Multi-Index Quasi-Monte Carlo

An algorithm for simulating PDEs with random coefficients

Pieterjan Robbe

joint work with D. Nuyens and S. Vandewalle

pieterjan.robbe@cs.kuleuven.be

August 16, 2016

Department of Computer Science

The missing link?

	Monte Carlo	Quasi-Monte Carlo
Multilevel	MLMC[a,b]	MLQMC[c,d,e]
Multi-index	MIMC[f]	?

a K. A. Cliffe, M. B. Giles, R. Scheichl and A. L. Teckentrup, *Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients*, Computing and Visualization in Science, 14 (2011), pp. 3-15. [Cliffe, 2011]

b A. Barth, C. Schwab and N. Zollinger, *Multilevel Monte Carlo Finite Element Method for Elliptic PDEs with Stochastic Coefficients*, Numerische Mathematik, 119 (2011), pp. 123-161. [Barth, 2011]

c F. Y. Kuo, R. Scheichl, C. Schwab, I. H. Sloan and E. Ullmann, *Multilevel Quasi-Monte Carlo Methods for Lognormal Diffusion Problems*, Mathematics of Computation, under revision, 2016. [Kuo, 2016]

d F. Y. Kuo and D. Nuyens, *Application of Quasi-Monte Carlo Methods to Elliptic PDEs with Random Diffusion Coefficients – A Survey of Analysis and Implementation*, Foundations of Computational Mathematics, to appear, 2016. [Kuo, 2016/2]

e I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, C. Schwab and I. H. Sloan, *Quasi-Monte Carlo Finite Element Methods for Elliptic PDEs with Lognormal Random Coefficients*, Numerische Mathematik, 131 (2014), pp. 329-368. [Graham, 2014]

f A.-L. Haji-Ali, F. Nobile and R. Tempone, *Multi-Index Monte Carlo: When Sparsity Meets Sampling*, Numerische Mathematik 132 (2016), pp. 767-806. [Haji-Ali, 2016]

Contents

PDEs with Random Coefficients

PDEs with random coefficients

- Let $(\Omega, \mathcal{A}, P)$ be a complete probability space and $D \subset \mathbb{R}^d$ a bounded domain
- We look for solutions $u : D \to \mathbb{R}$ that solve almost surely (a.s.)

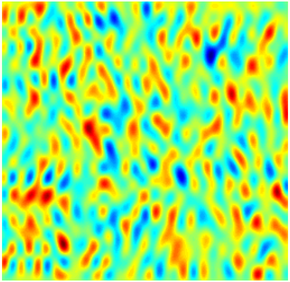
$$-\nabla \cdot (a(\boldsymbol{x}, \omega)\nabla u(\boldsymbol{x}, \omega)) = f(\boldsymbol{x}) \quad \text{for } \boldsymbol{x} \in D \text{ and } \omega \in \Omega$$

with given boundary conditions

$$\begin{aligned} u(\boldsymbol{x}, \cdot) &= u_1(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \partial_1 D \\ n(\boldsymbol{x}) \cdot (a(\boldsymbol{x}, \cdot)\nabla u(\boldsymbol{x}, \cdot)) &= u_2(\boldsymbol{x}) & \text{for } \boldsymbol{x} \in \partial_2 D \end{aligned}$$

- The **goal** is to compute statistics of the functional $G = \mathcal{G}(u)$ for some sufficiently smooth a and $\mathcal{G}$

Flow through porous media



- Represent permeability of the porous medium as a lognormal **random field** $k(\boldsymbol{x}, \omega) = \exp(Z(\boldsymbol{x}, \omega))$, where $Z(\boldsymbol{x}, \omega)$ is an underlying Gaussian random field
- A Gaussian random field is characterised by its **mean** $\mu(\boldsymbol{x}) = \mathbb{E}[Z(\boldsymbol{x})]$ and **covariance function** $\mathrm{cov}(Z(\boldsymbol{x}), Z(\boldsymbol{y}))$
- Every fixed $\omega \in \Omega$ yields a deterministic **realisation** of the random field

On the choice of covariance function

- A common choice is the **exponential covariance** function

$$C(\rho) = \sigma^2 \exp\left(\frac{\rho}{\lambda}\right)$$

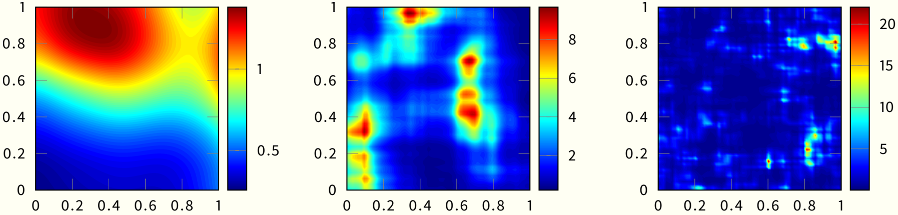
  with $\rho = \|\boldsymbol{x} - \boldsymbol{y}\|_p$ and p the usual ℓ_p-norm
- We choose the more general **Matérn covariance** function

$$C(\rho) = \sigma^2 \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\sqrt{2\nu}\frac{\rho}{\lambda}\right)^{\nu} K_\nu \left(\sqrt{2\nu}\frac{\rho}{\lambda}\right)$$

  with Γ the Gamma function and K_ν the modified Bessel function of the second kind
- The Matérn covariance reduces to the exponential case when $\nu = 1/2$, $\nu = \infty$ is the squared exponential case.

On the choice of covariance function

- Some examples of $k(\boldsymbol{x}, \omega)$ for different parameter sets



$\lambda = 1$	$\lambda = 0.3$	$\lambda = 0.075$
$\sigma^2 = 1$	$\sigma^2 = 1$	$\sigma^2 = 1$
$\nu = 2.5$	$\nu = 1$	$\nu = 0.5$

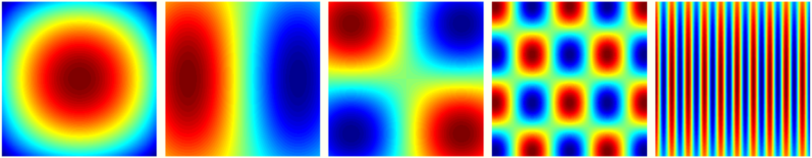
The KL expansion

- Classical technique to take samples from $k(\boldsymbol{x},\omega)$ is the Karhunen–Loève or **KL-expansion**

$$k(\boldsymbol{x},\omega) = \bar{k} + \exp\left(\sum_{r=1}^{\infty} \sqrt{\theta_r} f_r(\boldsymbol{x})\xi(\omega)\right)$$

with θ_r and f_r the solutions of the Fredholm equation

$$\int_D C(\boldsymbol{x},\boldsymbol{y}) f_r(\boldsymbol{y})\mathrm{d}\boldsymbol{y} = \lambda_r f_r(\boldsymbol{y}), \quad \boldsymbol{x} \in D$$

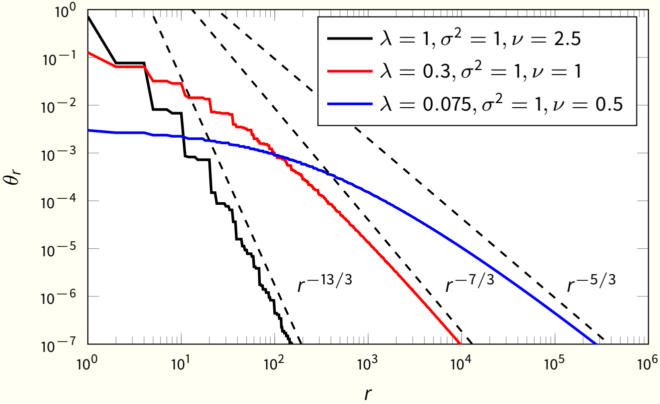


$r = 1$ $r = 2$ $r = 4$ $r = 15$ $r = 88$

The KL expansion

- Approximation quality of the KL expansion determined by eigenvalue decay rate

$$\theta_r \sim \mathcal{O}\left(r^{-\frac{4\nu+d}{d}}\right)$$



Darcy flow

- Steady-state flow through porous media given by **Darcy's law**

$$-\nabla \cdot k(\boldsymbol{x},\omega)\nabla p(\boldsymbol{x},\omega) = f(\boldsymbol{x})$$

  with k the hydraulic conductivity and p the pressure head
- This is an elliptic **PDE with random coefficients**!
- Let us choose $D = [0,1]^3$ and the functionals

$$G^1 = p(\boldsymbol{x}^*), \;\; \boldsymbol{x}^* = (0.5, 0.5, 0.5)$$

$$\text{and} \quad G^2 = -\int_0^1 \int_0^1 k \left.\frac{\partial p}{\partial x}\right|_{x=1} \mathrm{d}y\mathrm{d}z$$

- Our **goal** is to compute $\mathbb{E}[G^1]$ and $\mathbb{E}[G^2]$

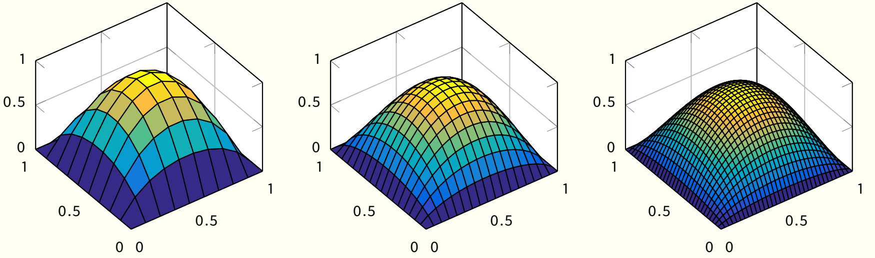
From MLMC to MIMC

Numerical approximation

- We assume that different approximations of u are available based on the discretization parameter h
- For MLMC, one usually assumes a **geometrical** mesh hierarchy

$$h_\ell = h_0 2^{-\ell}$$

- The approximations are indexed by a **level** $\ell \in \mathbb{N}$



$\ell = 0$ $\qquad$ $\ell = 1$ $\qquad$ $\ell = 2$

Multilevel Monte Carlo

- Suppose we denote by G_ℓ the approximation of G at level ℓ
- Let $\Delta G_\ell := \begin{cases} G_\ell & \text{if } \ell = 0 \\ G_\ell - G_{\ell-1} & \text{if } \ell > 0 \end{cases}$
- Using the **telescoping sum** identity,

$$\mathbb{E}[G_L] = \sum_{\ell=0}^{L} \mathbb{E}[\Delta G_\ell]$$

  the MLMC [Giles, 2008] estimator can be expressed as

$$\hat{G}_{\mathrm{ML}} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \Delta G_\ell(\omega_{\ell,n})$$

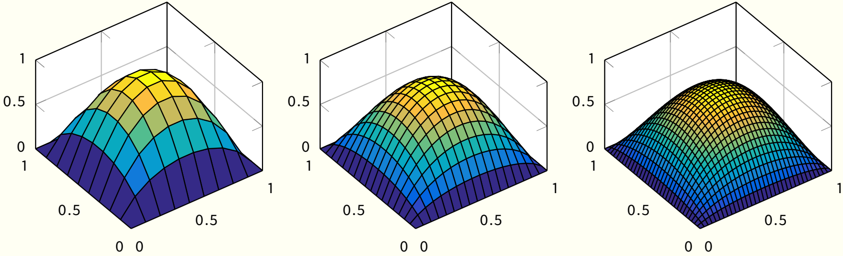
- **Main point**: MLMC is a recursive control variate strategy that uses a hierarchy of coarser grids

Numerical approximation (2)

- We assume that different approximations of u are available based on the discretization parameters h_i, $i = 1 \dots d$
- Similar to MLMC, assume a **geometrical** mesh hierarchy

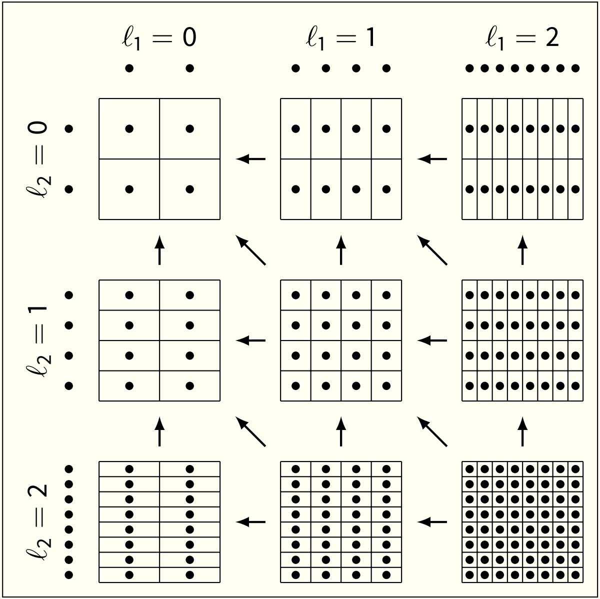
$$h_{\ell_i} = h_0 2^{-\ell_i}$$

- Approximations are indexed by a **multi-index** $\vec{\boldsymbol{\ell}} = (\ell_i)_{i=1}^{d} \in \mathbb{N}$



$\ell = (0, 0)$ $\ell = (1, 1)$ $\ell = (2, 2)$

Multi-Index Monte Carlo



Multi-Index Monte Carlo

- Suppose we denote by $G_{\vec{\ell}}$ the approximation of G at index $\vec{\ell}$
- Let $\Delta G_{\vec{\ell}} := \left(\bigotimes_{i=1}^{d} \Delta_i\right) G_{\vec{\ell}}$ with

$$\Delta_i G_{\vec{\ell}} = \begin{cases} G_{\vec{\ell}} & \text{if} \quad \ell_i = 0 \\ G_{\vec{\ell}} - G_{\vec{\ell}-\vec{e}_i} & \text{if} \quad \ell_i > 0 \end{cases} \quad \text{for } i = 1 \ldots d$$

- The MIMC estimator [Haji-Ali, 2016] can be expressed as

$$\hat{G}_{\mathrm{MI}} = \sum_{\vec{\ell} \in \mathcal{I}} \frac{1}{N_{\vec{\ell}}} \sum_{n=1}^{N_{\vec{\ell}}} \Delta G_{\vec{\ell}}(\omega_{\vec{\ell},n})$$

where $\mathcal{I} \subset \mathbb{N}^d$ is an appropriate set of indices, the **index set**

Example

- For $d = 2$ and $\vec{\ell} = (1, 2)$, we have

$$\begin{aligned}
\Delta G_{(1,2)}(\omega) &= \Delta_2 \left(\Delta_1 G_{(1,2)}(\omega)\right) \\
&= \Delta_2 \left(G_{(1,2)}(\omega) - G_{(0,2)}(\omega)\right) \\
&= (G_{(1,2)}(\omega) - G_{(0,2)}(\omega)) - (G_{(1,1)}(\omega) - G_{(0,1)}(\omega)) \\
&= G_{(1,2)}(\omega) - G_{(0,2)}(\omega) - G_{(1,1)}(\omega) + G_{(0,1)}(\omega)
\end{aligned}$$

- **Key point** is that these 4 solutions are based on the same realisation of the random field $k(\cdot, \omega)$
- In general, computing a single realisation of $\Delta G_{\vec{\ell}}$ requires the solution of 2^d different PDEs, where the cost is dominated by the solution at index $\vec{\ell}$

Complexity analysis

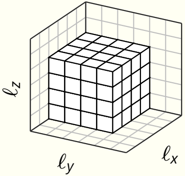
- Classical MSE error splitting: $\text{MSE} = \mathbb{V}[\hat{G}] + \mathbb{E}[(\hat{G} - G)]^2$
- By independence, we have $\mathbb{V}[\hat{G}] = \sum_{\vec{\ell} \in \mathcal{I}} \frac{V_{\vec{\ell}}}{N_{\vec{\ell}}}$ with $V_{\vec{\ell}} := \mathbb{V}[\Delta G_{\vec{\ell}}]$
- The second term can be approximated as

$$\mathbb{E}[(\hat{G} - G)] = \left| \sum_{\vec{\ell} \notin \mathcal{I}} \mathbb{E}[\Delta G_{\vec{\ell}}] \right| \approx \left| \sum_{\vec{\ell} \in \partial \mathcal{I}} \mathbb{E}[\Delta G_{\vec{\ell}}] \right|$$

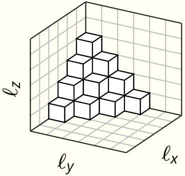
  with $\partial \mathcal{I}$ the **boundary** of the index set $\mathcal{I}$. Note: this requires sufficient decay of the $\mathbb{E}[\Delta G_{\vec{\ell}}]$ (e.g., exponential, see [Haji-Ali, 2016])
- **Conclusion**: error and work complexity depends on choice of index set!

On the choice of index set

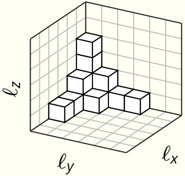
$$\mathcal{I}(L) = \left\{ \vec{\boldsymbol{\ell}} \in \mathbb{N}^d : \ell_i \leq L \text{ for all } 1 \leq i \leq d \right\}$$



$$\mathcal{I}(L) = \left\{ \vec{\boldsymbol{\ell}} \in \mathbb{N}^d : \sum_{i=1}^{d} \ell_i \leq L \right\}$$

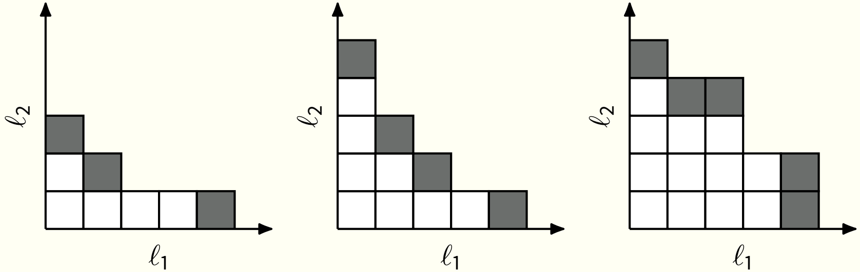


$$\mathcal{I}(L) = \left\{ \vec{\boldsymbol{\ell}} \in \mathbb{N}^d : \prod_{i=1}^{d} \max(1, \ell_i) \leq L \right\}$$



Boundary of an arbitrary index set

- $\mathcal{I}$ is valid when it is **downward closed**
- An index $\vec{\ell}$ **dominates** an index $\vec{\tau}$ in a certain direction *i* if $\ell_i > \tau_i$
- An index $\vec{\ell}$ is an **interior index** if it is dominated by another index in the direction of the position of its maximum entry
- The **boundary** $\partial\mathcal{I}$ of an index set $\mathcal{I}$ are these indices $\vec{\ell} \in \mathcal{I}$ that are part of the index set but are not interior indices



From MIMC to MIQMC

The MIQMC estimator

- Let us assume that we have a **randomly shifted rank-1 lattice rule** available that approximates

$$I_s(f) = \int_{[0,1]^s} f(\boldsymbol{y}) \mathrm{d}\boldsymbol{y} \approx \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{N} \sum_{n=0}^{N-1} f(\{\boldsymbol{t}_n + \boldsymbol{\Delta}_k\})$$

(see tutorials on Sunday, or first plenary talk by D. Nuyens)

- In our setting, the “integral” is the expectation

$$\begin{aligned} \mathbb{E}[G_{\vec{\boldsymbol{\ell}}}] &= \int_{\mathbb{R}^\infty} G_{\vec{\boldsymbol{\ell}}}(u(\cdot, \xi_1, \xi_2 \ldots)) \mathrm{d}\Phi(\xi_1, \xi_2 \ldots) \\ &\approx \int_{[0,1]^s} G_{\vec{\boldsymbol{\ell}}}(u(\cdot, \Phi^{-1}(y_1), \ldots, \Phi^{-1}(y_s), 0, \ldots) \mathrm{d}\boldsymbol{y} \end{aligned}$$

with Φ the cumulative normal density and we apply the change of variables component-wise

$$\boldsymbol{\xi} = \Phi^{-1}(\boldsymbol{y}) = (\Phi^{-1}(y_1), \Phi^{-1}(y_2), \ldots) \in \mathbb{R}^{\mathbb{N}} \text{ and } \boldsymbol{y} \in (0,1)^{\mathbb{N}}$$

The MIQMC estimator

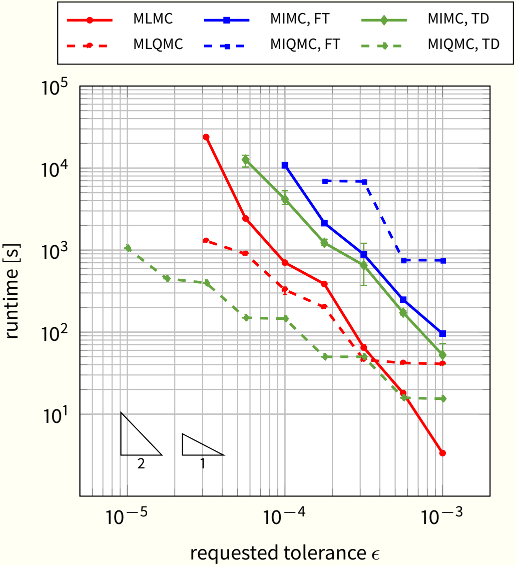
- The MIQMC estimator reads

$$\hat{G}^{\star}_{\mathrm{MI}} = \sum_{\vec{\ell}\in\mathcal{I}} \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{N_{\vec{\ell}}} \sum_{n=0}^{N_{\vec{\ell}}-1} \Delta G_{\vec{\ell}}(\Phi^{-1}(\{\boldsymbol{t}_n + \boldsymbol{\Delta}_{k,\vec{\ell}}\}))$$

- For linear functionals G in a single- or multilevel setting, it can be show that the integrand belongs to a weighted Sobolev space with (S)POD-weights, see [Graham, 2014], [Kuo, 2016] and [Kuo, 2016/2]
- Software for generating these lattice rules: QMC4PDE
- The convergence rate depends on decay rate of eigenvalues, but is independent of the number of dimensions s

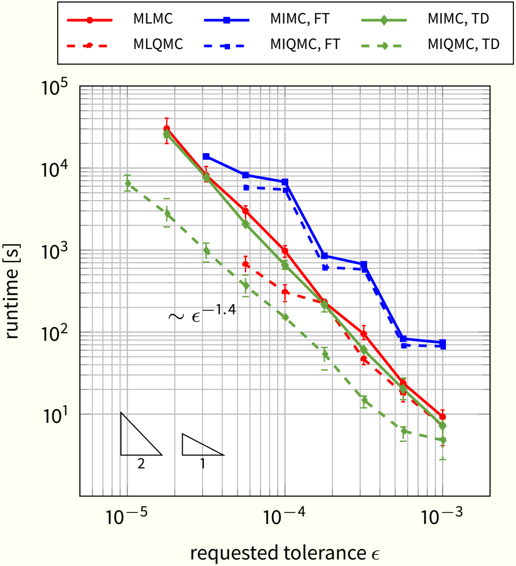
Results

Numerical results (1)



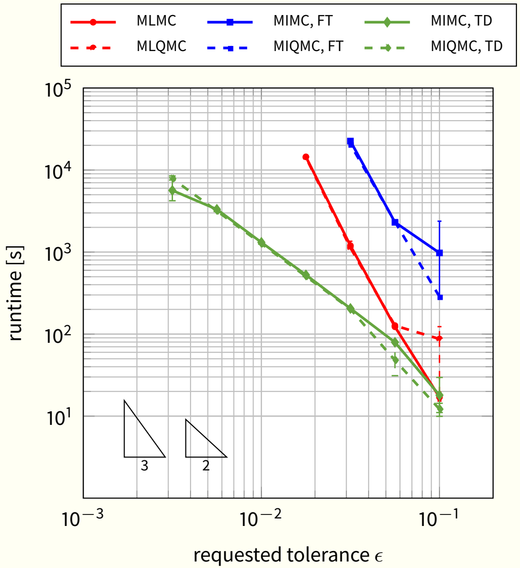
- Three-dimensional problem with Dirichlet boundary conditions
- Quantity of interest is G^1 (point evaluation)
- Lognormal random field with Matérn covariance, $\lambda = 1, \nu = 2.5, s = 12$
- Solve deterministic PDE using FV, CG with AMG preconditioner
- Coarsest mesh has $4^3 = 64$ DOF, finest mesh has $128^3 \approx 2M$ DOF

Numerical results (2)



- Three-dimensional problem with Dirichlet boundary conditions
- Quantity of interest is G^1 (point evaluation)
- Lognormal random field with Matérn covariance, $\lambda = 0.3, \nu = 1, s = 201$
- Solve deterministic PDE using FV, CG with AMG preconditioner
- Coarsest mesh has $4^3 = 64$ DOF, finest mesh has $128^3 \approx 2$M DOF

Numerical results (3)



- Three-dimensional **flow cell** geometry
- Quantity of interest is G^2 (flux through boundary)
- Lognormal random field with Matérn covariance, $\boldsymbol{\lambda = 0.075}$, $\sigma^2 = 1$, $\boldsymbol{\nu = 0.5}$, $\boldsymbol{s = 3000}$
- Solve deterministic PDE using FV, CG with AMG preconditioner
- Coarsest mesh has $4^3 = 64$ DOF, finest mesh has $128^3 \approx 2$M DOF

Conclusions

Closing Thoughts

- MIQMC is a QMC extension of MIMC [Haji-Ali, 2016] or a MIMC extension of MLQMC [Kuo, 2016]
- We show numerical evidence that the MIQMC estimator achieves a convergence rate as good as MIMC, and as good as (or better than) MLQMC for smooth problems
- Cost and error analysis of the estimator must take into account the shape of the index set (can be complicated)
- A general-purpose library is available on
  `https://github.com/PieterjanRobbe/MultilevelEstimators.jl`
  preprint is available on ArXiv
  `http://arxiv.org/abs/1608.03157`
- Future work: adaptive construction of the index set $\mathcal{I}$