Enhanced Multi-Index Monte Carlo by means of Multiple Semi-Coarsened Multigrid for Anisotropic Diffusion Problems

Pieterjan Robbe, Dirk Nuyens, Stefan Vandewalle

19th Copper Mountain Conference On Multigrid Methods

Outline

• Introduction

• Multiple Semi-Coarsened Multigrid

• Unbiased Multi-Index Monte Carlo

• Numerical Example

• Conclusions

PDEs with random coefficients

- consider the elliptic PDE with random coefficients

$$\begin{aligned} -\nabla \cdot (a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) &= f(\mathbf{x}) && \text{with } \mathbf{x} \in D, \omega \in \Omega \\ u(\mathbf{x}, \cdot) &= 0 && \text{on } \partial D \end{aligned}$$

where D is a bounded domain in $\mathbb{R}^d$ and Ω is the sample space of a probability space $(\Omega, \mathcal{F}, P)$

PDEs with random coefficients

- consider the elliptic PDE with random coefficients

$$\begin{aligned} -\nabla \cdot (a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) &= f(\mathbf{x}) && \text{with } \mathbf{x} \in D, \omega \in \Omega \\ u(\mathbf{x}, \cdot) &= 0 && \text{on } \partial D \end{aligned}$$

where D is a bounded domain in $\mathbb{R}^d$ and Ω is the sample space of a probability space $(\Omega, \mathcal{F}, P)$

- take $D = [0, 1]^d$ with $d = 2$ and $f(\mathbf{x}) = 1$

PDEs with random coefficients

- consider the elliptic PDE with random coefficients

$$\begin{aligned} -\nabla \cdot (a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) &= f(\mathbf{x}) && \text{with } \mathbf{x} \in D, \omega \in \Omega \\ u(\mathbf{x}, \cdot) &= 0 && \text{on } \partial D \end{aligned}$$

where D is a bounded domain in $\mathbb{R}^d$ and Ω is the sample space of a probability space $(\Omega, \mathcal{F}, P)$

- take $D = [0, 1]^d$ with $d = 2$ and $f(\mathbf{x}) = 1$
- we are particularly interested in the case where $a(\mathbf{x}, \omega)$ is an **anisotropic** random field

Anisotropic random fields

- stationary non-isotropic Matérn model

$$C(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} (2\sqrt{\nu}r)^\nu K_\nu(2\sqrt{\nu}r)$$

with

$$r = \sqrt{(\mathbf{x} - \mathbf{y})^\top Q (\mathbf{x} - \mathbf{y})} \quad \text{and} \quad Q = R^\top \Lambda R$$

- transformation matrices

$$\Lambda = \begin{pmatrix} \eta\lambda & \\ & \lambda \end{pmatrix}^{-1} \quad \text{and} \quad R = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

Anisotropic random fields

- stationary non-isotropic Matérn model

$$C(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} (2\sqrt{\nu}r)^\nu K_\nu(2\sqrt{\nu}r)$$

with

$$r = \sqrt{(\mathbf{x} - \mathbf{y})^\top Q (\mathbf{x} - \mathbf{y})} \quad \text{and} \quad Q = R^\top \Lambda R$$

- transformation matrices

$$\Lambda = \begin{pmatrix} \eta\lambda & \\ & \lambda \end{pmatrix}^{-1} \quad \text{and} \quad R = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

- variation of smoothness ν , length scale λ , anisotropy parameter η and rotation parameter θ allows for the modeling of many different fields

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/1 \quad \theta = 0^\circ$$

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/2 \quad \theta = 0^\circ$$

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/4 \quad \theta = 0^\circ$$

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/8 \quad \theta = 0^\circ$$

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/8 \quad \theta = 5^\circ$$

Anisotropic random fields

$$\nu = 1/2$$

$$\lambda = 1/4$$

$$\eta = 1/8$$

$$\theta = 10^\circ$$

Anisotropic random fields

$$\nu = 1/2$$

$$\lambda = 1/4$$

$$\eta = 1/8$$

$$\theta = 30^\circ$$

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/8 \quad \theta = 45^\circ$$

Anisotropic random fields

$$\nu = 1/2 \quad \lambda = 1/4 \quad \eta = 1/4 \quad \theta = 45^\circ$$

Anisotropic random fields

$$\nu = 1/2$$

$$\lambda = 1/4$$

$$\eta = 1/2$$

$$\theta = 45^\circ$$

Anisotropic random fields

$$\nu = 1/2$$

$$\lambda = 1/4$$

$$\eta = 1/1$$

$$\theta = 45^\circ$$

Sampling-based methods

- goal is to approximate statistics of a quantity of interest Q
- the expected value of Q can be computed by

$$\mathbb{E}[Q] \approx \mathcal{Q}_N = \frac{1}{N} \sum_{n=1}^N Q^{(n)}$$

Sampling-based methods

- goal is to approximate statistics of a quantity of interest Q
- the expected value of Q can be computed by

$$\mathbb{E}[Q] \approx \mathcal{Q}_N = \frac{1}{N} \sum_{n=1}^N Q^{(n)}$$

- classic Monte Carlo (MC) method requires $N = O(\varepsilon^{-2})$ expensive PDE solves to reach accuracy ε
- several attempts to improve MC: variance reduction techniques, quasi-Monte Carlo, multilevel Monte Carlo, multi-index Monte Carlo...

Two challenges

- **robust** solvers that compute accurate solutions for a wide range of input random fields
 - ⇒ Multiple Semi-coarsened Multigrid (MSG)
- **efficient** sampling-based methods to compute statistics of a quantity of interest
 - ⇒ Unbiased Multi-Index Monte Carlo (uMIMC)

Two challenges

- **robust** solvers that compute accurate solutions for a wide range of input random fields
 - ⇒ Multiple Semi-coarsened Multigrid (MSG)
- **efficient** sampling-based methods to compute statistics of a quantity of interest
 - ⇒ Unbiased Multi-Index Monte Carlo (uMIMC)
- efficiency achieved by
 - learning of distribution of number of samples on all indices (unbiased)
 - recycling of coarse samples from MSG method

Outline

- Introduction
- Multiple Semi-Coarsened Multigrid
- Unbiased Multi-Index Monte Carlo
- Numerical Example
- Conclusions

Failure of standard Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/1$, $\theta = 0$

64×64

Failure of standard Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/2$, $\theta = 0$

64×64

Failure of standard Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/4$, $\theta = 0$

64×64

Failure of standard Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/8$, $\theta = 0$

64×64

Failure of standard Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/16$, $\theta = 0$

64×64

Multiple Semi-Coarsened Multigrid

- standard Multigrid is inadequate, how to improve?
 - use a more powerful smoother: line smoother, ILU [Hackbusch 85]
 - use another coarse grid correction strategy [Mulder, 89], [Oosterlee, 95]

Multiple Semi-Coarsened Multigrid

- standard Multigrid is inadequate, how to improve?
 - use a more powerful smoother: line smoother, ILU [Hackbusch 85]
 - use another coarse grid correction strategy [Mulder, 89], [Oosterlee, 95]

- define a sequence of coarse grids as

$$G^{(p,q)} = \{(i \cdot \Delta x, j \cdot \Delta y) : i = 0, 1, \dots, 2^p, j = 0, 1, \dots, 2^q\}$$

with $\Delta x = 2^{-p}$ and $\Delta y = 2^{-q}$, $p = 1, \dots, \bar{p}$, and $q = 1, \dots, \bar{q}$

Multiple Semi-Coarsened Multigrid

- simple averaging for restriction

$$r^{(p,q)} = \begin{cases} \mathcal{I}_{p+1,q}^{p,q} r^{(p+1,q)} & \text{if } q = \bar{q}, \\ \mathcal{I}_{p,q+1}^{p,q} r^{(p,q+1)} & \text{if } p = \bar{p}, \\ \frac{1}{2} \mathcal{I}_{p+1,q}^{p,q} r^{(p+1,q)} + \frac{1}{2} \mathcal{I}_{p,q+1}^{p,q} r^{(p,q+1)} & \text{otherwise,} \end{cases}$$

Multiple Semi-Coarsened Multigrid

- simple averaging for restriction

$$r^{(p,q)} = \begin{cases} \mathcal{I}_{p+1,q}^{p,q} r^{(p+1,q)} & \text{if } q = \bar{q}, \\ \mathcal{I}_{p,q+1}^{p,q} r^{(p,q+1)} & \text{if } p = \bar{p}, \\ \frac{1}{2} \mathcal{I}_{p+1,q}^{p,q} r^{(p+1,q)} + \frac{1}{2} \mathcal{I}_{p,q+1}^{p,q} r^{(p,q+1)} & \text{otherwise,} \end{cases}$$

- weighted average for interpolation

$$u^{(p,q)} = \begin{cases} \mathcal{I}_{p-1,q}^{p,q} u^{(p-1,q)} & \text{if } q = 1, \\ \mathcal{I}_{p,q-1}^{p,q} u^{(p,q-1)} & \text{if } p = 1, \\ \kappa_x^{p,q} \mathcal{I}_{p-1,q}^{p,q} u^{(p-1,q)} + \kappa_y^{p,q} \mathcal{I}_{p,q-1}^{p,q} u^{(p,q-1)} & \text{otherwise,} \end{cases}$$

see [Naik, Van Rosendale, 93], [de Zeeuw, 90], [Dendy, 82], [Oosterlee, 95]

Success of Multiple Semi-Coarsened Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/1$, $\theta = 0$

64×64

Success of Multiple Semi-Coarsened Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/2$, $\theta = 0$

64×64

Success of Multiple Semi-Coarsened Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/4$, $\theta = 0$

64×64

Success of Multiple Semi-Coarsened Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/8$, $\theta = 0$

64×64

Success of Multiple Semi-Coarsened Multigrid

- W(2,2) cycles for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/16$, $\theta = 0$

64×64

Full Multigrid

- use solution on coarse grids as good initial guess for iterative method on the fine grid (**nested iteration**) see [Hackbusch, 85], [Brandt, 77]
- typically uses higher-order (cubic) interpolation

Full Multigrid

- use solution on coarse grids as good initial guess for iterative method on the fine grid (**nested iteration**) see [Hackbusch, 85], [Brandt, 77]
- typically uses higher-order (cubic) interpolation
- Full Multiple Semi-Coarsened Multigrid (FMSG) on grid $G^{(p,q)}$ yields solutions on all coarser grids $G^{(p',q')}$ with $p' \leq p$ and $q' \leq q$

Outline

- Introduction
- Multiple Semi-Coarsened Multigrid
- Unbiased Multi-Index Monte Carlo
- Numerical Example
- Conclusions

Multi-Index Monte Carlo

- in multi-index methods, we use a hierarchy of approximations for the quantity of interest Q_ℓ , with $\ell = (\ell_1, \ell_2, \dots, \ell_d)$ a d -tuple
- we assume $Q_\ell \rightarrow Q$ as $\ell \rightarrow \infty$ componentwise

Multi-Index Monte Carlo

- in multi-index methods, we use a hierarchy of approximations for the quantity of interest Q_ℓ , with $\ell = (\ell_1, \ell_2, \dots, \ell_d)$ a d -tuple
- we assume $Q_\ell \rightarrow Q$ as $\ell \rightarrow \infty$ componentwise
- multi-index differences are computed as

$$\Delta Q_\ell := \left(\bigotimes_{i=1}^d \Delta_i \right) Q_\ell \quad \text{with} \quad \Delta_i Q_\ell = \begin{cases} Q_\ell - Q_{\ell - \mathbf{e}_i} & \text{if } \ell_i > 0 \\ Q_\ell & \text{if } \ell_i = 0 \end{cases},$$

where $\mathbf{e}_i$ is the unit vector in direction i

Multi-Index Monte Carlo

- in multi-index methods, we use a hierarchy of approximations for the quantity of interest Q_ℓ , with $\ell = (\ell_1, \ell_2, \dots, \ell_d)$ a d -tuple
- we assume $Q_\ell \rightarrow Q$ as $\ell \rightarrow \infty$ componentwise
- multi-index differences are computed as

$$\Delta Q_\ell := \left(\bigotimes_{i=1}^d \Delta_i \right) Q_\ell \quad \text{with} \quad \Delta_i Q_\ell = \begin{cases} Q_\ell - Q_{\ell - \mathbf{e}_i} & \text{if } \ell_i > 0 \\ Q_\ell & \text{if } \ell_i = 0 \end{cases},$$

where $\mathbf{e}_i$ is the unit vector in direction i

- for example, with $d = 2$ and $\ell = (1, 2)$

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

Multi-Index Monte Carlo

- in general, taking a sample of ΔQ_ℓ requires 2^d PDE solves

Multi-Index Monte Carlo

- Multi-Index Monte Carlo estimator from [Haji-Ali, Nobile, Tempone, 16], is just the sum of MC estimators for the multi-index differences

$$\mathbb{E}[Q] = \sum_{\ell \in \mathbb{N}_0^d} \mathbb{E}[\Delta Q_\ell] \quad \Rightarrow \quad \mathcal{Q} = \sum_{\ell \in \mathcal{I}_d(L)} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \Delta Q_\ell^{(n)}$$

Multi-Index Monte Carlo

- Multi-Index Monte Carlo estimator from [Haji-Ali, Nobile, Tempone, 16], is just the sum of MC estimators for the multi-index differences

$$\mathbb{E}[Q] = \sum_{\ell \in \mathbb{N}_0^d} \mathbb{E}[\Delta Q_\ell] \quad \Rightarrow \quad \mathcal{Q} = \sum_{\ell \in \mathcal{I}_d(L)} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \Delta Q_\ell^{(n)}$$

- difference ΔQ_ℓ must be based on same random sample to ensure strong positive correlation, such that

$$\ell \rightarrow \infty \text{ componentwise} \quad \Rightarrow \quad Q_\ell \rightarrow Q \quad \Rightarrow \quad \Delta Q_\ell \rightarrow 0$$

- variance of difference, and thus the number of samples required, will decrease rapidly with ℓ

Multi-Index Monte Carlo

- Multi-Index Monte Carlo estimator from [Haji-Ali, Nobile, Tempone, 16], is just the sum of MC estimators for the multi-index differences

$$\mathbb{E}[Q] = \sum_{\ell \in \mathbb{N}_0^d} \mathbb{E}[\Delta Q_\ell] \quad \Rightarrow \quad \mathcal{Q} = \sum_{\ell \in \mathcal{I}_d(L)} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \Delta Q_\ell^{(n)}$$

- difference ΔQ_ℓ must be based on same random sample to ensure strong positive correlation, such that

$$\ell \rightarrow \infty \text{ componentwise} \quad \Rightarrow \quad Q_\ell \rightarrow Q \quad \Rightarrow \quad \Delta Q_\ell \rightarrow 0$$

- variance of difference, and thus the number of samples required, will decrease rapidly with ℓ
- $d = 1$ is known as Multilevel Monte Carlo [Giles, 08]

Variance bias trade-off

$$\mathcal{I}_d(L) = \{\ell : |\ell| \leq L\}$$

$$\mathcal{I}_d(L) = \{\ell : \Pi(\ell_j + 1) \leq L + 1\}$$

$$\mathcal{I}_d(L) = \{\ell : \max(1, \ell_j) \leq L\}$$

- accuracy is monitored by controlling the mean square error (MSE)

$$\text{MSE}(\mathcal{Q}) = \mathbb{V}[\mathcal{Q}] + (\mathbb{E}[\mathcal{Q}] - Q)^2$$

Variance bias trade-off

$$\mathcal{I}_d(L) = \{\ell : |\ell| \leq L\}$$

$$\mathcal{I}_d(L) = \{\ell : \Pi(\ell_j + 1) \leq L + 1\}$$

$$\mathcal{I}_d(L) = \{\ell : \max(1, \ell_j) \leq L\}$$

- accuracy is monitored by controlling the mean square error (MSE)

$$\text{MSE}(\mathcal{Q}) = \mathbb{V}[\mathcal{Q}] + (\mathbb{E}[\mathcal{Q}] - Q)^2$$

- bias is computed by heuristic

$$|\mathbb{E}[\mathcal{Q}] - Q| \approx \left| \sum_{\ell \in \partial \mathcal{I}_d} \mathbb{E}[\Delta Q_\ell] \right|$$

where $\partial \mathcal{I}_d$ is the boundary of index set [Haji-Ali, Nobile, Tempone, 16]

Unbiased Multi-Index Monte Carlo

- the unbiased MIMC estimator is

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq \mathbf{L}^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}$$

where

- $\mathbf{L}$ is a d -variate discrete random variable with pmf $\pi_{\mathbf{L}}$, and
- $\sigma(\ell) = \mathbb{P}[\ell \leq \mathbf{L}]$ is the probability that $\mathbf{L}$ is at least ℓ

Unbiased Multi-Index Monte Carlo

- the unbiased MIMC estimator is

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_\ell^{(n)}$$

where

- L is a d -variate discrete random variable with pmf π_L , and
- $\sigma(\ell) = \mathbb{P}[\ell \leq L]$ is the probability that L is at least ℓ
- it can be shown that $\mathbb{E}[Q^u] = \mathbb{E}[Q]$ (benefit I)

Unbiased Multi-Index Monte Carlo

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}$$

Unbiased Multi-Index Monte Carlo

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_\ell^{(n)}$$

Unbiased Multi-Index Monte Carlo

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_\ell^{(n)}$$

Unbiased Multi-Index Monte Carlo

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}$$

Unbiased Multi-Index Monte Carlo

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_\ell^{(n)}$$

Unbiased Multi-Index Monte Carlo

$$Q^u = \frac{1}{N} \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}$$

Sample reuse

- the estimator is just the average of N terms

$$Y^{(n)} = \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}, \quad n = 1, \dots, N$$

Sample reuse

- the estimator is just the average of N terms

$$Y^{(n)} = \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}, \quad n = 1, \dots, N$$

- every $Y^{(n)}$ can be computed from just a single PDE solve using FMSG (**benefit II**)
- the estimator that reuses samples is denoted by Q^r

Sample reuse

- the estimator is just the average of N terms

$$Y^{(n)} = \sum_{\ell \leq L^{(n)}} \frac{1}{\sigma(\ell)} \Delta Q_{\ell}^{(n)}, \quad n = 1, \dots, N$$

- every $Y^{(n)}$ can be computed from just a single PDE solve using FMSG (**benefit II**)
- the estimator that reuses samples is denoted by Q^r
- can use the sample variance of $Y^{(n)}$ to compute the variance (and thus the MSE) $\mathbb{V}[Q^r]$ if L is i.i.d

$$\mathbb{V}[Q^r] \approx \frac{1}{N(N-1)} \sum_{n=1}^N \left(Y^{(n)} - \frac{1}{N} \sum_{n=1}^N Y^{(n)} \right)^2$$

Choice of probability mass function π_L

- rewrite the estimator as

$$Q^r = \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{N\sigma(\ell)} \Delta Q_\ell^{(n)} = \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{N'_\ell} \Delta Q_\ell^{(n)}$$

where N'_ℓ is now a discrete density of sample sizes

Choice of probability mass function π_L

- rewrite the estimator as

$$Q^r = \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{N\sigma(\ell)} \Delta Q_\ell^{(n)} = \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{N'_\ell} \Delta Q_\ell^{(n)}$$

where N'_ℓ is now a discrete density of sample sizes

- a standard MIMC result is

$$N'_\ell \approx \sqrt{\frac{V_\ell}{C_\ell}}$$

with $V_\ell := \mathbb{V}[\Delta Q_\ell]$ and C_ℓ the cost to compute one sample

Choice of probability mass function π_L

- rewrite the estimator as

$$Q^r = \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{N\sigma(\ell)} \Delta Q_\ell^{(n)} = \sum_{n=1}^N \sum_{\ell \leq L^{(n)}} \frac{1}{N'_\ell} \Delta Q_\ell^{(n)}$$

where N'_ℓ is now a discrete density of sample sizes

- a standard MIMC result is

$$N'_\ell \approx \sqrt{\frac{V_\ell}{C_\ell}}$$

with $V_\ell := \mathbb{V}[\Delta Q_\ell]$ and C_ℓ the cost to compute one sample

- 2 possibilities
 - on-the-fly computed estimates for V_ℓ and C_ℓ to yield empirical pmf π_L
 - propose a model for V_ℓ and C_ℓ and compute π_L a priori

Assumptions

- standard Multi-Index Monte Carlo assumptions

$$E_\ell := |\mathbb{E}[\Delta Q_\ell]| \leq c_E \prod_{j=1}^d 2^{-\alpha_j \ell_j}$$

$$V_\ell := \mathbb{V}[\Delta Q_\ell] \leq c_V \prod_{j=1}^d 2^{-\beta_j \ell_j}$$

$$C_\ell := \text{cost}(\Delta Q_\ell) \leq c_C \prod_{j=1}^d 2^{\gamma_j \ell_j}$$

see [Haji-Ali, Nobile, Tempone, 16]

Theoretical cost reduction

- with these assumptions, we can proof that the optimal choice for π_L is a multivariate geometric distribution

Theoretical cost reduction

- with these assumptions, we can proof that the optimal choice for π_L is a multivariate geometric distribution

Proposition

The cost reduction factor of the MSG-MIMC algorithm with uniform refinements is given by

$$\frac{\text{cost}(Q^r)}{\text{cost}(Q^u)} = 1 - \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} (-1)^{|u|+1} \prod_{j \in u} 2^{-\frac{1}{2}(1+\beta_j)}$$

where β_j is the rate of the decrease in variance in direction j .

Outline

- Introduction
- Multiple Semi-Coarsened Multigrid
- Unbiased Multi-Index Monte Carlo
- Numerical Example
- Conclusions

Numerical example

- model problem with **anisotropic** diffusion coefficient
- $a(\mathbf{x}, \omega)$ is lognormal random field with Matérn covariance
- put $\theta = \mathcal{U}[-30^\circ, 30^\circ]$ and $\eta = \mathcal{U}[1/16, 1]$ as **hyperparameters**

Numerical example

- model problem with **anisotropic** diffusion coefficient
- $a(\mathbf{x}, \omega)$ is lognormal random field with Matérn covariance
- put $\theta = \mathcal{U}[-30^\circ, 30^\circ]$ and $\eta = \mathcal{U}[1/16, 1]$ as **hyperparameters**
- coarsest mesh has 8×8 unknowns, finest mesh has 128×128
- FMG with an **adaptive** number of $W(2,2)$ -cycles
- quantity of interest is point evaluation of $u(\mathbf{x})$
- compare the number of equivalent PDE solves on the finest level for different accuracies (root mean square error) using MIMC with (Q^r) and without (Q^u) sample recycling

Numerical example

- standard cost for $\nu = 1/2$, $\lambda = 1/4$, $\eta \in [1/16, 1]$, $\theta \in [-30^\circ, 30^\circ]$

Outline

- Introduction
- Multiple Semi-Coarsened Multigrid
- Unbiased Multi-Index Monte Carlo
- Numerical Example
- Conclusions

Conclusions

- focus on elliptic PDE with random anisotropic coefficients
- we have shown that this problem requires both a **robust** deterministic solver and an **efficient** sampling method to compute statistics of a quantity of interest
- efficiency is obtained by unbiased estimation (**benefit I**)
[Rhee, Glynn, 15], [Detommaso, Dodwell, Scheichl, 18] and sample recycling (**benefit II**)

Conclusions

- focus on elliptic PDE with random anisotropic coefficients
- we have shown that this problem requires both a **robust** deterministic solver and an **efficient** sampling method to compute statistics of a quantity of interest
- efficiency is obtained by unbiased estimation (**benefit I**)
[Rhee, Glynn, 15], [Detommaso, Dodwell, Scheichl, 18] and sample recycling (**benefit II**)
- how tight is the theoretically computed reduction factor?
- can we gain by using QMC?

Conclusions

- focus on elliptic PDE with random anisotropic coefficients
- we have shown that this problem requires both a **robust** deterministic solver and an **efficient** sampling method to compute statistics of a quantity of interest
- efficiency is obtained by unbiased estimation (**benefit I**)
[Rhee, Glynn, 15], [Detommaso, Dodwell, Scheichl, 18] and sample recycling (**benefit II**)
- how tight is the theoretically computed reduction factor?
- can we gain by using QMC?

Thank you for your attention!

Multiple Semi-Coarsened Multigrid

- convergence factor for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/16$, $\theta = 5$

Multiple Semi-Coarsened Multigrid

- convergence factor for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/16$, $\theta = 10$

Multiple Semi-Coarsened Multigrid

- convergence factor for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/16$, $\theta = 20$

Multiple Semi-Coarsened Multigrid

- convergence factor for $\nu = 1/2$, $\lambda = 1/4$, $\eta = 1/16$, $\theta = 30$

Cost analysis

- suppose N_ℓ samples are required on each index $\ell \in \mathcal{I}_d$
- when recycling samples, only

$$N'_\ell = N_\ell + \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} (-1)^{|u|} N_{\ell + \mathbf{e}_u}$$

samples remain to be taken on index ℓ

Cost analysis

- suppose N_ℓ samples are required on each index $\ell \in \mathcal{I}_d$
- when recycling samples, only

$$N'_\ell = N_\ell + \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} (-1)^{|u|} N_{\ell + \mathbf{e}_u}$$

samples remain to be taken on index ℓ

- using the relation

$$N_{\ell + \mathbf{e}_u} = \left(\prod_{j \in u} 2^{-\frac{1}{2}(\gamma_j + \beta_j)} \right) N_\ell$$

Cost analysis

- we find

$$\text{cost}(\mathcal{Q}^r) = \sum_{\ell \in \mathcal{I}_d} N'_\ell C_\ell$$

Cost analysis

- we find

$$\begin{aligned}\text{cost}(Q^r) &= \sum_{\ell \in \mathcal{I}_d} N'_\ell C_\ell \\ &= \left(1 + \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} \left((-1)^{|u|} \prod_{j \in u} 2^{-\frac{1}{2}(\gamma_j + \beta_j)} \right) \right) \sum_{\ell \in \mathcal{I}_d} N_\ell C_\ell \\ &= \left(1 + \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} \left((-1)^{|u|} \prod_{j \in u} 2^{-\frac{1}{2}(\gamma_j + \beta_j)} \right) \right) \text{cost}(Q^u)\end{aligned}$$

Cost analysis

- we find

$$\begin{aligned}\text{cost}(Q^r) &= \sum_{\ell \in \mathcal{I}_d} N'_\ell C_\ell \\ &= \left(1 + \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} \left((-1)^{|u|} \prod_{j \in u} 2^{-\frac{1}{2}(\gamma_j + \beta_j)} \right) \right) \sum_{\ell \in \mathcal{I}_d} N_\ell C_\ell \\ &= \left(1 + \sum_{\substack{u \subseteq \{1, \dots, d\} \\ u \neq \emptyset}} \left((-1)^{|u|} \prod_{j \in u} 2^{-\frac{1}{2}(\gamma_j + \beta_j)} \right) \right) \text{cost}(Q^u)\end{aligned}$$

- with an FMG solver, we have $\gamma_j = 1$