An Introduction to Multilevel Monte Carlo Methods for Uncertainty Quantification in Earth Science

Tobias Schwedes

PhD candidate, Imperial College London

Supervisors: Ben Calderhead, David Ham, Simon Funke, Matthew Piggott

1st October 2016

Introduction

- Investigating complex physical phenomena requires modelling
- Models are subject to unavoidable uncertainties (in measurements, due to neglect of physical effects, ...)
- Reliable statements can still be achieved by:
  - Quantitative representation of uncertainties
  - Techniques to estimate quantities of interest in presence of these uncertainties
- Often we need to evaluate integrals for which Monte Carlo methods prove as suitable estimation techniques!

Imperial College London

Formalism for physical model

$$\mathcal{F}(\alpha) = X$$

- $\mathcal{F}$: forward model (ODE, PDE, SDE, SPDE, ...)
- α : random input parameters (initial conditions, model parameters, ...)
- X : solution to $\mathcal{F}$ and α - random as such!

Note that α and X are generally infinite-dimensional objects, thus their exact generation is in practice rarely possible!

Discretised representation and problem formulation

- Assume α can be approximated by a random vector $\vec{\alpha}_n \in \mathbb{R}^n$
- Assume there is a discretisation $\mathcal{F}_h$ of $\mathcal{F}$ with associated solution random vector $\vec{X}_m \in \mathbb{R}^m$

$\Longrightarrow$ Discrete formalism:

$$\mathcal{F}_h(\vec{\alpha}_n) = \vec{X}_m$$

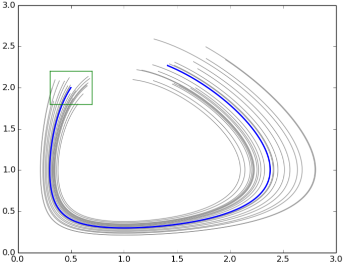
- Let $Q_m := \mathcal{G}(\vec{X}_m)$ be some quantity of interest, where $\mathcal{G}$ is a deterministic and continuous functional
- For some generally inaccessible random variable Q with $Q_m \to Q$, we want to **estimate** $\mathbb{E}[Q]$

Imperial College London

Predator-prey model

System of first-order, non-linear, ordinary differential equations (ODEs), modelling the dynamics of two interacting species (predator and prey) in time:

$$\dot{\vec{u}} = \begin{pmatrix} \dot{u_1} \\ \dot{u_2} \end{pmatrix} = \begin{pmatrix} u_1(1 - u_2) \\ u_2(u_1 - 1) \end{pmatrix} .$$



- Initial state $\vec{u}_0$ uncertain, modelled as $\vec{u}_0 \sim \mathcal{U}(\Gamma)$
- $\Gamma := \bar{\vec{u}}_0 + [-\varepsilon, \varepsilon]^2$ for some deterministic $\bar{\vec{u}}_0 \in (0, \infty)^2$
- **Aim:** estimate $\mathbb{E}[u_1(T)]$ for some $T > 0$

Imperial College London

Predator-prey model (continued)

With the notation from above, it holds

- Input parameter: $\alpha = \vec{u}_0$

Predator-prey model (continued)

With the notation from above, it holds

- Input parameter: $\alpha = \vec{u}_0$
- Solution: $X = \vec{u}(0 < t < T) = \mathcal{F}(\vec{u}_0)$

Imperial College London

Predator-prey model (continued)

With the notation from above, it holds

- Input parameter: $\alpha = \vec{u}_0$
- Solution: $X = \vec{u}(0 < t < T) = \mathcal{F}(\vec{u}_0)$
- Inaccessible quantity of interest: $Q = u_1(T)$

Imperial College London

Predator-prey model (continued)

With the notation from above, it holds

- Input parameter: $\alpha = \vec{u}_0$
- Solution: $X = \vec{u}(0 < t < T) = \mathcal{F}(\vec{u}_0)$
- Inaccessible quantity of interest: $Q = u_1(T)$
- Discretised forward model $\mathcal{F}_h$ is the solution operator of an explicit forward Euler scheme with step size $h = T/m$ for some $m \in \mathbb{N}$

Imperial College London

Predator-prey model (continued)

With the notation from above, it holds

- Input parameter: $\alpha = \vec{u}_0$
- Solution: $X = \vec{u}(0 < t < T) = \mathcal{F}(\vec{u}_0)$
- Inaccessible quantity of interest: $Q = u_1(T)$
- Discretised forward model $\mathcal{F}_h$ is the solution operator of an explicit forward Euler scheme with step size $h = T/m$ for some $m \in \mathbb{N}$
- $\vec{X}_m = \mathcal{F}_h(\vec{u}_0)$ is m-dimensional with ith component being the solution of the Euler scheme for u_1 at time $i \cdot T/m$

Imperial College London

Predator-prey model (continued)

With the notation from above, it holds

- Input parameter: $\alpha = \vec{u}_0$
- Solution: $X = \vec{u}(0 < t < T) = \mathcal{F}(\vec{u}_0)$
- Inaccessible quantity of interest: $Q = u_1(T)$
- Discretised forward model $\mathcal{F}_h$ is the solution operator of an explicit forward Euler scheme with step size $h = T/m$ for some $m \in \mathbb{N}$
- $\vec{X}_m = \mathcal{F}_h(\vec{u}_0)$ is m-dimensional with ith component being the solution of the Euler scheme for u_1 at time $i \cdot T/m$
- Q_m is the last value of $\vec{X}_m$

Imperial College London

Standard Monte Carlo

$$\mathcal{MC}_N(Q) := \frac{1}{N}\sum_{i=1}^{N} Q_m^{(i)} \overset{N \text{ large}}{\approx} \mathbb{E}[Q_m] \overset{m \text{ large}}{\approx} \mathbb{E}[Q]$$

- $Q_m^{(i)}, i = 1, ..., N$ are iid samples of Q_m

Imperial College London

Standard Monte Carlo

$$\mathcal{MC}_N(Q) := \frac{1}{N}\sum_{i=1}^{N} Q_m^{(i)} \overset{N \text{ large}}{\approx} \mathbb{E}[Q_m] \overset{m \text{ large}}{\approx} \mathbb{E}[Q]$$

- $Q_m^{(i)}, i = 1, \ldots, N$ are iid samples of Q_m
- $\mathcal{MC}_N(Q)$ is a consistent, unbiased, asymptotically normal estimator for $\mathbb{E}[Q_m]$ with mean squared error (MSE)

$$\text{MSE}\left[\mathcal{MC}_N(Q)\right] = \underbrace{\frac{1}{N}\text{Var}[Q_m]}_{\text{sampling error}} + \underbrace{\Big(\mathbb{E}\left[Q_m\right] - \mathbb{E}[Q]\Big)^2}_{\text{numerical error}},$$

Imperial College London

Standard Monte Carlo

$$\mathcal{MC}_N(Q) := \frac{1}{N}\sum_{i=1}^{N} Q_m^{(i)} \overset{N \text{ large}}{\approx} \mathbb{E}[Q_m] \overset{m \text{ large}}{\approx} \mathbb{E}[Q]$$

- $Q_m^{(i)}, i = 1, ..., N$ are iid samples of Q_m
- $\mathcal{MC}_N(Q)$ is a consistent, unbiased, asymptotically normal estimator for $\mathbb{E}[Q_m]$ with mean squared error (MSE)

$$\text{MSE}\left[\mathcal{MC}_N(Q)\right] = \underbrace{\frac{1}{N}\text{Var}[Q_m]}_{\text{sampling error}} + \underbrace{\Big(\mathbb{E}\left[Q_m\right] - \mathbb{E}[Q]\Big)^2}_{\text{numerical error}},$$

- Rate of convergence is $\mathcal{O}(N^{-1/2})$, **independently of dimensionality**!

Imperial College London

Standard Monte Carlo

$$\mathcal{MC}_N(Q) := \frac{1}{N}\sum_{i=1}^{N} Q_m^{(i)} \overset{N \text{ large}}{\approx} \mathbb{E}[Q_m] \overset{m \text{ large}}{\approx} \mathbb{E}[Q]$$

- $Q_m^{(i)}, i = 1, ..., N$ are iid samples of Q_m
- $\mathcal{MC}_N(Q)$ is a consistent, unbiased, asymptotically normal estimator for $\mathbb{E}[Q_m]$ with mean squared error (MSE)

$$\text{MSE}\left[\mathcal{MC}_N(Q)\right] = \underbrace{\frac{1}{N}\text{Var}[Q_m]}_{\text{sampling error}} + \underbrace{\Big(\mathbb{E}\left[Q_m\right] - \mathbb{E}[Q]\Big)^2}_{\text{numerical error}},$$

- Rate of convergence is $\mathcal{O}(N^{-1/2})$, **independently of dimensionality**!

Problem: Requiring $N = \mathcal{O}(\varepsilon^{-2})$ samples to achieve ε accuracy while each sample requires solving $\mathcal{F}_h(\vec{\alpha}_n) = \vec{X}_m$

Imperial College London

Multi-level Monte Carlo (MLMC)

- **Idea:** use samples from different resolutions instead of just a single one to reduce computational cost (Giles, 2008 & Heinrich, 2001)

Imperial College London

Multi-level Monte Carlo (MLMC)

- **Idea:** use samples from different resolutions instead of just a single one to reduce computational cost (Giles, 2008 & Heinrich, 2001)
- $Q_{m_\ell} = \mathcal{G}(\vec{X}_{m_\ell})$ with m_ℓ discretisation points ($\ell = 0, ..., L$), where $m_0 < ... < m_L =: m$

Imperial College London

Multi-level Monte Carlo (MLMC)

- **Idea:** use samples from different resolutions instead of just a single one to reduce computational cost (Giles, 2008 & Heinrich, 2001)
- $Q_{m_\ell} = \mathcal{G}(\vec{X}_{m_\ell})$ with m_ℓ discretisation points ($\ell = 0, ..., L$), where $m_0 < ... < m_L =: m$
- Due to linearity of expectation, $\Delta_\ell := Q_{m_\ell} - Q_{m_{\ell-1}}$ and $\Delta_0 := Q_{m_0}$,

$$\mathbb{E}[Q_m] = \mathbb{E}[Q_{m_0}] + \sum_{\ell=1}^{L} \mathbb{E}\left[Q_{m_\ell} - Q_{m_{\ell-1}}\right] = \sum_{\ell=1}^{L} \mathbb{E}\left[\Delta_\ell\right]$$

Imperial College London

Multi-level Monte Carlo (MLMC)

- **Idea:** use samples from different resolutions instead of just a single one to reduce computational cost (Giles, 2008 & Heinrich, 2001)
- $Q_{m_\ell} = \mathcal{G}(\vec{X}_{m_\ell})$ with m_ℓ discretisation points ($\ell = 0, ..., L$), where $m_0 < ... < m_L =: m$
- Due to linearity of expectation, $\Delta_\ell := Q_{m_\ell} - Q_{m_{\ell-1}}$ and $\Delta_0 := Q_{m_0}$,

$$\mathbb{E}[Q_m] = \mathbb{E}[Q_{m_0}] + \sum_{\ell=1}^{L} \mathbb{E}\left[Q_{m_\ell} - Q_{m_{\ell-1}}\right] = \sum_{\ell=1}^{L} \mathbb{E}\left[\Delta_\ell\right]$$

- Estimate each correction term $\mathbb{E}[\Delta_\ell]$ independently from all other levels via standard MC, leading to

$$\mathcal{MLMC}_{\vec{N}}(Q) := \sum_{\ell=1}^{L} \mathcal{MC}_{N_\ell}(\Delta_\ell)$$

Imperial College London

MLMC (continued)

- Samples $Q_{m_\ell}^{(i)}$ and $Q_{m_{\ell-1}}^{(i)}$ in $\Delta_\ell^{(i)} := Q_{m_\ell}^{(i)} - Q_{m_{\ell-1}}^{(i)}$ are generated by the same underlying random parameters
- For the MSE, it holds

$$\text{MSE}\left[\mathcal{MLMC}_{\vec{N}}(Q)\right] = \sum_{\ell=1}^{L} \frac{1}{N_\ell} \text{Var}\left[\Delta_\ell\right] + \Big(\mathbb{E}\left[Q_m\right] - \mathbb{E}[Q]\Big)^2.$$

- Assuming $Q_m \to Q$ in L^2 implies $\text{Var}[\Delta_\ell] \to 0$, indicating that N_ℓ decreases in ℓ (fewer samples on finer resolutions)
- However sampling on coarse levels is cheap such that overall cost of achieving ε accuracy is reduced compared to standard MC

Imperial College London

Predator-prey model (MLMC)

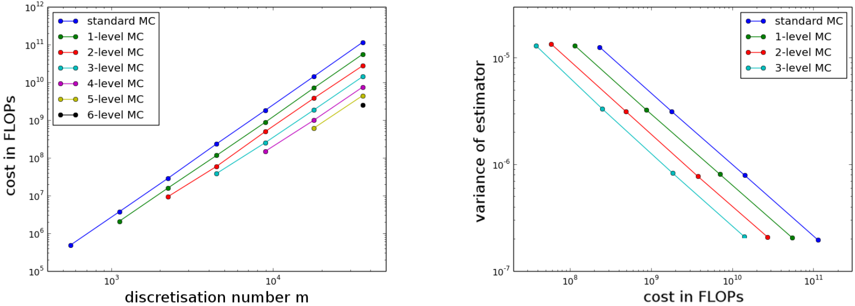


Figure : *Left* plots of cost in FLOPs versus number of discretisation points m for a decreasing tolerance of the MSE. *Right* corresponding estimator variance versus cost in FLOPs.

Imperial College London

Further work

- **Unbiased MLMC** (Rhee & Glynn, 2012 & McLeish, 2010): Randomise number of levels L and thereby remove numerical error!
- **Quasi-Monte Carlo (QMC)**: Use deterministic low-discrepancy samples instead of pseudo-random numbers
  - Convergence rates close to $\mathcal{O}(N^{-1})$
  - Biased, lack of practical error estimates
- **Randomised QMC (rQMC)**: Use QMC, but randomise samples and thereby overcome QMC drawbacks
- **Multilevel rQMC (MLQMC)** (Giles & Waterhouse, 2009): Instead of estimating $\mathbb{E}[\Delta_\ell]$ by standard MC, use rQMC
  - Further increase of convergence rate

Imperial College London

Predator-prey model (method comparison)

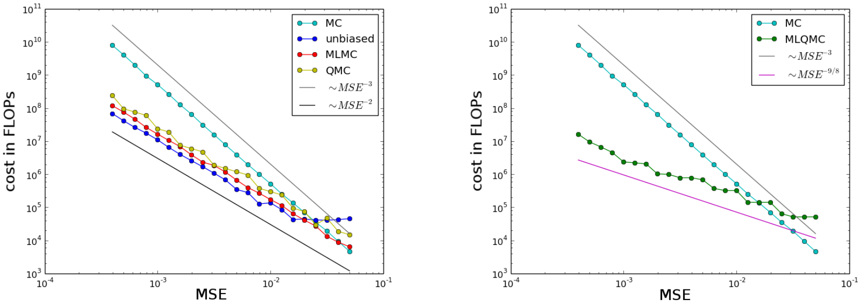


Figure : *Left* plots of cost in FLOPs versus MSE for MLMC, QMC, unbiased multilevel method and *right* MLQMC comparing to standard Monte Carlo.

Predator-prey model (method comparison)

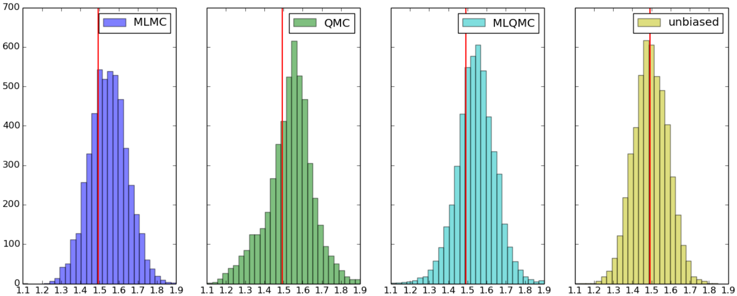
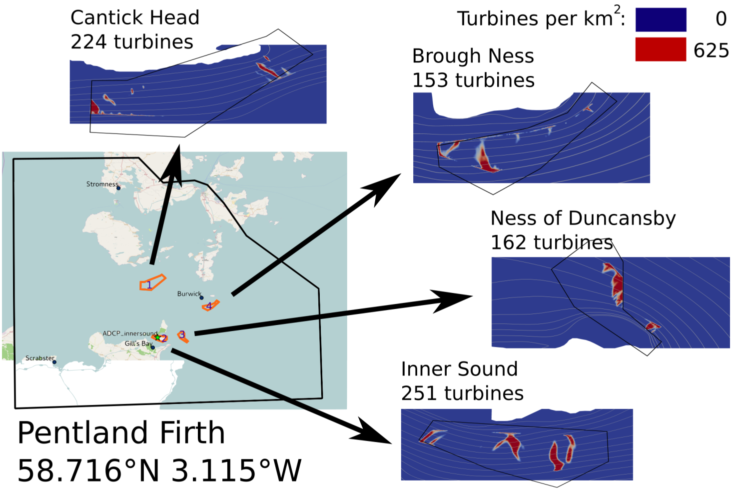


Figure : Histograms for MLMC, QMC, MLQMC and unbiased multilevel estimation, each based on 5000 estimates for $\mathbb{E}[u_1(T)]$ with a fixed RMSE of $\varepsilon = 0.1$. The gold standard value is given by the red line.

Application to earth science problems

- Same methods in principle applicable in the context of any differential equation setup (general formalism)
- Reduction in complexity order / computational cost of standard MC
- Large-scale UQ problems in earth science become tractable that otherwise would not be due to tremendous computational cost
- **Example**: Reliable prediction for energy production of tidal turbines in the presence of uncertainties in the flow model
  - Domain of size of hundreds of kilometres diameter

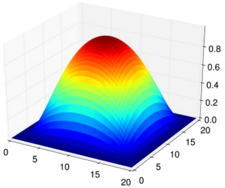
Application to tidal turbine array assessment

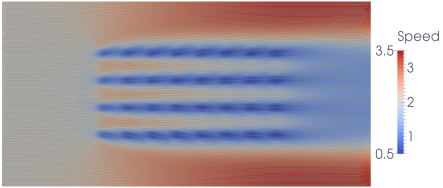


Imperial College London

Application to tidal turbine array assessment (continued)

- Forward model $\mathcal{F}$ given by shallow water equations (SWE)
- Discretised forward model $\mathcal{F}_h$ finite element SWE solver
- Random input α given by random initial & boundary conditions, bottom friction field, water depth at rest, source terms ...
- Turbines represented by locally increased bottom friction field (bump function)

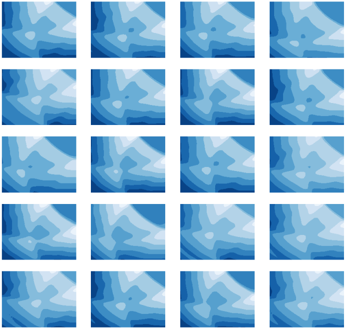




(Opentidalfarm.org)

Imperial College London

Proof of concept: MLMC for SWE



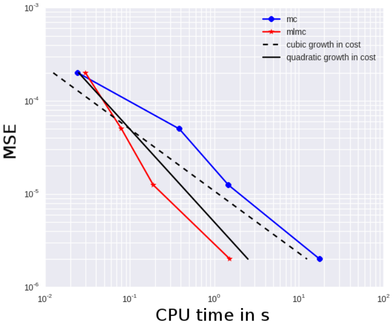


Figure : *Left* plots of an ensemble of realisations of free surface solutions for SWE on high resolution. *Right* MSE versus cost in CPU time in s for MLMC

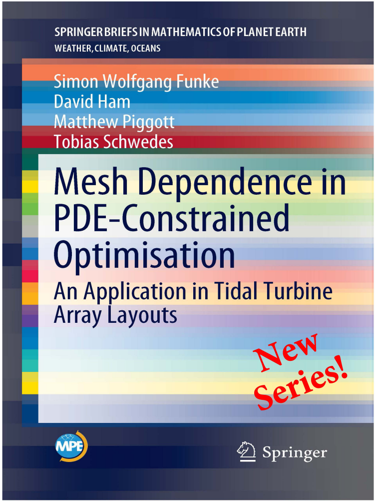
(plots by Alistair Gregory, PyClaw 2D shallow water)

Imperial College London

Future work

- Represent all relevant sources of uncertainty by random fields
- Apply MLMC for general energy estimates
- Incorporate unbiased and QMC approaches
- Use adaptively refined meshes
- Use multi-fidelity approaches

Imperial College London



SpringerBriefs in Mathematics of Planet Earth: Weather, Climate, Oceans

Imperial College London

References

- Giles, Michael B. "Multilevel Monte Carlo path simulation." Operations Research 56.3 (2008): 607-617.
- Heinrich, Stefan. "Multilevel Monte Carlo methods." International Conference on Large-Scale Scientific Computing. Springer Berlin Heidelberg, 2001.
- Cliffe, K. Andrew, et al. "Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients." Computing and Visualization in Science 14.1 (2011): 3-15.
- Rhee, Chang-Han, and Peter W. Glynn. "A new approach to unbiased estimation for SDE's." Proceedings of the Winter Simulation Conference. Winter Simulation Conference, 2012.
- McLeish, Don. "A general method for debiasing a Monte Carlo estimator." Monte Carlo Meth. and Appl. 17.4 (2011): 301-315.
- Dick, Josef, Frances Y. Kuo, and Ian H. Sloan. "High-dimensional integration: the quasi-Monte Carlo way." Acta Numerica 22 (2013): 133-288.
- Giles, Michael B., and Ben J. Waterhouse. "Multilevel quasi-Monte Carlo path simulation." Advanced Financial Modelling, Radon Series on Computational and Applied Mathematics (2009): 165-181.