Rapid covariance-based sampling for finite element approximations of linear SPDE in MLMC

Andreas Petersson

MCQMC 2018, Rennes, July 5, 2018

Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, Sweden

Introduction

Introduction to stochastic partial differential equations

Let $(\Omega, \mathcal{A}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ be a complete filtered probability space and let H be a real, separable Hilbert space. Consider the SPDE

$$\begin{aligned} \mathrm{d}X(t) &= (AX(t) + F(t, X(t)))\,\mathrm{d}t + G(t)\,\mathrm{d}W(t) \quad \text{for } t \in (0, T], \\ X(0) &= x_0. \end{aligned} \tag{1}$$

- $-A : \mathcal{D}(-A) \to H$ densely defined, self-adjoint and positive definite operator with a compact inverse $\implies$ A is the generator of a C_0-semigroup $E = (E(t), t \geq 0)$
- W is a Q-Wiener process in H
- $F(t, \cdot)$ is affine linear

Main idea

- **Goal:** Efficiently estimate $\mathbb{E}[\phi(X(T))]$ for a smooth functional ϕ.
- **Method:** Sample an approximation $X^T_{h,\Delta t}$ (with, e.g., finite elements and backward Euler) many times in a Monte Carlo simulation. Expensive!
- **Idea:** Consider only equations such that $X^T_{h,\Delta t}$ is *Gaussian*, compute its covariance and use this for sampling rather than sampling $X^T_{h,\Delta t}$ directly.

Outline

Introduction

Theoretical framework and discretization

Monte Carlo (MC) and multilevel Monte Carlo (MLMC)

Computing covariances and cross-covariances

Computational complexity and simulation

Model example: Stochastic advection-diffusion equation

Example

- $H = L^2(D)$, convex polygonal domain $D \subset \mathbb{R}^d$, $d = 1, 2, 3$
- $A = \Delta$ with Dirichlet zero boundary conditions
- $F(t, f) = b(t, \cdot) \cdot \nabla f(\cdot)$, $b : D \times [0, T] \to \mathbb{R}^d$ is a sufficiently smooth vector field
- $G = I$

The SPDE (1) is interpreted as

$$\begin{aligned}
\mathrm{d}X(t,x) &= (\Delta X(t,x) + b(t,x) \cdot \nabla X(t,x))\, \mathrm{d}t + \mathrm{d}W(t,x), && t \in (0,T], x \in D, \\
X(t,x) &= 0, && t \in (0,T], x \in \partial D, \\
X(0,x) &= x_0(x), && x \in D.
\end{aligned}$$

Model example: The Wiener process

Example

- $W(t) - W(s) \sim N(0, (t-s)Q)$ for a positive semidefinite trace class covariance operator $Q \in L(H)$
- $\mathrm{Tr}(Q) < \infty \implies Q^{1/2} \in \mathcal{L}_2(H) \implies$ there exists a symmetric square integrable function $q : D \times D \to \mathbb{R}$ such that

$$Qf = \int_D q(\cdot, y) f(y) \, \mathrm{d}y$$

- Similarly, a symmetric, positive semidefinite continuous function q on $D \times D$, defines a covariance operator Q.

Theoretical framework and discretization

Theoretical framework

Theoretical framework and discretization from [Kruse, 2014]:
$X = (X(t))_{t\in[0,T]}$ is a *mild solution*, i.e., $\sup_{t\in[0,T]} \|X(t)\|_{L^2(\Omega;H)} < \infty$ and for all $t \in [0,T]$,

$$X(t) = E(t)x_0 + \int_0^t E(t-s)F(s,X(s))\,\mathrm{d}s + \int_0^t E(t-s)G(s)\,\mathrm{d}W(s),$$

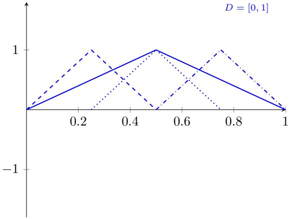
P-a.s.

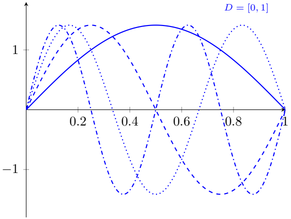
- $\dot{H}^r = \mathrm{dom}((-A)^{\frac{r}{2}})$ Hilbert space with norm $\|\cdot\|_r = \|A^{-\frac{r}{2}}\cdot\|$
- For the model example, $\dot{H}^1 = H_0^1(D)$ and $\dot{H}^2 = H^2(D) \cap H_0^1(D)$, where $H^k(D)$ is the Sobolev space of order k on D and $H_0^1(D)$ is the subset of functions that are zero on the boundary.

Spatial discretization

$(V_h)_{h\in(0,1]}$ is a family of subspaces of $\dot{H}^1$ equipped with the inner product of H such that $N_h = \dim(V_h) < \infty$, $h \in (0,1]$. Assume:

- $\|P_h x\|_1 \leq C\|x\|_1$ for all $h \in (0,1]$ and $x \in \dot{H}^1$
- $\|R_h x - x\| \leq Ch^s\|x\|_s$ for all $h \in (0,1]$ and x in $\dot{H}^s$, $s \in \{0,1\}$





$-A_h : V_h \to V_h$ is defined by the relationship

$$\langle -A_h f_h, g_h\rangle = \langle f_h, g_h\rangle_1 = \left\langle (-A)^{\frac{1}{2}} f_h, (-A)^{\frac{1}{2}} g_h \right\rangle$$

for all $f_h, g_h \in V_h$.

Full discretization

Backward Euler method: Let a uniform time grid be given by $t_j = j\Delta t$ for $j = 0, \ldots, N_{\Delta t}$, where $N_{\Delta t} \in \mathbb{N}$ and $\Delta t = T N_{\Delta t}^{-1}$. The approximation $(X_{h,\Delta t}^{t_j})_{j=0}^{N_{\Delta t}}$ of the SPDE (1) is given by the recursion

$$X_{h,\Delta t}^{t_{j+1}} - X_{h,\Delta t}^{t_j} = \left(A_h X_{h,\Delta t}^{t_{j+1}} + P_h F(t_j, X_{h,\Delta t}^{t_j}) \right) \Delta t + P_h G(t_j) \Delta W^j, \quad (2)$$

where $\Delta W^j = W(t_{j+1}) - W(t_j)$ and $j = 0, \ldots, N_{\Delta t} - 1$.

Assumptions and strong convergence

Assumption 1

The parameters of the SPDE (1) fulfill:

- W is an $(\mathcal{F}_t)_{t\in[0,T]}$-adapted Q-Wiener process with $\mathrm{Tr}(Q) < \infty$.
- $\exists C > 0$ s.t. $G : [0,T] \to \mathcal{L}_2^0 = \mathcal{L}_2(Q^{1/2}(H), H)$ satisfies
$$\|G(t_1) - G(t_2)\|_{\mathcal{L}_2^0} \le C|t_1 - t_2|^{1/2}, \forall t_1, t_2 \in [0,T].$$
- $F : [0,T] \times H \to \dot{H}^{-1}$ is affine in H, i.e., for each $t \in [0,T]$ there exists $F_t^1 \in \mathcal{L}(H, \dot{H}^{-1})$ and $F_t^2 \in \dot{H}^{-1}$ s.t. $F(t,f) = F_t^1 f + F_t^2, \forall f \in H$. Furthermore, $\exists C > 0$ s.t. F satisfies
$$\|F(t_1, f) - F(t_2, f)\|_{-1} \le C(1 + \|f\|)|t_1 - t_2|^{1/2}, \forall f \in H, t_1, t_2 \in [0,T],$$
$$\|F_t^1\|_{\mathcal{L}(H,\dot{H}^{-1})} \le C, \forall t \in [0,T].$$
- The initial value x_0 is a (possibly degenerate) $\mathcal{F}_0$-measurable $\dot{H}^1$-valued Gaussian random variable.

Assumptions and strong convergence

Existence, uniqueness of X and *strong convergence*:

Theorem (Kruse, 2014)

Under Assumption 1 : $\forall p \geq 1$, $\sup_{h,\Delta t}(\|X_{h,\Delta t}^T\|_{L^p(\Omega;H)}) < \infty$ *and* $\exists C > 0$, *s.t.* $\forall h, \Delta t \in (0,1]$

$$\|X(T) - X_{h,\Delta t}^T\|_{L^p(\Omega;H)} \leq C\left(h + \Delta t^{1/2}\right).$$

and $X_{h,\Delta t}^T$ is guaranteed to be Gaussian.

Assumptions and weak convergence

- The following assumption is only needed to tune the MC and MLMC estimator.

Assumption 2

The parameters of the SPDE (1) fulfill, for some $\delta \in [1/2, 1]$:

- $\exists C > 0$ s.t. $G : [0, T] \to \mathcal{L}_2^0$ satisfies

$$\|G(t_1) - G(t_2)\|_{\mathcal{L}_2^0} \leq C|t_1 - t_2|^{\delta}, \forall t_1, t_2 \in [0, T].$$

- $\exists C > 0$ s.t. $F : [0, T] \to H$ satisfies

$$\|F(t_1) - F(t_2)\| \leq C|t_1 - t_2|^{\delta}, \forall t_1, t_2 \in [0, T].$$

Furthermore, the functional $\phi : H \to \mathbb{R}$ is 2 times continuously Fréchet-differentiable with derivatives of polynomial growth.

Assumptions and weak convergence

Under these assumptions, we have *weak convergence*:

Theorem (Kruse, 2014)

Under Assumptions 1 and 2: there exists $C > 0$, *such that for all* $h, \Delta t \in (0, 1]$

$$\left|\mathbb{E}\left[\phi(X(T))\right] - \mathbb{E}[\phi(X^T_{h,\Delta t})]\right| \leq C\,(1 + |\log(h)|)\left(h^2 + \Delta t^{\delta}\right).$$

Monte Carlo (MC) and multilevel Monte Carlo (MLMC)

Monte Carlo (MC)

- **Goal:** Efficiently estimate $\mathbb{E}[\phi(X(T))]$
- Replace $\phi(X(T))$ with $\phi(X^T_{h,\Delta t})$, but how do we estimate $\mathbb{E}[\cdot]$?
- We denote the MC estimator by

$$E_N\left[\phi(X^T_{h,\Delta t})\right] = \frac{1}{N}\sum_{i=1}^{N}\phi(X^T_{h,\Delta t})^{(i)} \approx \mathbb{E}\left[\phi(X^T_{h,\Delta t})\right]$$

- Two algorithms: path-based and covariance-based

Two MC algorithms

For $X^T_{h,\Delta t} = \sum_{k=1}^{N_h} x_k \Phi_k^h$, where $\mathbf{\Phi}^h = (\Phi_k^h)_{k=1}^{N_h}$ is a basis of V_h, write $\bar{\mathbf{x}}_h^T = [x_1, x_2, \ldots, x_{N_h}]'$.

Algorithm 1 Path-based MC method of computing an estimate $E_N[\phi(X^T_{h,\Delta t})]$ of $\mathbb{E}[\phi(X(T))]$

$result = 0$

for $i = 1$ to N **do**

  Sample a realization $W^{(i)}$ of the Q-Wiener process W

  Compute $\bar{\mathbf{x}}_h^T = [x_1, x_2, \ldots, x_{N_h}]'$ by solving the matrix equations corresponding to the backward Euler scheme (2) driven by $W^{(i)}$

  Compute $\phi(X^T_{h,\Delta t}) = \phi\left(\sum_{k=1}^{N_h} x_k \Phi_k^h\right)$

  $result = result + \phi(X^T_{h,\Delta t})^{(i)}/N$

end for

$E_N\left[\phi(X^T_{h,\Delta t})\right] = result$

Algorithm 2 Covariance-based MC method of computing an estimate $E_N[\phi(X^T_{h,\Delta t})]$ of $\mathbb{E}[\phi(X(T))]$

```
Form the mean vector μ and covariance matrix Σ of x̄_h^T by finding the mean and covariance of X_{h,Δt}^T
result = 0
for i = 1 to N do
    Sample x̄_h^T = [x_1, x_2, ..., x_{N_h}]' ~ N(μ, Σ)
    Compute φ(X_{h,Δt}^T) = φ( Σ_{k=1}^{N_h} x_k Φ_k^h )
    result = result + φ(X_{h,Δt}^T)^{(i)} / N
end for
E_N[φ(X_{h,Δt}^T)] = result
```

Form the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of $\bar{\mathbf{x}}_h^T$ by finding the mean and covariance of $X^T_{h,\Delta t}$

$result = 0$

for $i = 1$ to N **do**

  Sample $\bar{\mathbf{x}}_h^T = [x_1, x_2, \ldots, x_{N_h}]' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

  Compute $\phi(X^T_{h,\Delta t}) = \phi\left(\sum_{k=1}^{N_h} x_k \Phi_k^h\right)$

  $result = result + \phi(X^T_{h,\Delta t})^{(i)}/N$

end for

$E_N\left[\phi(X^T_{h,\Delta t})\right] = result$

Multilevel Monte Carlo (MLMC)

- Due to [Giles, 2008], for SPDE [Barth & Lang, 2012] and [Barth, Lang & Schwab, 2013]
- A sequence $(X_\ell^T)_{\ell\in\mathbb{N}_0}$, indexed by levels $\ell \in \mathbb{N}_0$, of approximations by $X_\ell^T = X_{h_\ell,\Delta t_\ell}^T$
- $(h_\ell)_{\ell\in\mathbb{N}_0}$ and $(\Delta t_\ell)_{\ell\in\mathbb{N}_0}$ are decreasing sequences of mesh sizes and time step sizes
- MLMC estimator $E^L\left[\phi\left(X_L^T\right)\right]$ is given by

$$E^L[\phi(X_L^T)] = E_{N_0}[\phi(X_0^T)] + \sum_{\ell=1}^{L} E_{N_\ell}[\phi(X_\ell^T) - \phi(X_{\ell-1}^T)].$$

- Using weak and strong convergence rate results, the sample sizes $(N_\ell)_{\ell\in\mathbb{N}_0}$ can be chosen so that E^L outperforms E_N

Two MLMC algorithms

We set $X_{-1}^T = 0$.

Algorithm 3 Path-based MLMC method of computing an estimate $E^L\left[\phi\left(X_L^T\right)\right]$ of $\mathbb{E}[\phi(X(T))]$ `[Barth & Lang, 2012]`

1: $result = 0$
2: **for** $\ell = 0$ to L **do**
3: **for** $i = 1$ to N_ℓ **do**
4: Sample a realization $W^{(i)}$ of the Q-Wiener process W
5: Compute $\bar{\mathbf{x}}_{h_{\ell-1}}^T$ and $\bar{\mathbf{x}}_{h_\ell}^T$ by solving the matrix equations corresponding to the backward Euler scheme (2) driven by $W^{(i)}$
6: $result = result + \left(\phi(X_\ell^T) - \phi(X_{\ell-1}^T)\right)/N_\ell$

$$= result + \left(\phi\Big(\sum_{k=1}^{N_{h_\ell}} x_k^\ell \Phi_k^{h_\ell}\Big) - \phi\Big(\sum_{k=1}^{N_{h_{\ell-1}}} x_k^{\ell-1}\Phi_k^{h_{\ell-1}}\Big)\right)\Big/N_\ell$$

7: **end for**
8: **end for**
9: $E^L\left[\phi(X_L^T)\right] = result$

Two MLMC algorithms

Algorithm 4 Covariance-based MLMC method of computing an estimate $E^L\left[\phi\left(X_L^T\right)\right]$ of $\mathbb{E}[\phi(X(T))]$

$result = 0$

for $\ell = 0$ to L **do**

  Compute the covariance matrix $\boldsymbol{\Sigma}_\ell$ and mean vector $\boldsymbol{\mu}_\ell$ of $\bar{\mathbf{x}}_\ell = \left[\left[\bar{\mathbf{x}}^T_{h_{\ell-1}}\right]', \left[\bar{\mathbf{x}}^T_{h_\ell}\right]'\right]'$ by computing the means, covariances and cross-covariance of the pair $(X^T_{\ell-1}, X^T_\ell)$

  for $i = 1$ to N_ℓ **do**

    Sample $\bar{\mathbf{x}}_\ell \sim N(\boldsymbol{\mu}_\ell, \boldsymbol{\Sigma}_\ell)$

    $$\begin{aligned} result &= result + \left(\phi(X_\ell^T) - \phi(X_{\ell-1}^T)\right)/N_\ell \\ &= result + \left(\phi\Big(\sum_{k=1}^{N_{h_\ell}} x_k^\ell \Phi_k^{h_\ell}\Big) - \phi\Big(\sum_{k=1}^{N_{h_{\ell-1}}} x_k^{\ell-1}\Phi_k^{h_{\ell-1}}\Big)\right)\Big/N_\ell \end{aligned}$$

  end for

end for

$E^L\left[\phi(X_L^T)\right] = result$

Computing covariances and cross-covariances

Covariance computation

- Abbreviations: $R_{h,\Delta t} = (I_H - \Delta t A_h)$, $F^{1,j}_{h,\Delta t} = \Big(I_H + \Delta t P_h F^1_{t_j}\Big)$, and $F^{2,j}_{h,\Delta t} = \Delta t P_h F^2_{t_j}$
- (2) becomes $R_{h,\Delta t} X^{t_{j+1}}_{h,\Delta t} = F^{1,j}_{h,\Delta t} X^{t_j}_{h,\Delta t} + F^{2,j}_{h,\Delta t} + P_h G(t_j) \Delta W^j$
- For a Hilbert space U we write $U^{\otimes 2} = U \otimes U$
- $\Sigma^T = \mathrm{Cov}(X^T_{h,\Delta t}) = \mathbb{E}[(X^T_{h,\Delta t})^{\otimes 2}] - \mathbb{E}[X^T_{h,\Delta t}]^{\otimes 2}$

Theorem

Under Assumption 1: the mean $\mu^T \in V_h$ and covariance $\Sigma^T \in V_h^{\otimes 2}$ of $X^T_{h,\Delta t}$ are given by the recursions

$$R_{h,\Delta t} \mu^{t_{j+1}} = F^{1,j}_{h,\Delta t} \mu^{t_j} + F^{2,j}_{h,\Delta t} \tag{3}$$

and

$$(R_{h,\Delta t})^{\otimes 2} \Sigma^{t_{j+1}} = (F^{1,j}_{h,\Delta t})^{\otimes 2} \Sigma^{t_j} + \mathbb{E}\left[\left(P_h G(t_j) \Delta W^j\right)^{\otimes 2}\right] \tag{4}$$

for $j = 0, 1, \ldots, N_{\Delta t} - 1$.

Proof sketch

- Based on an idea from stability analysis of SDE approximations (see e.g. [Buckwar & Sickenberger, 2012])
- Assume F linear for simplicity, i.e., $F_t^2 = 0$ for all $t \in [0, T]$
- $R_{h,\Delta t} X_{h,\Delta t}^{t_{j+1}} = F_{h,\Delta t}^{1,j} X_{h,\Delta t}^{t_j} + P_h G(t_j) \Delta W^j$
- $\left(R_{h,\Delta t}\right)^{\otimes 2} \left(X_{h,\Delta t}^{t_{j+1}}\right)^{\otimes 2}$

$$= \left(F_{h,\Delta t}^{1,j}\right)^{\otimes 2} \left(X_{h,\Delta t}^{t_j}\right)^{\otimes 2} + F_{h,\Delta t}^{1,j} X_{h,\Delta t}^{t_j} \otimes P_h G(t_j) \Delta W^j + P_h G(t_j) \Delta W^j \otimes F_{h,\Delta t}^{1,j} X_{h,\Delta t}^{t_j} + \left(P_h G(t_j) \Delta W^j\right)^{\otimes 2}$$

- $F_{h,\Delta t}^{1,j} X_{h,\Delta t}^{t_j}$ is $\mathcal{F}_{t_j}$-measurable, $P_h G(t_j) \Delta W^j$ is independent of $\mathcal{F}_{t_j}$
- $\mathbb{E}[F_{h,\Delta t}^{1,j} X_{h,\Delta t}^{t_j} \otimes P_h G(t_j) \Delta W^j]$

$$= \mathbb{E}[F_{h,\Delta t}^{1,j} X_{h,\Delta t}^{t_j}] \otimes \mathbb{E}[P_h G(t_j) \Delta W^j] = 0$$

- Tensorizing (3) and subtracting it from this then yields (4)

Cross-covariance computation

- $\Sigma_T = \text{Cov}(X^T_{h',\Delta t'}, X^T_{h,\Delta t})$
  $= \mathbb{E}[X^T_{h',\Delta t'} \otimes X^T_{h,\Delta t}] - \mathbb{E}[X^T_{h',\Delta t'}] \otimes \mathbb{E}[X^T_{h,\Delta t}]$
- Two approximations, $(X^{t'_j}_{h',\Delta t'})_{j=0}^{N_{\Delta t'}}$ and $(X^{t_j}_{h,\Delta t})_{j=0}^{N_{\Delta t}}$, with $\Delta t' = K\Delta t$ for some $K \in \mathbb{N}$

Create an extension $(\hat{X}^{t_j}_{h',\Delta t})_{j=0}^{N_{\Delta t}}$ of $(X^{t'_j}_{h',\Delta t'})_{j=0}^{N_{\Delta t'}}$ to the fine time grid by

$$\hat{R}^j_{h',\Delta t'} \hat{X}^{t_{j+1}}_{h',\Delta t} = \hat{F}^{1,j}_{h',\Delta t'} \hat{X}^{t_j}_{h',\Delta t} + \hat{F}^{2,j}_{h',\Delta t'} + P_{h'} \hat{G}(t_j) \Delta W^j,$$

for $j = 0, 1, \ldots, N_{\Delta t} - 1$. $\hat{X}^{t_{j+1}}_{h',\Delta t'} = X^{t'_{(j+1)/K}}_{h',\Delta t'}$ when $j+1 = 0 \mod K$ if

$$\hat{R}^j_{h',\Delta t'} = \begin{cases} R_{h',\Delta t'} & \text{if } j+1 = 0 \mod K, \\ I_H & \text{otherwise,} \end{cases} \quad \hat{F}^{1,j}_{h',\Delta t} = \begin{cases} F^{1,j/K}_{h',\Delta t'} & \text{if } j+1 = 1 \mod K, \\ 0 & \text{otherwise,} \end{cases}$$

$$\hat{F}^{2,j}_{h',\Delta t} = \begin{cases} F^{2,j/K}_{h',\Delta t'} & \text{if } j+1 = 1 \mod K, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and } \hat{G}(t_j) = G(t_{j-(j \bmod K)}).$$

Cross-covariance computation

Theorem

Under Assumption 1: Cov$(X^T_{h',\Delta t'}, X^T_{h,\Delta t})$ *is given by* $\Sigma_T \in V_{h'} \otimes V_h$ *where the sequence* $\{\Sigma_{t_j}\}_{j=0}^{N_{\Delta t}}$ *fulfills*

$$\left(\hat{R}^j_{h',\Delta t'} \otimes R_{h,\Delta t}\right)\Sigma_{t_{j+1}} = \left(\hat{F}^{1,j}_{h,\Delta t} \otimes F^{1,j}_{h,\Delta t}\right)\Sigma_{t_j} + \mathbb{E}\left[P_{h'}\hat{G}(t_j)\Delta W^j \otimes P_h G(t_j)\Delta W^j\right].$$

Computational complexity and simulation

MC complexity

- Tuning the MC estimator using weak convergence:

$$\begin{aligned}
&\| \mathbb{E}[\phi(X(T))] - E_N[\phi(X^T_{h,\Delta t})]\|_{L^2(\Omega;\mathbb{R})} \\
&\quad \le \| \mathbb{E}[\phi(X(T))] - \mathbb{E}[\phi(X^T_{h,\Delta t})]\|_{L^2(\Omega;\mathbb{R})} \\
&\qquad + \| \mathbb{E}[\phi(X^T_{h,\Delta t})] - E_N[\phi(X^T_{h,\Delta t})]\|_{L^2(\Omega;\mathbb{R})} \\
&\quad \le C\left((1 + |\log(h)|)\left(h^2 + \Delta t^\delta\right) + N^{-1/2}\right)
\end{aligned}$$

- Balance the error by $\Delta t^\delta \simeq N^{-1/2} \simeq h^2$, so that

$$\| \mathbb{E}[\phi(X(T))] - E_N[\phi(X^T_{h,\Delta t})]\|_{L^2(\Omega;\mathbb{R})} \le C\,(1 + |\log(h)|)\,h^2$$

MC complexity continued

Assumption 3

For some $d = 1, 2, 3$, the cost of solving (2) once to get $X_{h,\Delta t}^{t_{j+1}}$ given $X_{h,\Delta t}^{t_j}$ for $j = 0, \dots, N_{\Delta t} - 1$ is assumed to be $\mathrm{O}(h^{-d})$ while the cost of solving the corresponding tensorized system (4) once is $\mathrm{O}(h^{-2d})$. The cost of a sample of $X_{h,\Delta t}^{T}$ using Cov$(X_{h,\Delta t}^{T})$ is assumed to be $\mathrm{O}(h^{-2d})$.

Proposition

Under Assumptions 1 and 2: if $N \simeq h^{-4}$ and $\Delta t \simeq h^{2/\delta}$, then $\exists C > 0$, s.t. $\forall h > 0$,

$$\| \mathbb{E}[\phi(X(T))] - E_N[\phi(X_{h,\Delta t}^{T})] \|_{L^2(\Omega;\mathbb{R})} \leq C \, (1 + |\log(h)|) \, h^2.$$

The cost of the path-based method is bounded by $\mathrm{O}(h^{-4-d-2/\delta})$, the cost of the covariance-based method is bounded by $\mathrm{O}(h^{-2d-2/\delta}) + \mathrm{O}(h^{-2d-4}) = \mathrm{O}(h^{-2d-4})$.

MLMC complexity

For the computational complexity analysis for the MLMC we follow the approach of [Lang, 2016] with some minor modifications.

Proposition

Under Assumptions 1 and 2: Let $(h_\ell)_{\ell\in\mathbb{N}_0}$ *be a sequence that satisfies* $h_\ell \simeq a^{-\ell}$ *for some* $a \in \mathbb{R}$ *and all* $\ell \in \mathbb{N}_0$*. Let* $(X_\ell^T)_{\ell\in\mathbb{N}_0}$ *be given by* $X_\ell^T = X_{h_\ell,\Delta t_\ell}^T$ *with* $\Delta t_\ell^\delta \simeq h_\ell^2$*.*

For $L \in \mathbb{N}, \ell = 1, \ldots, L$, $\epsilon > 0$*, set* $N_\ell = \lceil h_L^{-4} h_\ell^2 \ell^{1+\epsilon} \rceil$*, and* $N_0 = \lceil h_L^{-4} \rceil$*. Then* $\exists C > 0$ *s.t.* $\forall L \in \mathbb{N}$

$$\| \mathbb{E}\left[\phi\left(X(T)\right)\right] - E^L\left[\phi\left(X_L^T\right)\right] \|_{L^2(\Omega;\mathbb{R})} \leq C(1 + |\log(h_L)|)h_L^2.$$

The cost of finding $E^L\left[\phi\left(X_L^T\right)\right]$ *with the path-based method is bounded by* $\mathrm{O}(h_L^{-2-d-2/\delta} L^{2+\epsilon})$ *and with the covariance-based method by* $\mathrm{O}(\max(h_L^{-2d-2/\delta} L, h_L^{-2-2d} L^{2+\epsilon}))$*.*

Complexity remarks

	MC	MLMC
Path	$\mathrm{O}(h^{-4-d-2/\delta})$	$\mathrm{O}(h_L^{-2-d-2/\delta} L^{2+\epsilon})$
Cov	$\mathrm{O}(h^{-2d-4})$	$\mathrm{O}(\max(h_L^{-2d-2/\delta} L, h_L^{-2-2d} L^{2+\epsilon}))$

- $\delta = 1/2 \implies$ Cov + MC should be used for $d = 1, 2$, for $d = 3$ use Path + MLMC
- $\delta = 1 \implies$ Cov + MLMC should be used for $d = 1, 2$, for $d = 3$ use Path + MLMC

- Consider the stochastic heat equation driven by additive (i.e., $G = I$) Wiener noise

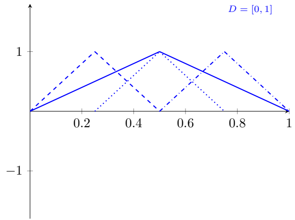
$$\begin{aligned} \mathrm{d}X(t) &= \Delta X(t)\,\mathrm{d}t + \mathrm{d}W(t), t \in (0,T] \\ X(0) &= x_0 \end{aligned}$$

  on $H = L^2(D)$ with $D = (0,1)$, for $t \in (0,T] = (0,1]$ with deterministic initial value $X(0) = x_0 = \sin(2\pi\cdot)$ and Dirichlet zero boundary conditions.

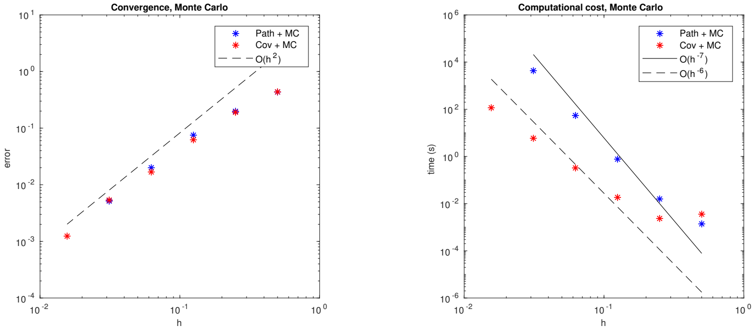
- The covariance function of W is for $x, y \in D$ is chosen to be $q(x,y) = 20\exp(-2|x-y|)$.
- $\phi(\cdot) = \|\cdot\|^2$.

Simulation, spatial discretization

- Let $(T_h)_{h\in(0,1]}$ be a family of uniform triangulations of $(0,1)$ with h being the mesh size and let V_h be the space of all functions that are continuous and piecewise linear on T_h and zero at the boundary of $(0,1)$.
- We use $\mathbf{\Phi}^h = (\Phi_i^h, i = 1, 2, \ldots, N_h)$, the standard basis of hat functions, on V_h and for $V_h^{\otimes 2}$ we use $\mathbf{\Phi}^{2,h} = (\Phi_i^{2,h}, i = 1, 2, \ldots, N_h^2)$ with $\Phi_i^{2,h} = \Phi^h_{\lfloor (i-1)/N_h \rfloor + 1} \otimes \Phi^h_{i - \lfloor (i-1)/N_h \rfloor N_h}$ for $i = 1, 2, \ldots, N_h^2$.



Simulation results, MC



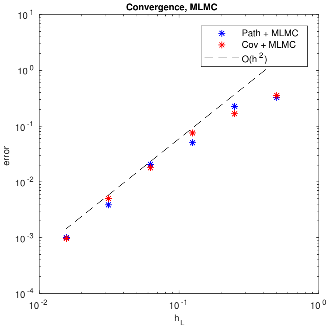
(a) MSE for MC.

(b) Computational costs in seconds.

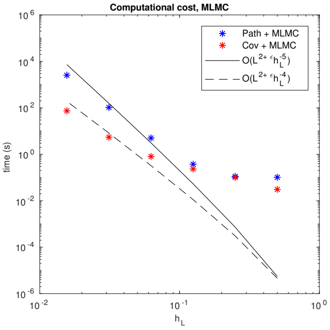
Figure: Convergence and computational costs of the MC estimator.

- Plotting the errors $\| \mathbb{E}[\phi(X(T))] - E_N[\phi(X^T_{h,\Delta t})]\|_{L^2(\Omega;\mathbb{R})}$ for $h = 2^{-1}, \dots, 2^{-6}$.

Simulation results, MLMC



(a) MSE for MLMC.



(b) Computational costs in seconds.

Figure: Convergence and computational costs of the MLMC estimator.

- Plotting the errors $\|\mathbb{E}[\phi(X(T))] - E^L[\phi(X_L^T)]\|_{L^2(\Omega;\mathbb{R})}$ for $h_L = 2^{-1}, \ldots, 2^{-6}$.

Concluding remarks

- Covariance-based sampling can significantly reduce the computational complexity of Monte Carlo methods for parabolic SPDE, and in MLMC it does not require nesting of $(V_h)_{h\in(0,1]}$.
- The sampling requires a Gaussian approximation, but the computation of the covariance can be done for more general SPDEs (affine linear multiplicative noise, Lévy noise...) and temporal approximations (general rational approximations of the semigroup).
- Future work: Consider hyperbolic equations and alternative methods of sampling given the covariance.
- Details at `arXiv:1806.11523 [math.PR]`.
- Thank you for listening!