.

[10] Liu, W. & Röckner, M. *Stochastic partial differential equations: an introduction.* (Springer, 2015).

Motivation

Regularity: The semigroup e^{-At} is smoothing:

$$\|e^{-At}\|_{\mathcal{L}(H)} = \|e^{-At}e_1\| = e^{-\lambda_1 t} < 1$$

So that

$$\|e^{-At}X_0\|_{\dot{H}^1} = \|A^{1/2}e^{-At}X_0\| \leq \|e^{-At}A^{1/2}X_0\| \leq \|A^{1/2}X_0\|.$$

And $\|A^{1/2}e^{-At}\|_{\mathcal{L}(H)} \leq Ct^{-1/2}$ used to bound $\dot{H}^1$-norm of other terms, e.g.,

$$\begin{aligned}\int_0^t \|e^{-A(t-s)}P_N F(X_N^M(s))\|_{L^p(\Omega,\dot{H}^1)}\,\mathrm{d}s \\ &= \int_0^t \|A^{1/2}e^{-A(t-s)}P_N F(X_N^M(s))\|_{L^p(\Omega,H)}\,\mathrm{d}s \\ &\leq c\int_0^t (t-s)^{-1/2}\|P_N F(X_N^M(s))\|_{L^p(\Omega,H)}\,\mathrm{d}s\end{aligned}$$

Motivation

Numerical error: bound by

$$\|X(T) - X_N^M(T)\| \leq \underbrace{\|X(T) - P_N X(T)\|}_{\text{spatial error}} + \underbrace{\|P_N X(T) - X_N^M(T)\|}_{\text{time error}}.$$

Spatial error:

$$\|(I - P_N)X(T)\| = \|A^{-1/2}(I - P_N)A^{1/2}X(T)\| \leq \|A^{-1/2}(I - P_N)\|_{\mathcal{L}(H)}$$

and

$$\|A^{-1/2}(I - P_N)\|_{\mathcal{L}(H)} = \|A^{-1/2}(I - P_N)e_{N+1}\| = \lambda_{N+1}^{-1/2}$$

Time error:

$$\|P_N X(T) - X_N^M(T)\|_{L^p(\Omega,H)} = \mathcal{O}(\sqrt{\Delta t})$$

is same rate as Euler–Maruyama has for N-dimensional SDE.

Negative norm bounds

Negative norm bounds on F,

$$\|F'(u)v\|_{\dot H^{-\eta}} \le C(1+\|u\|_{\dot H^\kappa}^{\lceil r\rceil})\|v\|_{\dot H^{-r}}, \quad r\in\{1,\kappa-\delta\}, \quad u\in \dot H^\kappa, v\in H,$$

follow from a duality argument applied to

$$\|F'(u)v\|_{\dot H^{r}} \le C(1+\|u\|_{\dot H^\kappa}^{\lceil r\rceil})\|v\|_{\dot H^{\eta}}, \quad r\in\{1,\kappa-\delta\}, \quad u\in \dot H^\kappa, v\in \dot H^\eta$$

in the case that $F'(u)$ is symmetric on H.
Case that $A=-\Delta$ on $H=L^2(\mathcal{D})$ with zero Dirichlet b.c. and F is a composition mapping: Identify $\dot H^r$ with $W^{r,2}$ or $W_0^{r,2}$ and note $(F'(u)v)(\cdot)=f'(u(\cdot))v(\cdot)$.
Use Sobolev embedding and multiplication theorems with $\eta\in(d/2,2)$ to deduce

Lemma

For $\kappa\le\min(\eta,1)$ and f twice differentiable with bounded derivatives:

$$\|F'(u)v\|_{W^{\kappa,2}} \lesssim \|u\|_{W^{\kappa,2}}\|v\|_{W^{\eta,2}} \quad u\in W^{\kappa,2}, v\in W^{\eta,2}.$$

For $\kappa\in(1,\eta), d\le 2$ and f thrice differentiable with bounded derivatives:

$$\|F'(u)v\|_{W^{\kappa-\delta,2}} \lesssim (1+\|u\|_{W^{\kappa,2}}^2)\|v\|_{W^{\eta,2}} \quad u\in W^{\kappa,2}, v\in W^{\eta,2}, \delta>0.$$