CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN CALCUL SCIENTIFIQUE

Multilevel Monte Carlo methods for uncertainty quantification

application to sensistivity analysis

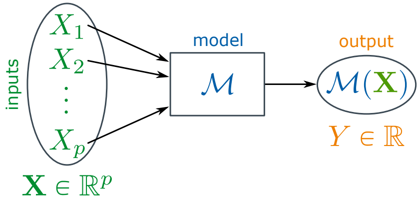
P. Mycek

Cerfacs, Toulouse, France.

RMR 2019 — June 19, 2019

Uncertainty quantification (UQ)

Simplified UQ framework:



The parameters $X_1, \ldots, X_p$ are *uncertain*

$\rightarrow$ modeled by random variables (prescribed probability distribution).

$\Rightarrow Y = \mathcal{M}(\mathbf{X})$ is a random variable (unknown distribution).

How uncertainty in the inputs impacts uncertainty in the output?

- Compute statistics of the output: $\mathbb{E}[Y]$, $\mathbb{V}[Y]$, $\ldots$
- Sensitivity analysis: which input parameters induce the most variability in Y?

Outline

Hoeffding decomposition / ANOVA

If the random variables $(X_i)_{i=1}^p$ are independent and $\mathbb{E}[\mathcal{M}(\mathbf{X})^2] < \infty$,

$$Y = \mathcal{M}(\mathbf{X}) = \mathcal{M}_\varnothing + \sum_{1 \leqslant i \leqslant p} \mathcal{M}_i(X_i) + \sum_{1 \leqslant i < j \leqslant p} \mathcal{M}_{ij}(\mathbf{X}_{ij}) + \ldots + \mathcal{M}_{12\ldots p}(\mathbf{X})$$

with $\mathcal{M}_\varnothing = \mathbb{E}[\mathcal{M}(\mathbf{X})]$, $\quad \mathbb{E}[\mathcal{M}_u(\mathbf{X}_u)] = 0$, $\quad \mathbb{E}[\mathcal{M}_u(\mathbf{X}_u)\mathcal{M}_{v \neq u}(\mathbf{X}_{v \neq u})] = 0.$

Variance decomposition

$$\mathbb{V}[Y] = \sum_{1 \leqslant i \leqslant p} \mathbb{V}[\mathcal{M}_i(X_i)] + \sum_{1 \leqslant i < j \leqslant p} \mathbb{V}[\mathcal{M}_{ij}(\mathbf{X}_{ij})] + \ldots + \mathbb{V}[\mathcal{M}_{12\ldots p}(\mathbf{X})]$$

Sobol' decomposition

$$1 = \sum_{1 \leqslant i \leqslant p} \underbrace{\frac{\mathbb{V}[\mathcal{M}_i(X_i)]}{\mathbb{V}[Y]}}_{S_i} + \sum_{1 \leqslant i < j \leqslant p} \underbrace{\frac{\mathbb{V}[\mathcal{M}_{ij}(\mathbf{X}_{ij})]}{\mathbb{V}[Y]}}_{S_{ij}} + \ldots + \underbrace{\frac{\mathbb{V}[\mathcal{M}_{12\ldots p}(\mathbf{X})]}{\mathbb{V}[Y]}}_{S_{12\ldots p}}$$

Sobol' indices

Goal: measuring the dependence (X_i, Y) by **shares of output variance**

- **first order** Sobol' indices:

$$S_i = \frac{\text{Variance of } Y \text{ due uniquely to } X_i}{\text{Variance of } Y}$$

- **second order** Sobol' indices:

$$S_{ij} = \frac{\text{Variance of } Y \text{ due to the interaction } (X_i, X_j)}{\text{Variance of } Y}$$

- **third order** Sobol' indices:

$$S_{ijk} = \frac{\text{Variance of } Y \text{ due to the interaction } (X_i, X_j, X_k)}{\text{Variance of } Y}$$

- **total order** Sobol' indices:

$$S_i^T = \frac{\text{Variance of } Y \text{ due to } X_i \text{ and to all its interactions}}{\text{Variance of } Y}$$

Properties and remarks

A few remarks:

- All the indices add up to 1:

$$\sum_{1 \leqslant i \leqslant p} S_i + \sum_{1 \leqslant i < j \leqslant p} S_{ij} + \sum_{1 \leqslant i < j < k \leqslant p} S_{ijk} + \ldots + S_{12\ldots p} = 1$$

- Number of Sobol' indices: $2^p - 1$
- Usually, we first consider S_i and S_i^T.
- If $\sum_{i=1}^{p} S_i \approx 1$, there is no input interaction.
- If $S_i^T \approx 0$, X_i has no influence on the output variance.
- If $S_i^T \gg S_i$, we consider S_{ij}.

Example #1

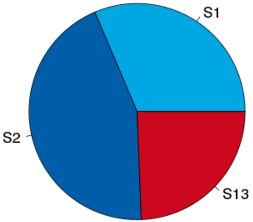
The Ishigami function

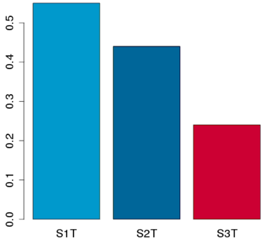
$$\mathcal{M}(\mathbf{X}) = \sin(X_1) + 7\sin^2(X_2) + 0.1\, X_3^4 \sin(X_1)$$

where $X_i \sim \mathcal{U}[-\pi, \pi]$ are independent uniform random variables.

Sobol' indices

- First order: $S_1 = 0.31$ $S_2 = 0.44$ $S_3 = 0.00$ $S_1 + S_2 + S_3 = 0.75$
- Total order: $S_1^T = 0.55$ $S_2^T = 0.44$ $S_3^T = 0.24$
- Higher order: $S_{12} = 0.00$ $S_{13} = 0.24$ $S_{23} = 0.00$ $S_{123} = 0.00$





Example #2

1D IVP:
$$\begin{cases} \frac{\mathrm{d}u}{\mathrm{d}t}(t,\omega) = -\Lambda(\omega)u(t,\omega), \quad t \in (0,1], \\ u(t=0,\cdot) = U_0(\omega). \end{cases}$$

Uncertain input parameters: $\mathbf{X} = (\Lambda, U_0)$

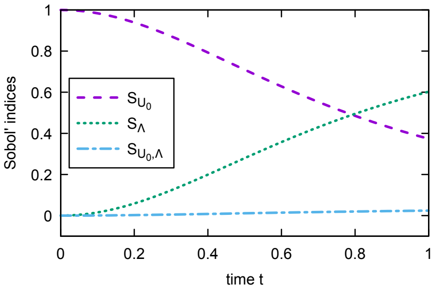
- growth rate $\Lambda \sim \mathcal{N}(\mu = 1,\ \sigma = 0.25)$.
- initial condition $U_0 \sim \mathcal{N}(\mu_0 = 10,\ \sigma_0 = 2)$.

Sobol' indices for $Y(t) = \mathcal{M}(\mathbf{X};t) \equiv u(t,\cdot) = U_0 e^{-\Lambda t}$

$$S_{U_0}(t) = \frac{\sigma_0^2 e^{-\sigma^2 t^2}}{\sigma_0^2 + \mu_0^2(1 - e^{-\sigma^2 t^2})}$$

$$S_{\Lambda}(t) = \frac{\mu_0^2(1 - e^{-\sigma^2 t^2})}{\sigma_0^2 + \mu_0^2(1 - e^{-\sigma^2 t^2})}$$

$$S_{U_0,\Lambda}(t) = \frac{\sigma_0^2(1 - e^{-\sigma^2 t^2})}{\sigma_0^2 + \mu_0^2(1 - e^{-\sigma^2 t^2})}$$



Sobol' indices: mathematical expression

Reminder: $\mathcal{M}(\mathbf{X}) = \mathcal{M}_\varnothing + \sum_{1 \leqslant i \leqslant p} \mathcal{M}_i(X_i) + \sum_{1 \leqslant i < j \leqslant p} \mathcal{M}_{ij}(\mathbf{X}_{ij}) + \cdots + \mathcal{M}_{1\ldots p}(\mathbf{X})$

with $\mathcal{M}_\varnothing = \mathbb{E}[\mathcal{M}(\mathbf{X})]$, $\mathbb{E}[\mathcal{M}_u(\mathbf{X}_u)] = 0$, $\mathbb{E}[\mathcal{M}_u(\mathbf{X}_u)\mathcal{M}_{v \neq u}(\mathbf{X}_{v \neq u})] = 0$.

Recursive construction:

$$\begin{cases} \mathcal{M}_\varnothing = \mathbb{E}[\mathcal{M}(\mathbf{X})], \\ \mathcal{M}_u(\mathbf{X}_u) = \mathbb{E}[\mathcal{M}(\mathbf{X})|\mathbf{X}_u] - \sum_{v \subsetneq u} \mathcal{M}_v(\mathbf{X}_v), \quad \forall u \subseteq \{1, \ldots, p\}, \ u \neq \varnothing. \end{cases}$$

Sobol' indices: $S_\varnothing = 0$ and $\forall \varnothing \subsetneq u \subseteq \{1, \ldots, p\}$,

$$S_u = \frac{\mathbb{V}[\mathcal{M}_u(\mathbf{X}_u)]}{\mathbb{V}[\mathcal{M}(\mathbf{X})]} = \frac{\mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|\mathbf{X}_u]]}{\mathbb{V}[\mathcal{M}(\mathbf{X})]} - \sum_{v \subsetneq u} S_v$$

Sobol' indices: summary

first order Sobol' indices:

$$S_i = \frac{\mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|X_i]]}{\mathbb{V}[\mathcal{M}(\mathbf{X})]}$$

second order Sobol' indices:

$$S_{ij} = \frac{\mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|X_i, X_j]]}{\mathbb{V}[\mathcal{M}(\mathbf{X})]} - S_i - S_j$$

third order Sobol' indices:

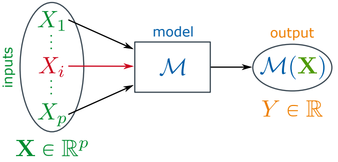
$$S_{ijk} = \frac{\mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|X_i, X_j, X_k]]}{\mathbb{V}[\mathcal{M}(\mathbf{X})]} - S_{ij} - S_{ik} - S_{jk} - S_i - S_j - S_k$$

and so on...

Outline

Pick-and-freeze formulation of Sobol' indices

Simplified UQ framework:



First-order Sobol' indices:

$$S_i = \frac{\mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|X_i]]}{\mathbb{V}[\mathcal{M}(\mathbf{X})]} = \frac{D_i}{D}$$

share of output variance attributable to X_i

Pick-and-freeze **formulation** for D_i:

$$Y = \mathcal{M}(X_1, \dots, X_{i-1}, X_i, X_{i+1}, \dots, X_s),$$

$$Y^{[i]} = \mathcal{M}(X'_1, \dots, X'_{i-1}, X_i, X'_{i+1}, \dots, X'_s).$$

where for $k \neq i$, X'_k is an i.i.d. copy of X_k.

The numerator can be rewritten as a **covariance**: $D_i = \mathbb{C}\big[Y, Y^{[i]}\big]$.

Monte Carlo (MC) estimation

MC expectation estimator of $\mathbb{E}[Y]$ (sample mean): $\boxed{\hat{E}_n[Y] = \frac{1}{n}\sum_{i=1}^{n} Y^{(i)}}$

where $\{Y^{(1)}, \dots, Y^{(n)}\}$ is an n-sample of Y (the $Y^{(i)}$'s are i.i.d. copies of Y).

MC covariance estimator of $\mathbb{C}[Y, Z]$ (unbiased):

$$\boxed{\hat{C}_n[Y, Z] = \frac{n}{n-1}\hat{E}_n\Big[(Y - \hat{E}_n[Y])(Z - \hat{E}_n[Z])\Big]}$$

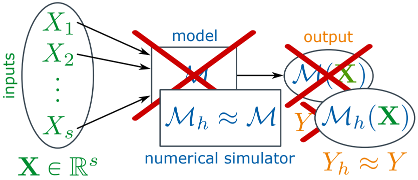
Slow convergence:

$$\mathrm{RMSE}(\hat{E}_n[Y]) = \mathbb{V}[Y]^{1/2} \Big/ \sqrt{n} = \mathcal{O}(1/\sqrt{n})$$

$$\mathrm{RMSE}(\hat{C}_n[Y, Z]) \leqslant \left(\mathbb{M}^4[Y]\,\mathbb{M}^4[Z]\right)^{1/4} \Big/ \sqrt{n-1} = \mathcal{O}(1/\sqrt{n})$$

Consequence:
to reduce the RMSE by a factor of α, one needs to increase n by a factor of α^2.

MC in numerical simulations



$$Y \approx Y_h = \mathcal{M}_h(\mathbf{X})$$

- Numerical scheme $\mathcal{M}_h$ (*e.g.* space and/or time discretization) for solving the mathematical model $\mathcal{M}$ (PDE, ODE, ...).
- h is a discretization parameter (*e.g.* "mesh size"). Ideally, $Y_h \to Y$ as $h \to 0$.
- To obtain an n-sample $\{Y_h^{(1)}, \dots, Y_h^{(n)}\}$ of Y_h:
  1. generate an n-sample $\{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}\}$ of $\mathbf{X}$ (known probability distribution);
  2. from each $\mathbf{X}^{(i)}$ generate $Y_h^{(i)} \equiv \mathcal{M}_h(\mathbf{X}^{(i)})$ $\to$ requires n runs of the simulator.
- Compute the MC estimator. **Ex**: sample mean $\hat{E}_n[Y_h] = \frac{1}{n}\sum_{i=1}^{n} \mathcal{M}_h(\mathbf{X}^{(i)})$.

Consequence for MC estimators

In practice: compute an estimation $\hat{E}_n[Y_h]$ of $\mathbb{E}[Y_h]$.

but we are still interested in the error w.r.t. the "exact" expectation $\mathbb{E}[Y]$.

Variance-bias decomposition of the MSE:

$$\text{MSE} \equiv \mathbb{E}\Big[\Big(\hat{E}_n[Y_h] - \mathbb{E}[Y]\Big)^2\Big] = \underbrace{\mathbb{V}(\hat{E}_n[Y_h])}_{\substack{\text{sampling error} \\ = \mathbb{V}[Y_h]/n}} + \underbrace{|\mathbb{E}[Y_h - Y]|^2}_{\text{(squared) discretization bias}}$$

Also true for other statistics:

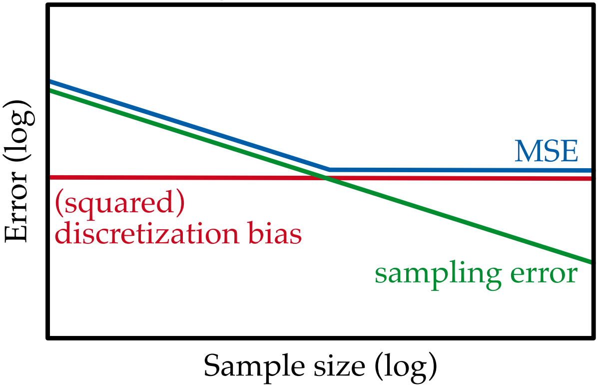
- denote by $\hat{\theta}_{h,n}$ the **unbiased** estimator of $\theta_h = \mathbb{E}[Y_h]$, $\mathbb{V}[Y_h]$, $\mathbb{C}[Y_h, Z_h]$, ...
- assume $\theta_h \to \theta$ as $h \to 0$, where $\theta = \mathbb{E}[Y]$, $\mathbb{V}[Y]$, $\mathbb{C}[Y, Z]$, ...

$$\boxed{\text{MSE} \equiv \mathbb{E}[(\hat{\theta}_{h,n} - \theta)^2] = \underbrace{\mathbb{V}(\hat{\theta}_{h,n})}_{\text{sampling error}} + \underbrace{|\theta_h - \theta|^2}_{\text{(squared) discretization bias}}}$$

Balancing discretization and sampling errors

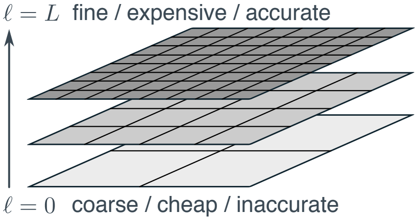
$$\text{MSE} \equiv \mathbb{E}[(\hat{\theta}_{h,n} - \theta)^2] = \underbrace{\mathbb{V}(\hat{\theta}_{h,n})}_{\text{sampling error}} + \underbrace{|\theta_h - \theta|^2}_{\text{(squared) discretization bias}}$$

For a given discretization:



Outline

Multilevel MC (MLMC)



Hierarchy of levels:

- Sequence of meshes with sizes $h_0 > h_1 > \ldots > h_L$.
  Ex: $h_\ell \approx 2^{-\ell}$, *i.e.* $h_\ell = h_{\ell-1}/2$.
- Corresponding numerical simulators $\mathcal{M}_0, \ldots, \mathcal{M}_L$.
- Corresponding discrete outputs $Y_0, \ldots, Y_L$, where $Y_\ell = \mathcal{M}_\ell(\mathbf{X})$.

MLMC estimation — arbitrary parameter

Goal: estimate an arbitrary parameter θ (*e.g.* $\mathbb{E}[Y]$, $\mathbb{V}[Y]$, $\mathbb{C}[Y,Z]$, ...).

Discretization: we can only estimate discrete versions θ_ℓ (*e.g.* $\mathbb{E}[Y_\ell]$, ...).

Multilevel rewriting: for a given finest level L,

$$\theta_L = \theta_0 + \sum_{\ell=1}^{L} \theta_\ell - \theta_{\ell-1}, \qquad \textit{e.g. } \mathbb{E}[Y_L] = \underbrace{\mathbb{E}[Y_0]}_{\text{coarse approx.}} + \sum_{\ell=1}^{L} \underbrace{\mathbb{E}[Y_\ell] - \mathbb{E}[Y_{\ell-1}]}_{\text{correction at level } \ell}$$

MLMC estimation: each term is estimated by a single-level MC estimator

$$\boxed{\hat{\theta}_L^{\mathrm{ML}} = \hat{\theta}_{0,n_0}^{(0)} + \sum_{\ell=1}^{L} \hat{\theta}_{\ell,n_\ell}^{(\ell)} - \hat{\theta}_{\ell-1,n_\ell}^{(\ell)}}$$

- $\hat{\theta}_{\ell,n_\ell}^{(\ell)}$ is an estimator of θ_ℓ using the n_ℓ-sample $\{\mathcal{M}_\ell(\mathbf{X}^{(\ell,i)})\}_{i=1,\dots,n_\ell}$.
- $\hat{\theta}_{\ell-1,n_\ell}^{(\ell)}$ is an estimator of $\theta_{\ell-1}$ using the n_ℓ-sample $\{\mathcal{M}_{\ell-1}(\mathbf{X}^{(\ell,i)})\}_{i=1,\dots,n_\ell}$.

MLMC error

Variance-bias decomposition of the MSE:

$$\mathrm{MSE} \equiv \mathbb{E}\Big[(\hat{\theta}_L^{\mathrm{ML}} - \theta)^2\Big] = \underbrace{\mathbb{V}(\hat{\theta}_L^{\mathrm{ML}})}_{\text{sampling error}} + \underbrace{|\theta_L - \theta|^2}_{\text{(squared) discretization bias}}$$

Important remarks:

- The discretization bias **only depends on the finest level** L.
- The correction mechanism may help reduce the variance.

Example: for the expectation, $\mathbb{V}(\hat{E}_L^{\mathrm{ML}}[Y]) = \dfrac{\mathbb{V}[Y_0]}{n_0} + \displaystyle\sum_{\ell=1}^{L} \dfrac{\mathbb{V}[Y_\ell - Y_{\ell-1}]}{n_\ell}$

- large sample size n_0 on $\ell = 0$;
- smaller sample size n_ℓ on finer correction levels: $\mathbb{V}[Y_\ell - Y_{\ell-1}] \downarrow 0$ as $\ell \to \infty$.
  $\to$ variance reduction technique.

MLMC theorem (assumptions)

Reminder: $\hat{\theta}_L^{\mathrm{ML}} = \hat{\theta}_{0,n_0}^{(0)} + \sum_{\ell=1}^{L} \hat{\theta}_{\ell,n_\ell}^{(\ell)} - \hat{\theta}_{\ell-1,n_\ell}^{(\ell)}$.

Assumptions:

1. There exists a fixed $N \in \mathbb{N}_0$ such that $\boxed{\mathbb{V}[\hat{\theta}_{\ell,n_\ell}^{(\ell)} - \hat{\theta}_{\ell-1,n_\ell}^{(\ell)}] \leqslant \mathcal{V}_\ell/(n_\ell - N).}$

   Ex: for the expectation, $N = 0$ and $\mathcal{V}_\ell = \mathbb{V}[Y_\ell - Y_{\ell-1}]$.

2. There exist constants $\alpha, \beta, \gamma > 0$ such that, for any $\ell \geqslant 0$,

$$\boxed{|\theta_\ell - \theta| \lesssim h_\ell^\alpha, \qquad \mathcal{V}_\ell \lesssim h_\ell^\beta, \qquad \mathcal{C}_\ell \lesssim h_\ell^{-\gamma}.}$$

   Ex: for the expectation, this means

$$|\mathbb{E}[Y_\ell - Y]| \lesssim h_\ell^\alpha, \qquad \mathbb{V}[Y_\ell - Y_{\ell-1}] \lesssim h_\ell^\beta, \qquad \mathcal{C}_\ell \lesssim h_\ell^{-\gamma}.$$

MLMC theorem

Under the previous assumptions:

for any tolerance ε, there exist $L \geqslant 0$ and $(n_\ell)_{\ell=0}^{L}$ such that $\mathrm{RMSE} \leqslant \varepsilon$ and

$$\mathrm{Cost}_\varepsilon(\hat{\theta}_L^{\mathrm{ML}}) \lesssim \varepsilon^{-\frac{\gamma}{\alpha}} + \begin{cases} \varepsilon^{-2} & \text{if } \beta > \gamma \\ \varepsilon^{-2}\log(\varepsilon)^2 & \text{if } \beta = \gamma \\ \varepsilon^{-2-\frac{\gamma-\beta}{\alpha}} & \text{if } \beta < \gamma \end{cases}$$

where $\mathrm{Cost}_\varepsilon(\hat{\theta}_L^{\mathrm{ML}})$ is the total cost of computing the multilevel estimator.

Note: If $\min(\beta, \gamma) \leqslant 2\alpha$, the first term $\varepsilon^{-\frac{\gamma}{\alpha}}$ can be neglected.

Outline

MLMC estimation — covariance

Multilevel covariance estimator:

$$\boxed{\hat{C}_L^{\mathrm{ML}}[Y,Z] = \hat{C}_{n_0}^{(0)}[Y_0,Z_0] + \sum_{\ell=1}^{L} \hat{C}_{n_\ell}^{(\ell)}[Y_\ell,Z_\ell] - \hat{C}_{n_\ell}^{(\ell)}[Y_{\ell-1},Z_{\ell-1}]}$$

where

- $\hat{C}_{n_\ell}^{(\ell)}[Y_\ell,Z_\ell]$ is the single-level **unbiased** MC estimator of $\mathbb{C}[Y_\ell,Z_\ell]$.
- $\hat{C}_{n_\ell}^{(\ell)}[Y_{\ell-1},Z_{\ell-1}]$ is the single-level **unbiased** MC estimator of $\mathbb{C}[Y_{\ell-1},Z_{\ell-1}]$.

Sufficient conditions for the covariance MLMC theorem

Reminder: for the MLMC theorem, we need

1. $|\theta_\ell - \theta| = |\mathbb{C}[Y_\ell, Z_\ell] - \mathbb{C}[Y, Z]| \lesssim h_\ell^\alpha$.
2. $\mathcal{V}_\ell \lesssim h_\ell^\beta$.

We can show that

$$\left.\begin{aligned} h_\ell \eqsim s^{-\ell} \text{ for some fixed } s > 1 \\ |\mathbb{C}[Y_\ell, Z_\ell] - \mathbb{C}[Y_{\ell-1}, Z_{\ell-1}]| \lesssim h_\ell^\alpha \end{aligned}\right\} \forall \ell \geqslant 1 \implies |\mathbb{C}[Y_\ell, Z_\ell] - \mathbb{C}[Y, Z]| \lesssim h_\ell^\alpha$$

and

$$\left.\begin{aligned} \{\mathbb{M}^4[Y_\ell]\}_{\ell \geqslant 0} \text{ is uniformly bounded} \\ \{\mathbb{M}^4[Z_\ell]\}_{\ell \geqslant 0} \text{ is uniformly bounded} \\ \sqrt{\mathbb{M}^4[Y_\ell - Y_{\ell-1}]} + \sqrt{\mathbb{M}^4[Z_\ell - Z_{\ell-1}]} \lesssim h_\ell^\beta \end{aligned}\right\} \implies \mathcal{V}_\ell \lesssim h_\ell^\beta.$$

Theorem for MLMC covariance estimation

Under the assumptions that

1. $h_\ell \eqsim s^{-\ell}$ for some fixed $s > 1$,
2. $\{\mathbb{M}^4[Y_\ell]\}_{\ell \geqslant 0}$ and $\{\mathbb{M}^4[Z_\ell]\}_{\ell \geqslant 0}$ are uniformly bounded,
3. there exist constants $\alpha, \beta, \gamma > 0$ such that, $\forall \ell \geqslant 0$,
   - $|\mathbb{C}[Y_\ell, Z_\ell] - \mathbb{C}[Y_{\ell-1}, Z_{\ell-1}]| \lesssim h_\ell^\alpha$,
   - $\sqrt{\mathbb{M}^4[Y_\ell - Y_{\ell-1}]} + \sqrt{\mathbb{M}^4[Z_\ell - Z_{\ell-1}]} \lesssim h_\ell^\beta$,
   - $\mathcal{C}_\ell \lesssim h_\ell^{-\gamma}$,

we have:
for any tolerance ε, there exist $L \geqslant 0$ and $(n_\ell)_{\ell=0}^L$ such that $\mathrm{RMSE} \leqslant \varepsilon$ and

$$\mathrm{Cost}_\varepsilon(\hat{C}_L^{\mathrm{ML}}[Y, Z]) \lesssim \varepsilon^{-\frac{\gamma}{\alpha}} + \begin{cases} \varepsilon^{-2} & \text{if } \beta > \gamma \\ \varepsilon^{-2} \log(\varepsilon)^2 & \text{if } \beta = \gamma \\ \varepsilon^{-2-\frac{\gamma-\beta}{\alpha}} & \text{if } \beta < \gamma \end{cases}$$

Note: If $\min(\beta, \gamma) \leqslant 2\alpha$, the first term $\varepsilon^{-\frac{\gamma}{\alpha}}$ can be neglected.

Outline

Numerical illustration: 1D initial value problem

1D IVP: $\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t,\omega) = -\Lambda(\omega)u(t,\omega), \quad t \in (0,1], \\ u(t=0,\cdot) = U_0(\omega). \end{cases}$

- **Uncertain input parameters**: $\mathbf{X} = (\Lambda, U_0)$
  - growth rate $\Lambda \sim \mathcal{N}(\mu = 1,\ \sigma = 0.25)$.
  - initial condition $U_0 \sim \mathcal{N}(\mu_0 = 10,\ \sigma_0 = 2)$.
- **Output of interest**: $Y = \mathcal{M}(\Lambda, U_0) \equiv u(t=1,\cdot) = U_0 e^{-\Lambda}$.
- **Goal**: estimate numerators of Sobol' indices [Reminder: $S_i = D_i/D$]

$$\begin{aligned} D_\Lambda &\equiv \mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|\Lambda]] \\ &= \mathbb{C}[\mathcal{M}(\Lambda, U_0),\ \mathcal{M}(\Lambda, U_0')] \\ &= 0.929. \end{aligned}$$

and

$$\begin{aligned} D_{U_0} &\equiv \mathbb{V}[\mathbb{E}[\mathcal{M}(\mathbf{X})|U_0]] \\ &= \mathbb{C}[\mathcal{M}(\Lambda, U_0),\ \mathcal{M}(\Lambda', U_0)] \\ &= 0.577. \end{aligned}$$

Numerical scheme

1D IVP:
$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t,\omega) = -\Lambda(\omega)u(t,\omega), \quad t \in (0,1], \\ u(t=0,\cdot) = U_0(\omega). \end{cases}$$

Sequence of levels: "mesh sizes" (time-step sizes) $\{h_\ell \equiv \delta t_\ell = 1/m_\ell\}_{\ell \geqslant 0}$, where $m_\ell = 16 \times 2^\ell$ is the number of time-steps used to discretize $(0,1]$.

Note: $h_\ell \eqsim 2^{-\ell}$.

Backward Euler scheme using m_ℓ time-steps:

$$u_\ell^0(\Lambda, U_0) = U_0 \quad \text{and} \quad \forall k = 1, \dots, m_\ell, \quad u_\ell^k(\Lambda, U_0) = \frac{u_\ell^{k-1}(\Lambda, U_0)}{1 - \Lambda/m_\ell}.$$

Note: $\mathcal{C}_\ell \lesssim h_\ell^{-\gamma}$ with $\gamma = 1$.

Discrete output: $Y \approx Y_\ell = \mathcal{M}_\ell(\Lambda, U_0) \equiv u_\ell^{m_\ell}(\Lambda, U_0)$.

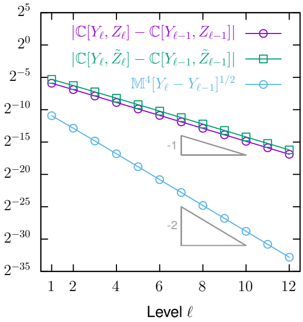
Convergence rates

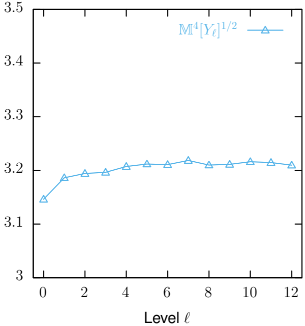
Notations:

$Y_\ell \equiv \mathcal{M}_\ell(\Lambda, U_0)$

$Z_\ell \equiv \mathcal{M}_\ell(\Lambda', U_0)$

$\tilde{Z}_\ell \equiv \mathcal{M}_\ell(\Lambda, U_0')$





$$|\mathbb{C}[Y_\ell, Z_\ell] - \mathbb{C}[Y_{\ell-1}, Z_{\ell-1}]| \lesssim 2^{-\alpha\ell}, \quad \alpha \approx 1$$

$$|\mathbb{C}[Y_\ell, \tilde{Z}_\ell] - \mathbb{C}[Y_{\ell-1}, \tilde{Z}_{\ell-1}]| \lesssim 2^{-\tilde{\alpha}\ell}, \quad \tilde{\alpha} \approx 1$$

$$\sqrt{\mathbb{M}^4[Y_\ell - Y_{\ell-1}]} + \sqrt{\mathbb{M}^4[Z_\ell - Z_{\ell-1}]} = 2 \times \sqrt{\mathbb{M}^4[Y_\ell - Y_{\ell-1}]} \lesssim 2^{-\beta\ell}, \quad \beta \approx 2$$

$$\sqrt{\mathbb{M}^4[Y_\ell - Y_{\ell-1}]} + \sqrt{\mathbb{M}^4[\tilde{Z}_\ell - \tilde{Z}_{\ell-1}]} = 2 \times \sqrt{\mathbb{M}^4[Y_\ell - Y_{\ell-1}]} \lesssim 2^{-\beta\ell}$$

$\{\mathbb{M}^4[Y_\ell]\}_{\ell\geqslant 0}$ is uniformly bounded (and so are $\{\mathbb{M}^4[Z_\ell]\}_{\ell\geqslant 0}$ and $\{\mathbb{M}^4[\tilde{Z}_\ell]\}_{\ell\geqslant 0}$).

Consequences for MLMC

We have: $\alpha \approx 1$, $\beta \approx 2$ and $\gamma \approx 1$, so $\beta > \gamma$ and $\min(\beta, \gamma) = \gamma \leqslant 2\alpha$.

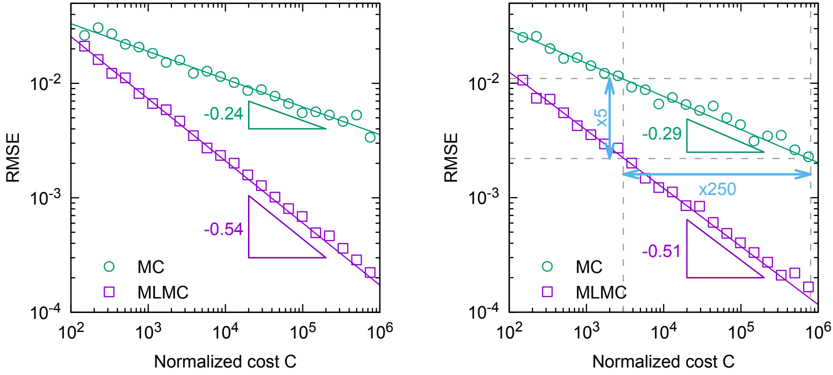
According to the MLMC theorem:

For any tolerance ε, there exist $L \geqslant 0$ and $(n_\ell)_{\ell=0}^{L}$ such that

$$\mathrm{RMSE} \leqslant \varepsilon, \quad \text{and } \mathrm{Cost}_\varepsilon(\hat{C}_L^{\mathrm{ML}}[Y, Z]) \lesssim \varepsilon^{-2}.$$

On the other hand, the cost of standard MC is $\mathcal{O}(\varepsilon^{-3})$.

Illustration on the toy SDE example



(a) Estimation of D_Λ (b) Estimation of D_{U_0}

$$\mathrm{Cost}(\mathrm{MC}) \lesssim \varepsilon^{-3} \iff \mathrm{RMSE}(\mathrm{MC}) \lesssim \mathrm{Cost}(\mathrm{MC})^{-1/3}$$

$$\mathrm{Cost}(\mathrm{ML}) \lesssim \varepsilon^{-2} \iff \mathrm{RMSE}(\mathrm{ML}) \lesssim \mathrm{Cost}(\mathrm{ML})^{-1/2}$$

Outline

Conclusion

Summary:

- Sobol' indices are a common tool for sensitivity analysis.
- Their numerator can be expressed as a covariance.
- MLMC methods reduce the overall sampling cost.
- The methodology can be applied to higher order Sobol' indices.
- Not always easy to estimate α and β from few samples.
- Other sources of error? (geometry, ...).

Extensions:

- MLMC theory can be extended to the estimation of higher-order moments.
- MLMC can be applied to problems in higher dimension (multiindex MC).
- Adaptive algorithms exist to pick the number of levels and sample sizes.
- Advanced sampling methods can be combined with MLMC (e.g. QMC).