Multilevel Monte Carlo methods for uncertainty quantification in subsurface flow

Aretha Teckentrup

Department of Mathematical Sciences
University of Bath

Joint work with:

Rob Scheichl and Elisabeth Ullmann (Bath),
Julia Charrier (Marseille), Mike Giles (Oxford),
Andrew Cliffe, Minho Park (Nottingham),
Christian Ketelsen and Panayot Vassilevski (LLNL)

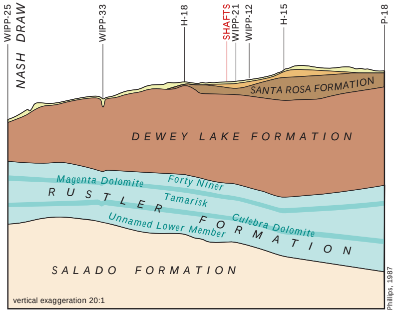
May 23, 2013

Outline

- Motivation and model problem: uncertainty quantification in radioactive waste disposal
- Standard Monte Carlo
- **Multilevel Monte Carlo**
- **Multilevel Markov chain Monte Carlo**

Application: WIPP test case

US Dept Energy Radioactive Waste Isolation Pilot Plant (WIPP) in New Mexico



Cross-section at WIPP

Cross section through the rock at the WIPP site

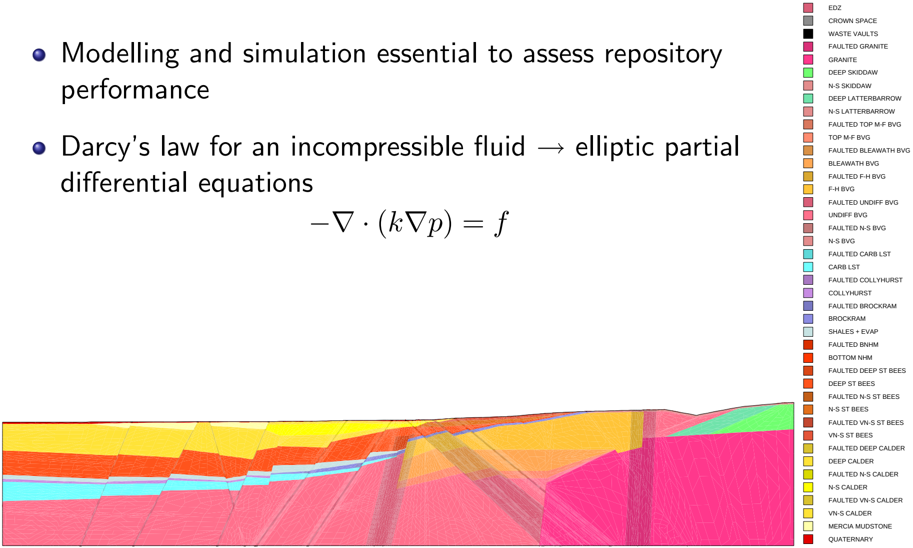
- Crucial to assess the risk of radionuclides reentering the human environment
- Culebra Dolomite layer acts as principal pathway for transport of radionuclides

  (2D to reasonable approximation)

Uncertainty in Groundwater Flow

- Modelling and simulation essential to assess repository performance
- Darcy's law for an incompressible fluid → elliptic partial differential equations

$$-\nabla \cdot (k \nabla p) = f$$

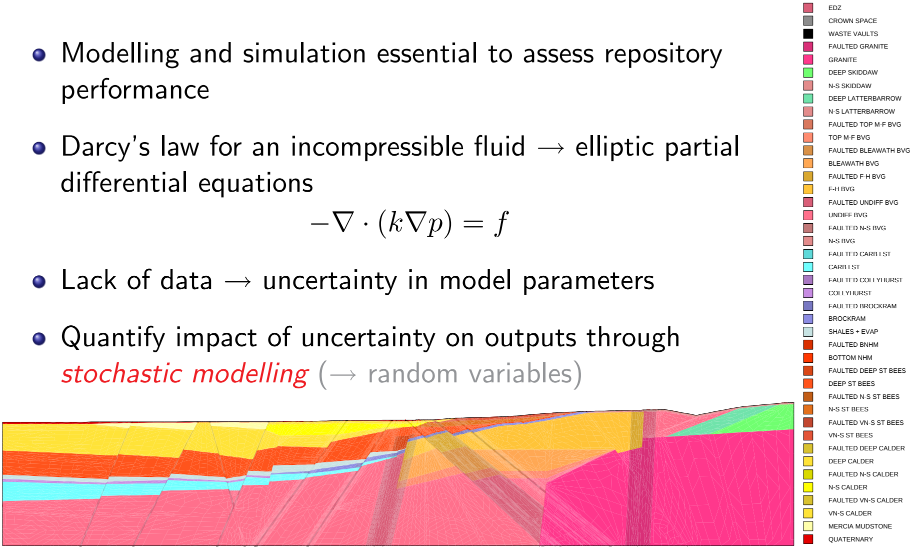


Uncertainty in Groundwater Flow

- Modelling and simulation essential to assess repository performance
- Darcy's law for an incompressible fluid → elliptic partial differential equations

$$-\nabla \cdot (k \nabla p) = f$$

- Lack of data → uncertainty in model parameters
- Quantify impact of uncertainty on outputs through *stochastic modelling* (→ random variables)



Typical model

- Typical simplified model for k is a log–normal random field, $k = \exp[g]$, where g is a scalar, isotropic Gaussian field.
- To sample from k, use Karhunen-Loève expansion:

$$\log k(x,\omega) \approx \sum_{j=1}^{J} \sqrt{\mu_j}\phi_j(x) Z_j(\omega),$$

with $Z_j(\omega)$ i.i.d. $N(0,1)$.

Stochastic modelling

- Many reasons for stochastic modelling in earth sciences:
  - **lack of data** (e.g. data assimilation for weather prediction)
  - **unresolvable scales** (e.g. atmospheric dispersion modelling)
- **Input:** best knowledge about system, statistics of input parameters, measured data with error statistics, etc...
- **Output:** statistics of quantities of interest or of entire state space

Stochastic modelling

- Many reasons for stochastic modelling in earth sciences:
  - **lack of data** (e.g. data assimilation for weather prediction)
  - **unresolvable scales** (e.g. atmospheric dispersion modelling)
- **Input:** best knowledge about system, statistics of input parameters, measured data with error statistics, etc...
- **Output:** statistics of quantities of interest or of entire state space

$$\mathbf{Z}_J(\omega) \in \mathbb{R}^J \xrightarrow{\text{Model}(M)} \mathbf{X}_M(\omega) \in \mathbb{R}^M \xrightarrow{\text{Output}} Q_{M,J}(\omega) \in \mathbb{R}$$

random input | state vector | quantity of interest

- e.g. $\mathbf{Z}_J$ multivariate Gaussian; $\mathbf{X}_M$ numerical solution of PDE; $Q_{M,J}$ a (non)linear functional of $\mathbf{X}_M$

Stochastic modelling

- Many reasons for stochastic modelling in earth sciences:
  - **lack of data** (e.g. data assimilation for weather prediction)
  - **unresolvable scales** (e.g. atmospheric dispersion modelling)
- **Input:** best knowledge about system, statistics of input parameters, measured data with error statistics, etc...
- **Output:** statistics of quantities of interest or of entire state space

$$\mathbf{Z}_J(\omega) \in \mathbb{R}^J \xrightarrow{\text{Model}(M)} \mathbf{X}_M(\omega) \in \mathbb{R}^M \xrightarrow{\text{Output}} Q_{M,J}(\omega) \in \mathbb{R}$$

random input — state vector — quantity of interest

- e.g. $\mathbf{Z}_J$ multivariate Gaussian; $\mathbf{X}_M$ numerical solution of PDE; $Q_{M,J}$ a (non)linear functional of $\mathbf{X}_M$
- $Q(\omega)$ inaccessible random variable s.t. $\mathbb{E}[Q_{M,J}] \xrightarrow{M,J\to\infty} \mathbb{E}[Q]$

Standard Monte Carlo

Usually interested in finding the expected value (or higher order moments) of output functional Q.

Standard Monte Carlo

Usually interested in finding the expected value (or higher order moments) of output functional Q.

The standard **Monte Carlo estimator** for this is

$$\mathbb{E}[Q] \approx \mathbb{E}[Q_{M,J}] \approx \frac{1}{N}\sum_{i=1}^{N} Q_{M,J}^{(i)} := \hat{Q}^{\mathrm{MC}},$$

where $Q_{M,J}^{(i)}$ is the ith i.i.d sample computed with Model(M).

Standard Monte Carlo

Usually interested in finding the expected value (or higher order moments) of output functional Q.

The standard **Monte Carlo estimator** for this is

$$\mathbb{E}\left[Q\right] \;\approx\; \mathbb{E}\left[Q_{M,J}\right] \;\approx\; \frac{1}{N}\sum_{i=1}^{N} Q_{M,J}^{(i)} := \hat{Q}^{\mathrm{MC}},$$

where $Q_{M,J}^{(i)}$ is the ith i.i.d sample computed with Model(M).

The **mean square error** can be shown to equal

$$\begin{aligned}\mathbb{E}\Big[\big(\hat{Q}_h^{\mathrm{MC}} - \mathbb{E}[Q]\big)^2\Big] \;&=\; \mathbb{V}[\hat{Q}_h^{\mathrm{MC}}] \;+\; \big(\mathbb{E}[\hat{Q}_h^{\mathrm{MC}}] - \mathbb{E}[Q]\big)^2 \\ &=\; \underbrace{\frac{\mathbb{V}[Q_{M,J}]}{N}}_{\text{sampling error}} \;+\; \underbrace{\Big(\mathbb{E}[Q_{M,J} - Q]\Big)^2}_{\text{model error ("bias")}}\end{aligned}$$

$\Rightarrow$ **very large N and M!**

Complexity of Standard Monte Carlo

Assuming that

(A1) $\left|\mathbb{E}[Q_{M,J} - Q]\right| = \mathcal{O}(M^{-\alpha})$ (model error)

(A2) $\text{Cost}(Q_{M,J}^{(i)}) = \mathcal{O}(M^{\gamma})$ (PDE solver)

there exist M and N such that the **total cost** to obtain a **mean square error**

$$\mathbb{E}\left[(\hat{Q}^{\mathrm{MC}} - \mathbb{E}[Q])^2\right] \;=\; \mathcal{O}(\varepsilon^2)$$

is

$$\text{Cost}(\hat{Q}^{\mathrm{MC}}) \;=\; \mathcal{O}(\varepsilon^{-2-\gamma/\alpha})$$

Complexity of Standard Monte Carlo

Assuming that

(A1) $\left|\mathbb{E}[Q_{M,J} - Q]\right| = \mathcal{O}(M^{-\alpha})$ (model error)

(A2) Cost$(Q_{M,J}^{(i)}) = \mathcal{O}(M^{\gamma})$ (PDE solver)

there exist M and N such that the **total cost** to obtain a **mean square error**

$$\mathbb{E}\left[(\hat{Q}^{\mathrm{MC}} - \mathbb{E}[Q])^2\right] \; = \; \mathcal{O}(\varepsilon^2)$$

is

$$\text{Cost}(\hat{Q}^{\mathrm{MC}}) \; = \; \mathcal{O}(\varepsilon^{-2\,-\gamma/\alpha})$$

- Typically $\alpha = 1/2$, $\gamma = 1$: for $\varepsilon = 10^{-3}$ we have Cost $= \mathcal{O}(10^{12})$!

Multilevel Monte Carlo [Heinrich, '01], [Giles, '07]

The multilevel method works on a **sequence of levels**, s.t. $M_\ell = sM_{\ell-1}$ and $J_\ell = sJ_{\ell-1}$, $\ell = 0, 1, \ldots, L$, and set $Q_\ell = Q_{M_\ell, J_\ell}$.

Multilevel Monte Carlo [Heinrich, '01], [Giles, '07]

The multilevel method works on a **sequence of levels**, s.t. $M_\ell = sM_{\ell-1}$ and $J_\ell = sJ_{\ell-1}$, $\ell = 0, 1, \ldots, L$, and set $Q_\ell = Q_{M_\ell, J_\ell}$.

Linearity of expectation gives us

$$\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}[Q_\ell - Q_{\ell-1}]$$

Define the following **multilevel MC estimator** for $\mathbb{E}[Q]$:

$$\widehat{Q}_L^{\mathrm{ML}} := \widehat{Q}_0^{\mathrm{MC}} \; + \; \sum_{\ell=1}^{L} \widehat{(Q_\ell - Q_{\ell-1})}^{\mathrm{MC}}$$

Terms are estimated **independently**, with N_ℓ samples on level ℓ.

Multilevel Monte Carlo [Heinrich, '01], [Giles, '07]

The **mean square error** of the this estimator is

$$\mathbb{E}\Big[\big(\hat{Q}_L^{\mathrm{ML}} - \mathbb{E}[Q]\big)^2\Big] = \underbrace{\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]}_{\text{sampling error}} + \underbrace{\big(\mathbb{E}[\hat{Q}_L^{\mathrm{ML}}] - \mathbb{E}[Q]\big)^2}_{\text{model error}}$$

Multilevel Monte Carlo [Heinrich, '01], [Giles, '07]

The **mean square error** of the this estimator is

$$\mathbb{E}\Big[\big(\hat{Q}_L^{\mathrm{ML}} - \mathbb{E}[Q]\big)^2\Big] = \underbrace{\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]}_{\text{sampling error}} + \underbrace{\big(\mathbb{E}[\hat{Q}_L^{\mathrm{ML}}] - \mathbb{E}[Q]\big)^2}_{\text{model error}}$$

$$= \frac{\mathbb{V}[Q_0]}{N_0} + \sum_{\ell=1}^{L} \frac{\mathbb{V}[Q_\ell - Q_{\ell-1}]}{N_\ell} + \big(\mathbb{E}[Q_L - Q]\big)^2$$

- N_0 still needs to be large, **but** samples are much cheaper to obtain on coarser level
- N_ℓ $(\ell > 0)$ much smaller, **since** $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ as $M_\ell \to \infty$

Complexity of Multilevel Monte Carlo

Assume **(A1)** (model error $\mathcal{O}(M_\ell^{-\alpha})$), **(A2)** (cost/sample $\mathcal{O}(M_\ell^{\gamma})$) and

(A3) $\mathbb{V}[Q_\ell - Q_{\ell-1}] = \mathcal{O}(M_\ell^{-\beta})$

with $2\alpha \geq \min(\beta, \gamma)$. Then there exist L and $\{N_\ell\}$ such that the **total cost** to obtain a **mean square error**

$$\mathbb{E}\left[(\hat{Q}_L^{\mathrm{ML}} - \mathbb{E}[Q])^2\right] \; = \; \mathcal{O}(\varepsilon^2)$$

is

$$\mathrm{Cost}(\hat{Q}_L^{\mathrm{ML}}) \; = \; \begin{cases} \mathcal{O}(\varepsilon^{-2}) & \text{if} \;\; \beta > \gamma \\ \mathcal{O}(\varepsilon^{-2} \, \log(\varepsilon)^2) & \text{if} \;\; \beta = \gamma \\ \mathcal{O}(\varepsilon^{-2 - (\gamma - \beta)/\alpha}) & \text{if} \;\; \beta < \gamma \end{cases}$$

- $\{N_\ell\}$ chosen to minimise cost for a fixed variance

Convergence analysis of MLMC

Can prove that for typical 2D model problems in subsurface flow, **(A1)** and **(A3)** are satisfied with $\alpha = 1/2$, $\beta = 1$. With an optimal linear solver $(\gamma = 1)$, the computational costs are bounded by:

d	**MLMC**	**MC**
2	$\mathcal{O}(\varepsilon^{-2})$	$\mathcal{O}(\varepsilon^{-4})$

Convergence analysis of MLMC

Can prove that for typical 2D model problems in subsurface flow, **(A1)** and **(A3)** are satisfied with $\alpha = 1/2$, $\beta = 1$. With an optimal linear solver $(\gamma = 1)$, the computational costs are bounded by:

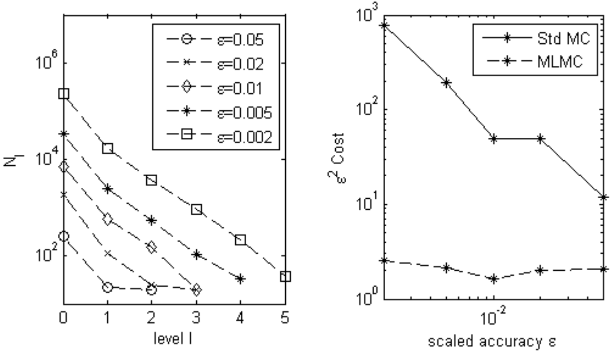
d	**MLMC**	**MC**
2	$\mathcal{O}(10^6)$	$\mathcal{O}(10^{12})$

0.1 % accuracy $\longrightarrow \varepsilon = 10^{-3}$

Numerical Example (multilevel MC)

$D = (0,1)^2$, $Q = \|p\|_{L^2(D)}$, $J_L = \infty$, $M_0 = 8^2$.

Typical **2D model problem**.



Left: Number of samples per level. *Right:* Total computational cost.

Can the multilevel idea be extended to Markov chain Monte Carlo?

Incorporating data - Bayesian approach

Recall:

$$\mathbf{Z}_J(\omega) \in \mathbb{R}^J \xrightarrow{\text{Model}(M)} \mathbf{X}_M(\omega) \in \mathbb{R}^M \xrightarrow{\text{Output}} Q_{M,J}(\omega) \in \mathbb{R}$$

random input — state vector — quantity of interest

- **"Prior"** in our model was multivariate Gaussian $\mathbf{Z}_J := [Z_1, \ldots, Z_J]$:

$$\mathcal{P}(\mathbf{Z}_J) \eqsim (2\pi)^{-J/2} \prod_{j=1}^J \exp\left(-\frac{Z_j^2}{2}\right)$$

- Usually **data** F_{obs} related to **outputs** (e.g. pressure) also available. To reduce uncertainty, incorporate F_{obs} $\rightarrow$ the **"posterior"**

Incorporating data - Bayesian approach

Bayes' Theorem: (RHS computable! Proportionality constant $1/\mathcal{P}(F_{\text{obs}})$ not!)

$$\underbrace{\pi^{M,J}(\mathbf{Z}_J)}_{\textbf{posterior}} := \mathcal{P}(\mathbf{Z}_J \,|\, F_{\text{obs}}) \eqsim \underbrace{\mathcal{L}_M(F_{\text{obs}} \,|\, \mathbf{Z}_J)}_{\textbf{likelihood}} \underbrace{\mathcal{P}(\mathbf{Z}_J)}_{\textbf{prior}}$$

Incorporating data - Bayesian approach

Bayes' Theorem: (RHS computable! Proportionality constant $1/\mathcal{P}(F_{\text{obs}})$ not!)

$$\underbrace{\pi^{M,J}(\mathbf{Z}_J)}_{\textbf{posterior}} := \mathcal{P}(\mathbf{Z}_J \,|\, F_{\text{obs}}) \;\eqsim\; \underbrace{\mathcal{L}_M(F_{\text{obs}} \,|\, \mathbf{Z}_J)}_{\textbf{likelihood}} \, \underbrace{\mathcal{P}(\mathbf{Z}_J)}_{\textbf{prior}}$$

- **Likelihood model** (e.g. Gaussian):

$$\mathcal{L}_M(F_{\text{obs}} \,|\, \mathbf{Z}_J) \;\eqsim\; \exp\left(\frac{-\|F_{\text{obs}} - F_M(\mathbf{Z}_J)\|^2}{\sigma^2_{\text{fid},M}}\right)$$

$F_M(\mathbf{Z}_J)$ ... model response; $\sigma_{\text{fid},M}$ ... fidelity parameter (M-dep.)

ALGORITHM 1. (Standard Metropolis Hastings MCMC)

- Choose $\mathbf{Z}_J^0$.
- At state $\mathbf{Z}_J^n$ generate proposal $\mathbf{Z}'_J$ from distribution $q(\mathbf{Z}'_J \,|\, \mathbf{Z}_J^n)$ (for simplicity symmetric, e.g. random walk).
- Accept sample $\mathbf{Z}'_J$ with probability $\alpha^{M,J} = \min\left(1, \frac{\pi^{M,J}(\mathbf{Z}'_J)}{\pi^{M,J}(\mathbf{Z}_J^n)}\right)$,

  i.e. $\mathbf{Z}_J^{n+1} = \mathbf{Z}'_J$ with probability $\alpha^{M,J}$; otherwise stay at $\mathbf{Z}_J^{n+1} = \mathbf{Z}_J^n$.

ALGORITHM 1. (Standard Metropolis Hastings MCMC)

- Choose $\mathbf{Z}_J^0$.
- At state $\mathbf{Z}_J^n$ generate proposal $\mathbf{Z}'_J$ from distribution $q(\mathbf{Z}'_J \,|\, \mathbf{Z}_J^n)$ (for simplicity symmetric, e.g. random walk).
- Accept sample $\mathbf{Z}'_J$ with probability $\alpha^{M,J} = \min\left(1, \frac{\pi^{M,J}(\mathbf{Z}'_J)}{\pi^{M,J}(\mathbf{Z}_J^n)}\right)$,

  i.e. $\mathbf{Z}_J^{n+1} = \mathbf{Z}'_J$ with probability $\alpha^{M,J}$; otherwise stay at $\mathbf{Z}_J^{n+1} = \mathbf{Z}_J^n$.

Samples $\mathbf{Z}_J^n$ used as usual for inference (even though not i.i.d.):

$$\mathbb{E}_{\pi^{M,J}}[Q] \;\approx\; \mathbb{E}_{\pi^{M,J}}[Q_{M,J}] \;\approx\; \frac{1}{N}\sum_{n=1}^{N} Q_{M,J}^{(n)} := \widehat{Q}^{\text{MetH}}$$

Pros:

- Produces a Markov chain $\{\mathbf{Z}_J^n\}_{n\in\mathbb{N}}$, with $\mathbf{Z}_J^n \sim \pi^{M,J}$ as $n \to \infty$.

Cons:

- Evaluating $\alpha^{M,J}$ expensive for large M.
- Acceptance rate $\alpha^{M,J}$ very low for large J $(< 10\%)$.

Multilevel Markov Chain Monte Carlo

Key ingredients in multilevel method:

- Models with less DOFs on coarser levels **much cheaper** to solve
- $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ as $\ell \to \infty$ $\Rightarrow$ **fewer samples** on finer levels
- **Telescoping sum:** $\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}[Q_\ell] - \mathbb{E}[Q_{\ell-1}]$

Multilevel Markov Chain Monte Carlo

Key ingredients in multilevel method:

- Models with less DOFs on coarser levels **much cheaper** to solve
- $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ as $\ell \to \infty$ $\Rightarrow$ **fewer samples** on finer levels
- **Telescoping sum:** $\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}[Q_\ell] - \mathbb{E}[Q_{\ell-1}]$

In MCMC setting target distribution depends on ℓ, so need to define multilevel estimator carefully!

$$\mathbb{E}_{\pi^L}[Q_L] = \mathbb{E}_{\pi^0}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}_{\pi^\ell}[Q_\ell] - \mathbb{E}_{\pi^{\ell-1}}[Q_{\ell-1}]$$

Multilevel Markov Chain Monte Carlo

Key ingredients in multilevel method:

- Models with less DOFs on coarser levels **much cheaper** to solve
- $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ as $\ell \to \infty$ $\Rightarrow$ **fewer samples** on finer levels
- **Telescoping sum:** $\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}[Q_\ell] - \mathbb{E}[Q_{\ell-1}]$

In MCMC setting target distribution depends on ℓ, so need to define multilevel estimator carefully!

$$\mathbb{E}_{\pi^L}[Q_L] = \underbrace{\mathbb{E}_{\pi^0}[Q_0]}_{\text{standard MCMC}} + \sum_{\ell=1}^{L} \underbrace{\mathbb{E}_{\pi^\ell}[Q_\ell] - \mathbb{E}_{\pi^{\ell-1}}[Q_{\ell-1}]}_{\text{2 level MCMC (NEW)}}$$

$$\widehat{Q}_L^{\text{MLMetH}} := \frac{1}{N_0}\sum_{n=1}^{N_0} Q_0(\mathbf{Z}_0^n) + \sum_{\ell=1}^{L} \frac{1}{N_\ell}\sum_{n=1}^{N_\ell} \left(Q_\ell(\mathbf{Z}_\ell^n) - Q_{\ell-1}(\mathbf{Z}_{\ell-1}^n)\right)$$

Multilevel Markov Chain Monte Carlo

Key ingredients in multilevel method:

- Models with less DOFs on coarser levels **much cheaper** to solve
- $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ as $\ell \to \infty$ $\Rightarrow$ **fewer samples** on finer levels
- **Telescoping sum:** $\mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^{L} \mathbb{E}[Q_\ell] - \mathbb{E}[Q_{\ell-1}]$

In MCMC setting target distribution depends on ℓ, so need to define multilevel estimator carefully!

$$\mathbb{E}_{\pi^L}[Q_L] = \underbrace{\mathbb{E}_{\pi^0}[Q_0]}_{\text{standard MCMC}} + \sum_{\ell=1}^{L} \underbrace{\mathbb{E}_{\pi^\ell}[Q_\ell] - \mathbb{E}_{\pi^{\ell-1}}[Q_{\ell-1}]}_{\text{2 level MCMC (NEW)}}$$

$$\widehat{Q}_L^{\mathrm{MLMetH}} := \frac{1}{N_0}\sum_{n=1}^{N_0} Q_0(\mathbf{Z}_0^n) + \sum_{\ell=1}^{L} \frac{1}{N_\ell}\sum_{n=1}^{N_\ell}\left(Q_\ell(\mathbf{Z}_\ell^n) - Q_{\ell-1}(\mathbf{Z}_{\ell-1}^n)\right)$$

Idea: Split $\mathbf{Z}_\ell^n = [\mathbf{z}_{\ell,\mathsf{C}}^n, \mathbf{z}_{\ell,\mathsf{F}}^n]$, where $\mathbf{z}_{\ell,\mathsf{C}}^n$ has length $J_{\ell-1}$.

ALGORITHM 2 (Two-level Metropolis Hastings MCMC for $Q_\ell - Q_{\ell-1}$)

At states $\mathbf{Z}^n_{\ell-1}, \mathbf{Z}^n_\ell$ (of the **independent** level ℓ and level $\ell - 1$ Markov chains)

1. Generate new state $\mathbf{Z}^{n+1}_{\ell-1}$ using Algorithm 1.

ALGORITHM 2 (Two-level Metropolis Hastings MCMC for $Q_\ell - Q_{\ell-1}$)

At states $\mathbf{Z}^n_{\ell-1}, \mathbf{Z}^n_\ell$ (of the **independent** level ℓ and level $\ell - 1$ Markov chains)

1. Generate new state $\mathbf{Z}^{n+1}_{\ell-1}$ using Algorithm 1.
2. Propose $\mathbf{Z}'_\ell = [\mathbf{Z}^{n+1}_{\ell-1}, \mathbf{z}'_{\ell,\mathsf{F}}]$ with $\mathbf{z}'_{\ell,\mathsf{F}}$ generated via random walk.

   (novel transition prob. q^{ML} depends on level $\ell - 1$ acceptance prob. $\alpha^{\ell-1}$)

ALGORITHM 2 (Two-level Metropolis Hastings MCMC for $Q_\ell - Q_{\ell-1}$)

At states $\mathbf{Z}^n_{\ell-1}, \mathbf{Z}^n_\ell$ (of the **independent** level ℓ and level $\ell - 1$ Markov chains)

1. Generate new state $\mathbf{Z}^{n+1}_{\ell-1}$ using Algorithm 1.
2. Propose $\mathbf{Z}'_\ell = [\mathbf{Z}^{n+1}_{\ell-1}, \mathbf{z}'_{\ell,\mathsf{F}}]$ with $\mathbf{z}'_{\ell,\mathsf{F}}$ generated via random walk.

   (novel transition prob. q^{ML} depends on level $\ell - 1$ acceptance prob. $\alpha^{\ell-1}$)
3. Accept $\mathbf{Z}'_\ell$ with probability

$$\alpha^\ell(\mathbf{Z}'_\ell \,|\, \mathbf{Z}^n_\ell) = \min\left(1, \frac{\pi^\ell(\mathbf{Z}'_\ell)\, \mathrm{q}^{\mathsf{ML}}(\mathbf{Z}^n_\ell \,|\, \mathbf{Z}'_\ell)}{\pi^\ell(\mathbf{Z}^n_\ell)\, \mathrm{q}^{\mathsf{ML}}(\mathbf{Z}'_\ell \,|\, \mathbf{Z}^n_\ell)}\right)$$

ALGORITHM 2 (Two-level Metropolis Hastings MCMC for $Q_\ell - Q_{\ell-1}$)

At states $\mathbf{Z}^n_{\ell-1}, \mathbf{Z}^n_\ell$ (of the **independent** level ℓ and level $\ell - 1$ Markov chains)

1. Generate new state $\mathbf{Z}^{n+1}_{\ell-1}$ using Algorithm 1.
2. Propose $\mathbf{Z}'_\ell = [\mathbf{Z}^{n+1}_{\ell-1}, \mathbf{z}'_{\ell,\mathsf{F}}]$ with $\mathbf{z}'_{\ell,\mathsf{F}}$ generated via random walk.

   (novel transition prob. q^{ML} depends on level $\ell - 1$ acceptance prob. $\alpha^{\ell-1}$)
3. Accept $\mathbf{Z}'_\ell$ with probability

$$\alpha^\ell(\mathbf{Z}'_\ell \,|\, \mathbf{Z}^n_\ell) = \min\left(1, \frac{\pi^\ell(\mathbf{Z}'_\ell)\pi^{\ell-1}(\mathbf{z}^n_{\ell,\mathsf{C}})}{\pi^\ell(\mathbf{Z}^n_\ell)\pi^{\ell-1}(\mathbf{Z}^{n+1}_{\ell-1})}\right)$$

   Follows quite easily & both terms have been computed previously.

Example: one step

Given $\mathbf{Z}^n_{\ell-1}$ and $\mathbf{Z}^n_\ell$, the possible $(n+1)^{th}$ states of the two chains are:

Level $\ell-1$ test	Level ℓ test	$\mathbf{Z}^{n+1}_{\ell-1}$	$\mathbf{Z}^{n+1}_{\ell}$
reject	accept	$\mathbf{Z}^n_{\ell-1}$	$[\mathbf{Z}^n_{\ell-1}, \mathbf{z}'_{\ell,\mathsf{F}}]$
accept	accept	$\mathbf{Z}'_{\ell-1}$	$[\mathbf{Z}'_{\ell-1}, \mathbf{z}'_{\ell,\mathsf{F}}]$
reject	reject	$\mathbf{Z}^n_{\ell-1}$	$[\mathbf{z}^n_{\ell,\mathsf{C}}, \mathbf{z}^n_{\ell,\mathsf{F}}]$
accept	reject	$\mathbf{Z}'_{\ell-1}$	$[\mathbf{z}^n_{\ell,\mathsf{C}}, \mathbf{z}^n_{\ell,\mathsf{F}}]$

Convergence analysis of MLMCMC

What can we prove?

- We have a genuine **Markov chain** on every level.
- Multilevel algorithm is **consistent** (no bias between levels).
- Multilevel algorithm **converges** for any initial state.
- **Same Complexity Theorem** as for Multilevel Monte Carlo.
  (completely abstract and applicable also in DA for NWP)
  - For typical 2D model problems in subsurface flow, we have $\alpha = 1/2$, $\beta = 1/2$. With an optimal linear solver $(\gamma = 1)$, the computational costs are bounded by:

d	**MLMCMC**	**MCMC**
2	$\mathcal{O}(\varepsilon^{-3})$	$\mathcal{O}(\varepsilon^{-4})$

Convergence analysis of MLMCMC

What can we prove?

- We have a genuine **Markov chain** on every level.
- Multilevel algorithm is **consistent** (no bias between levels).
- Multilevel algorithm **converges** for any initial state.
- **Same Complexity Theorem** as for Multilevel Monte Carlo.
  (completely abstract and applicable also in DA for NWP)
  - For typical 2D model problems in subsurface flow, we have $\alpha = 1/2$, $\beta = 1/2$. With an optimal linear solver $(\gamma = 1)$, the computational costs are bounded by:

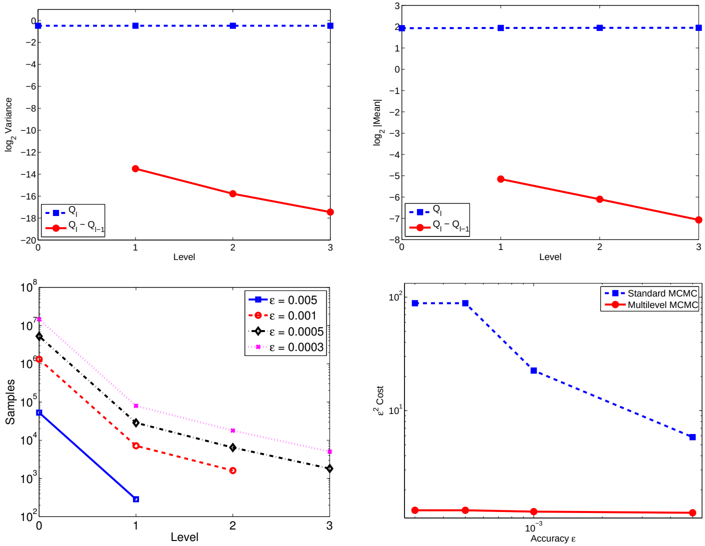
d	MLMCMC	MCMC
2	$\mathcal{O}(10^9)$	$\mathcal{O}(10^{12})$

0.1 % accuracy $\longrightarrow \varepsilon = 10^{-3}$

Numerical Example (multilevel MCMC)

$D = (0,1)^2$, $Q = k_{\text{eff}}$, $J_L = 169$, $M_0 = 16^2$.

Data (artificial): Pressure p at 9 random points in domain.



Conclusions

- Standard (Markov chain) Monte Carlo algorithms are often prohibitively expensive.
- Multilevel versions greatly reduce the cost.
- Multilevel algorithms are generally applicable.
- Full convergence analysis available.