Multilevel Monte Carlo for a Stochastic Optimal Control Problem

Q. Sun[†,‡], with J. Ming[†] and Q. Du[†,§]

BCAM Summer School on UQAP

July 5, 2016

[†] Division of Applied and Computational Mathematics, Beijing Computational Science Research Center
[‡] School of Mathematical Science, University of Science and Technology of China
[§] Department of Applied Physics and Applied Mathematics and Institute for Data Science, Columbia University

Stochastic Optimal Control Problem

- Stochastic elliptic problem
- Stochastic optimal control problem

Multilevel Monte Carlo Method

- Introduction to multilevel Monte Carlo method
- Analysis to multilevel Monte Carlo method

Numerical Experiment

- 2D stochastic optimal control problem

Concluding Remarks

Stochastic Optimal Control Problem

- Stochastic elliptic problem
- Stochastic optimal control problem

Multilevel Monte Carlo Method

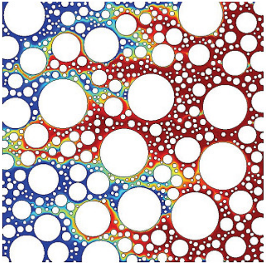
- Introduction to multilevel Monte Carlo method
- Analysis to multilevel Monte Carlo method

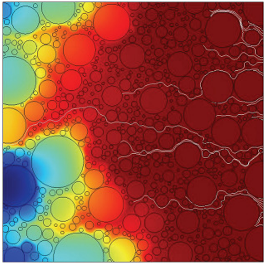
Numerical Experiment

- 2D stochastic optimal control problem

Concluding Remarks

Stochastic Elliptic Problems





Strong form: given $f(x) \in L^2(D)$, find the stochastic function $u(x,\omega) : \bar{D} \times \Omega \to \mathbb{R}$ such that for a.e. $\omega \in \Omega$

$$\begin{cases} -\nabla \cdot (a(x,\omega)\nabla u(x,\omega)) = f(x) & \text{in } D, \\ u(x,\omega) = 0 & \text{on } \partial D, \end{cases} \tag{1}$$

where the coefficient $a(x,\omega)$ could be, e.g., the permeability or conductivity field involving uncertainty.

§ Alexandra Foley, Can you use heat transfer to predict migration of contaminants?, COMSOL BLOG, 2013.

Existence and Uniqueness

Weak form: given $f(x) \in L^2(D)$, seek $u(x,\omega) \in \mathcal{H}_0^1(D)$ such that

$$\mathbb{E}\left[\int_D a\nabla u \cdot \nabla v dx\right] = \mathbb{E}\left[\int_D f v dx\right] \quad \forall v \in \mathcal{H}_0^1(D), \tag{2}$$

where $\mathcal{H}_0^1(D) = L^2(\Omega; H_0^1(D))$ is the stochastic Sobolev space.

If the stochastic coefficient $a(x,\omega)$ satisfies the integrability condition

$$\frac{1}{a^{\min}(\omega)} \in L^2(\Omega) \quad \text{and} \quad a^{\min}(\omega) > 0 \quad \text{for a.e. } \omega \in \Omega,$$

there exists a unique weak solution of (2), and the following estimate hold

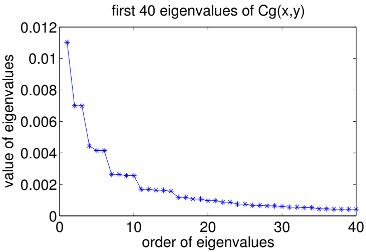
$$\|u\|_{\mathcal{H}_0^1(D)} \le C_D \left\| \frac{1}{a^{\min}(\omega)} \right\|_{L^2(\Omega)} \|f\|_{L^2(D)},$$

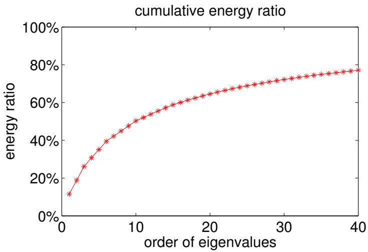
where $a^{\min}(\omega) = \min_{x \in \bar{D}} a(x,\omega)$ and C_D is the constant of Poincaré's inequality.

Representations of Random Fields

Log-normal coefficient: $a(x,\omega) = e^{g(x,\omega)}$ satisfies the integrability condition[§], where $g(x,\cdot)$ is a zero mean Gaussian random variable with $\text{Cov}_g(x,y) = \sigma^2 e^{-\frac{\|x-y\|}{\lambda}}$.

Truncated Karhunen-Loève expansion: $g(x,\omega) = \sum_{m=1}^{M} \sqrt{\lambda_m}\phi_m(x)X_m(\omega)$, where the truncation number M is proportional to the smoothness of $\text{Cov}_g(x,y)$.





$$\text{Cov}_g(\mathbf{x},\mathbf{y}) = 0.1 \times e^{-\frac{|x_1-y_1|+|x_2-y_2|}{0.2}} \text{ on } D = [0,1]^2$$

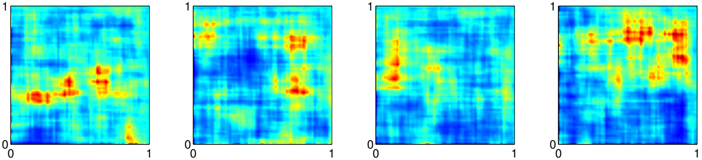
[§] J. Charrier, Strong and weak error estimates for elliptic PDE with random coefficients, SIAM J.N.A., 2012.

Representations of Random Fields - Continued

Grid-based representation[§]: given $\{x_i \in \bar{D}\}_{i=1}^N$ and positive-definite covariance matrix $\mathcal{G}_g = (\mathrm{Cov}_g(x_i, x_i))_{i,j=1}^N = \Theta\Theta^T$, then

$$(a(x_1,\omega),\ldots,a(x_N,\omega))^T = e^{\Theta(X_1(\omega),\ldots,X_N(\omega))^T}, \ X_i(\omega) \overset{iid}{\sim} N(0,1).$$

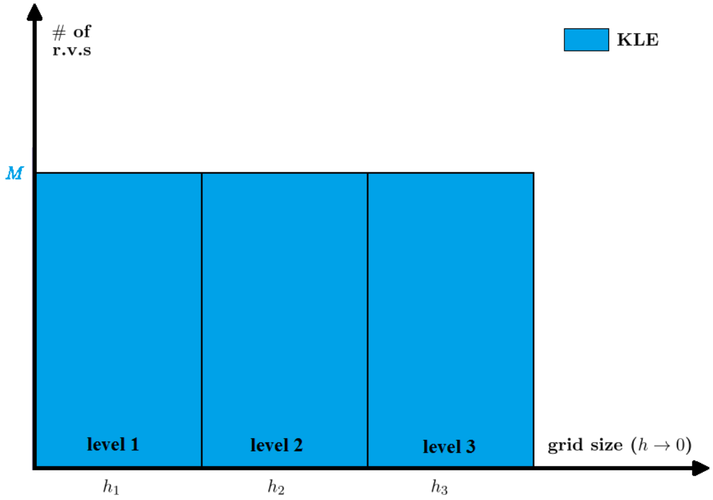
- GBR represents $a(x,\omega)$ exactly on gird nodes while KLE with truncation error.
- The dimension of r.v.s in GBR varies w.r.t. mesh size while KLE stays the same.
- KLE is cheaper than GBR on fine meshes but more expensive on coarse meshes.



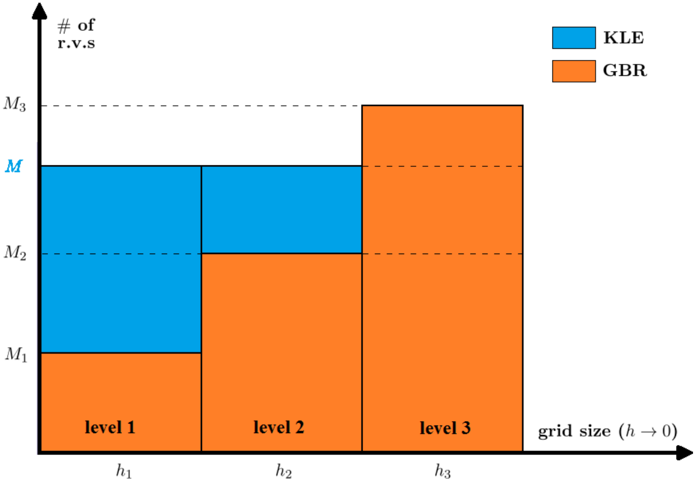
$$\mathrm{Cov}_g(\mathbf{x},\mathbf{y}) = 0.1 \times e^{-\frac{|x_1-y_1|+|x_2-y_2|}{0.2}} \text{ on } D = [0,1]^2$$

[§] I.H. Sloan et al, Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications, JCP, 2011.

Cost of KLE for one Realization



Cost of GBR for one Realization



Stochastic Optimal Control Problem

- Stochastic elliptic problem
- Stochastic optimal control problem

Multilevel Monte Carlo Method

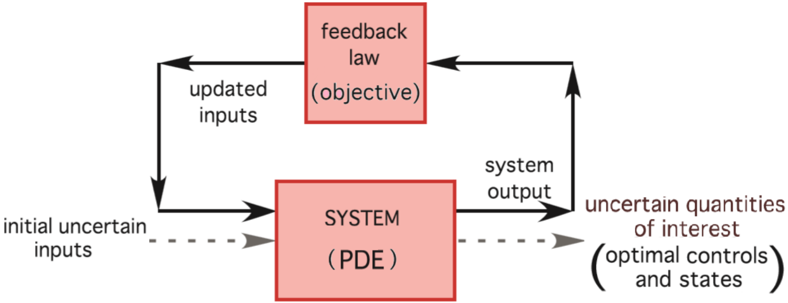
- Introduction to multilevel Monte Carlo method
- Analysis to multilevel Monte Carlo method

Numerical Experiment

- 2D stochastic optimal control problem

Concluding Remarks

Feedback control problems



Structure of SOCPs involves: state variables, control variables, an objective or cost functional and constraints.

SOCPs: find state and control variables (u, f) that minimize(or maximize) the objective functional $\mathcal{J}(u, f)$ subject to the constraints $G(u, f) = 0$ are satisfied.

§M. D. Gunzburger et al. An algorithmic introduction to numerical methods for PDEs with random inputs, BCAM, 2013.

Optimal Control Problem Governed by Elliptic SPDEs

Given a deterministic target function $U(x) \in L^2(D)$, find the optimal robust control $f^{\text{opt}}(x) \in \mathcal{A} \subset L^2(D)$ such that the objective functional

$$\mathcal{J}_\beta(u, f) = \mathbb{E}\left[\frac{1}{2}\int_D |u(x,\omega) - U(x)|^2 dx\right] + \frac{\beta}{2}\int_D |f(x)|^2 dx \tag{3}$$

subject to the elliptic SPDE

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

is minimized. Our quantity of interest (QoI) is the expected value $\mathbb{E}[u^{\text{opt}}]$.

Let $\mathcal{A}$ be a closed convex subset, there exists a unique optimal solution pair $(u^{\text{opt}}, f^{\text{opt}}) \in \mathcal{H}_0^1(D) \times \mathcal{A}$ of the stochastic optimal control problem

$$\text{Minimize (3) subject to state equation (1)}^{\S}. \tag{4}$$

§ J. Lee, Analyses and Finite Element Approximation of Stochastic Optimal Control Problems Constrained by Stochastic Elliptic PDEs, Iowa State University, 2008.

Surrogate Model

For numerical implementation, we adopt following surrogate model with Monte Carlo approximation of problem (4) to realize a gradient-based optimization algorithm later:

$$\text{minimize} \quad \widehat{\mathcal{J}}_\beta(u,f) = \frac{1}{2N}\sum_{i=1}^{N}\int_D |u(x,\omega_i) - U(x)|^2 dx + \frac{\beta}{2}\int_D |f(x)|^2 dx \tag{5}$$

$$\text{subject to} \quad \begin{cases} -\nabla\cdot(a(x,\omega_i)\nabla u(x,\omega_i)) = f(x) & \text{in } D \\ u(x,\omega_i) = 0 & \text{on } \partial D \end{cases} \quad \text{for } i = 1,\dots,N, \tag{6}$$

where our QoI $\mathbb{E}[u^{\text{opt}}]$ is approximated by $E_N[u^{\text{opt}}] = \frac{1}{N}\sum_{i=1}^{N} u^{\text{opt}}(x,\omega_i)$.

For any $N \in \mathbb{Z}_+$ and for $u^{\text{opt}} \in \mathcal{H}_0^1(D)$ satisfying $\frac{\partial \mathbb{E}[u^{\text{opt}}]}{\partial x} = \mathbb{E}[\frac{\partial u^{\text{opt}}}{\partial x}]$, there holds

$$\|E_N[u^{\text{opt}}] - \mathbb{E}[u^{\text{opt}}]\|_{\mathcal{H}_0^1(D)} \leq \sqrt{\frac{C_D}{N}} \left\| \frac{1}{a^{\min}(\omega)} \right\|_{L^2(\Omega)} \|f^{\text{opt}}\|_{L^2(D)}.$$

Euler-Lagrange Equations

Introducing the Lagrangian functional of problem (5-6),

$$\mathcal{L}(u, f, \xi) = \widehat{\mathcal{J}}_\beta(u, f) - \frac{1}{N}\sum_{i=1}^{N} a_i(u_i, \xi_i) + \frac{1}{N}\sum_{i=1}^{N}(f, \xi_i)$$

that defined for arbitrary $(u_i, f, \xi_i) \in H_0^1(D) \times \mathcal{A} \times H_0^1(D)$, $i = 1, \ldots, N$, where

$$(f, \xi_i) = \int_D f(x)\xi(x, \omega_i)dx, \quad a_i(u_i, \xi_i) = \int_D a(x, \omega_i)\nabla u(x, \omega_i) \cdot \nabla \xi(x, \omega_i)dx,$$

the equivalent system of problem (5-6) is given by: find $(u_i^{\text{opt}}, f^{\text{opt}}) \in H_0^1(D) \otimes L^2(D)$, $i = 1, \ldots, N$, such that

$$\begin{cases} a_i(u_i^{\text{opt}}, v_i) = (f^{\text{opt}}, v_i) & \forall v_i \in H_0^1(D) \text{ for } i = 1, \ldots, N, \quad \text{(state eqns)} \\ a_i(\xi_i^{\text{opt}}, \zeta_i) = (u_i^{\text{opt}} - U, \zeta_i) & \forall \zeta_i \in H_0^1(D) \text{ for } i = 1, \ldots, N, \quad \text{(adjoint eqns)} \\ (\beta f^{\text{opt}} + \frac{1}{N}\sum_{i=1}^{N} \xi_i^{\text{opt}}, z) = 0 & \forall z \in \mathcal{A}, \end{cases}$$

where our QoI is approximated by $E_N[u^{\text{opt}}] = \frac{1}{N}\sum_{i=1}^{N} u_i^{\text{opt}}$.

A Descent Direction for the Control Variate

Since the optimal solution of problem (5-6) satisfies

$$\mathcal{L}(u^{\text{opt}}, f^{\text{opt}}, \xi^{\text{opt}}) \leq \mathcal{L}(u^{\text{opt}}, g, \xi^{\text{opt}}) \quad \forall g \in \mathcal{A}^{\S},$$

given f_{old}, the descent direction could be constructed via the convex combination

$$f_{\text{new}} = \epsilon g + (1-\epsilon) f_{\text{old}} \;\text{ for some } \epsilon \in (0,1) \text{ and } g \in \mathcal{A},$$

such that $\mathcal{L}(u_{\text{old}}, f_{\text{new}}, \xi_{\text{old}}) \leq \mathcal{L}(u_{\text{old}}, f_{\text{old}}, \xi_{\text{old}})$. By letting $\epsilon \to 0^+$, we have

$$\int_D (g - f_{\text{old}})(\beta f_{\text{old}} + \frac{1}{N}\sum_{i=1}^{N} \xi_{\text{old}}^{(i)}) dx \leq 0.$$

Then, by choosing $g - f_{\text{old}} = -t(\beta f_{\text{old}} + \frac{1}{N}\sum_{i=1}^{N} \xi_{\text{old}}^{(i)})$, we derive

$$f_{\text{new}} = f_{\text{old}} - s(\beta f_{\text{old}} + \frac{1}{N}\sum_{i=1}^{N} \xi_{\text{old}}^{(i)}),$$

where $s = \epsilon t > 0$ is chosen to be a small stepsize.

§ M. D. Gunzburger et al, Finite-dimensional approximation of a class of constrained nonlinear optimal control problems, SIAM J.C.O., 1996.

Stochastic Optimal Control Problem

- Stochastic elliptic problem
- Stochastic optimal control problem

Multilevel Monte Carlo Method

- Introduction to multilevel Monte Carlo method
- Analysis to multilevel Monte Carlo method

Numerical Experiment

- 2D stochastic optimal control problem

Concluding Remarks

Multi-level Monte Carlo Method

Notations: suppose our QoI is $\mathbb{E}[Q]$ where $Q(x,\omega)$ solves a given problem, let $Q_\ell(x,\omega)$ denote the FE approximation of $Q(x,\omega)$ on mesh $\mathcal{T}_\ell$, and $Q_\ell^{(i)}(x)$ the corresponding realization $Q_\ell(x,\omega^{(i,\ell)})$ for $i=1,\dots,N_\ell$ and $\ell=0,\dots,L$.

Single-level Monte Carlo: given the mesh $\mathcal{T}_L$ and r.v.s $\{\omega^{(i,L)}\}_{i=1}^{N_L}$, SLMC gives

$$\widehat{Q}_L^{SL} = \frac{1}{N_L}\sum_{i=1}^{N_L} Q_L^{(i)} \text{ with } \|\text{MSE}(\widehat{Q}_L^{SL})\| \le \frac{\|\mathbb{V}[Q_L]\|}{N_L} + \|(\mathbb{E}[Q_L]-\mathbb{E}[Q])^2\|.$$

Multi-level Monte Carlo: given meshes $\{\mathcal{T}_\ell\}_{\ell=0}^{L}$ and r.v.s $\{\omega^{(i,\ell)}\}_{i=1,\ell=0}^{N_\ell,L}$, MLMC gives

$$\widehat{Q}_L^{ML} = \frac{1}{N_0}\sum_{i=1}^{N_0} Q_0^{(i)} + \sum_{\ell=1}^{L}\frac{1}{N_\ell}\sum_{i=1}^{N_\ell}(Q_\ell^{(i)} - Q_{\ell-1}^{(i)}),$$

$$\text{with } \|\text{MSE}(\widehat{Q}_L^{ML})\| \le \frac{\|\mathbb{V}[Q_0]\|}{N_0} + \sum_{\ell=1}^{L}\frac{\|\mathbb{V}[Q_\ell - Q_{\ell-1}]\|}{N_\ell} + \|(\mathbb{E}[Q_L]-\mathbb{E}[Q])^2\|.$$

Multi-level Monte Carlo Method - Continued

Instead of generating all realizations on one fixed mesh, MLMC uses hierarchical meshes

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

Alternatively, one could use quasi-Monte Carlo, collocation rule, etc., for the numerical approximation of expectations.

Notations: let $Q_{-1} := 0$, we define $\mathcal{V}_\ell := \|\mathbb{V}\left[Q_\ell - Q_{\ell-1}\right]\|$ and $\mathcal{C}_\ell$ the computational cost to obtain a single realization $Q_\ell^{(i)} - Q_{\ell-1}^{(i)}$ for $\ell = 0, \ldots, L$. Moreover, we assume the discrete error $\|(\mathbb{E}\left[Q_L\right] - \mathbb{E}\left[Q\right])^2\| = \mathcal{O}(h_L^\alpha)$ for some constant $\alpha > 0$.

Total computational cost of MLMC:

$$T_c = N_0\mathcal{C}_0 + \cdots + N_L\mathcal{C}_L.$$

Computational accuracy of MLMC:

$$e_L = N_0^{-1}\mathcal{V}_0 + \cdots + N_L^{-1}\mathcal{V}_L + \mathcal{O}(h_L^\alpha).$$

Sample Size Formulae

Sample size formulae are derived from two optimization perspectives: minimizing the computational error/cost with given computational complexity/accuracy:

$$\text{Opt 1}: \begin{cases} \text{Minimize } e_L = N_0^{-1}\mathcal{V}_0 + N_1^{-1}\mathcal{V}_1 + \ldots + N_L^{-1}\mathcal{V}_L + \mathcal{O}(h_L^\alpha), \\ \text{subject to } T_c^* = N_0\mathcal{C}_0 + N_1\mathcal{C}_1 + \ldots + N_L\mathcal{C}_L. \end{cases} \tag{7}$$

By solving this integer optimization problem, we derive

$$N_0 = \sqrt{\frac{\mathcal{V}_0}{\mathcal{C}_0}} \frac{T_c^*}{\sqrt{\mathcal{C}_0\mathcal{V}_0} + \ldots + \sqrt{\mathcal{C}_L\mathcal{V}_L}} \text{ and } N_i = N_j\sqrt{\frac{\mathcal{C}_j}{\mathcal{C}_i} \cdot \frac{\mathcal{V}_i}{\mathcal{V}_j}} \text{ for } 0 \le i \ne j \le L.$$

$$\text{Opt 2}: \begin{cases} \text{Minimize } T_c = N_0\mathcal{C}_0 + N_1\mathcal{C}_1 + \ldots + N_L\mathcal{C}_L, \\ \text{subject to } e_L^* - \mathcal{O}(h_L^\alpha) = N_0^{-1}\mathcal{V}_0 + N_1^{-1}\mathcal{V}_1 + \ldots + N_L^{-1}\mathcal{V}_L. \end{cases} \tag{8}$$

Similarly, we derive

$$N_0 = \sqrt{\frac{\mathcal{V}_0}{\mathcal{C}_0}} \left(\frac{\sqrt{\mathcal{V}_0\mathcal{C}_0} + \ldots + \sqrt{\mathcal{V}_L\mathcal{C}_L}}{e_L^* - \mathcal{O}(h_L^\alpha)} \right) \text{ and } N_i = N_j\sqrt{\frac{\mathcal{C}_j}{\mathcal{C}_i} \cdot \frac{\mathcal{V}_i}{\mathcal{V}_j}} \text{ for } 0 \le i \ne j \le L.$$

MLMC Analyses

WLOG, let the sequence of triangulations $\{\mathcal{T}_\ell\}_{\ell=0}^{L}$ satisfies the subdivision strategy which leads to $h_\ell = 2^{-1}h_{\ell-1}$ for $\ell = 1, \ldots, L$, suppose the following assumptions[§]

A1. $\mathcal{V}_\ell = \begin{cases} \|\mathbb{V}[Q_0]\| = c_0, \\ \|\mathbb{V}[Q_\ell - Q_{\ell-1}]\| = \mathcal{O}(h_\ell^\beta) \ \text{ for } \ \ell = 1, \ldots, L, \end{cases}$

A2. $\mathcal{C}_\ell = \mathcal{O}(h_\ell^{-\gamma}) \ \text{ for } \ \ell = 0, \ldots, L.$

hold for some positive constants β and γ, then for $1 \le i < j \le L$,

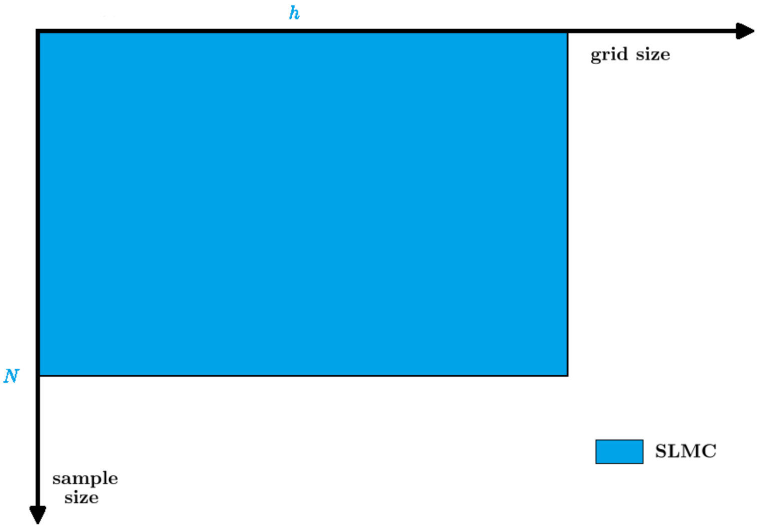
$$\frac{N_j}{N_i} = \sqrt{\frac{\mathcal{C}_i}{\mathcal{C}_j} \cdot \frac{\mathcal{V}_j}{\mathcal{V}_i}} = \left(\frac{h_j}{h_i}\right)^{\frac{\beta+\gamma}{2}} = 2^{-\frac{(j-i)(\beta+\gamma)}{2}} < 1, \tag{9}$$

which implies that the finer the mesh is, the fewer the realizations we need to compute.

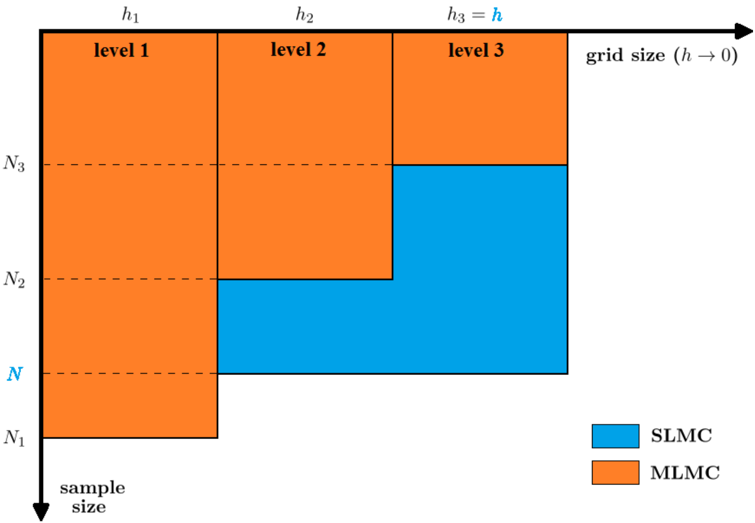
Now note that, the sample size of initial level could be determined by $\mathcal{C}_0$, $\mathcal{V}_0$ and $\mathcal{V}_1$ defined on coarse meshes by low cost, while the others could be derived via (9).

[§] M. B. Giles, Multilevel Monte Carlo methods, Acta Numerica, 2015.
[§] S. Pauli et al, Determining optimal multilevel Monte Carlo parameters with application to fault tolerance, CMA, 2015.

Total Cost of SLMC



Total Cost of MLMC



2D Stochastic Optimal Control Problem

Given a deterministic target function $U(\mathbf{x}) = \sin(2\pi x)(\cos(2\pi y) - 1)$ and $\beta = 10^{-4}$, find the optimal solution pair $(u^{\mathsf{opt}}, f^{\mathsf{opt}}) \in \mathcal{H}_0^1(D) \otimes L^2(D)$ such that

$$\widehat{\mathcal{J}}_\beta(u, f) = \frac{1}{2N} \sum_{i=1}^{N} \int_D |u(\mathbf{x}, \omega_i) - U(\mathbf{x})|^2 d\mathbf{x} + \frac{\beta}{2} \int_D |f(\mathbf{x})|^2 d\mathbf{x}$$

subject to the stochastic elliptic PDEs

$$\begin{cases} -\nabla \cdot (a(\mathbf{x}, \omega_i) \nabla u(\mathbf{x}, \omega_i)) = f(\mathbf{x}) & \text{in } D = [0,1]^2 \\ u(\mathbf{x}, \omega_i) = 0 & \text{on } \partial D \end{cases} \quad \text{for } \ i = 1, ..., N,$$

is minimized, where the coefficient $a(\mathbf{x}, \omega) = e^{g(\mathbf{x}, \omega)}$ is given as before.

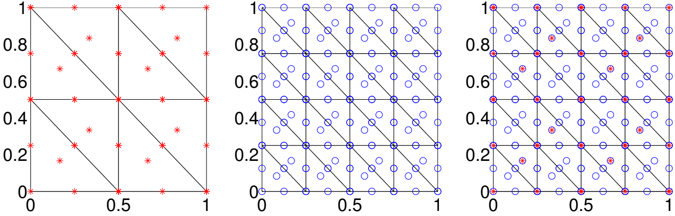
Using 4-level grids $\{h_\ell = 2^{-\ell-2}\}_{\ell=0}^{3}$, the MLMC approximation of $\mathbb{E}[u^{\mathsf{opt}}]$ is given by

$$\mathbb{E}\left[u_3^{\mathsf{opt}}\right] = \frac{1}{N_0} \sum_{i=1}^{N_0} u_0^{\mathsf{opt},(i)} + \sum_{\ell=1}^{3} \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} (u_\ell^{\mathsf{opt},(i)} - u_{\ell-1}^{\mathsf{opt},(i)}).$$

Representations of Random Fields Again

In order to utilize the characteristics of GBR, we use the following 7 points quadrature rule of order 3 that defined on a reference triangle with vertices $(0,0)$, $(1,0)$ and $(0,1)$.

weight	$9/40$	$1/40$	$1/40$	$1/40$	$1/15$	$1/15$	$1/15$
coordinate	$(1/3,1/3)$	$(1,0)$	$(0,1)$	$(0,0)$	$(0,1/2)$	$(1/2,0)$	$(1/2,1/2)$



Nested quadrature points from nested finite element spaces.

Sensitive Analysis

Two key observations:

- the optimal solution $u^{\text{opt}}(\mathbf{x}, \omega)$ of surrogate model problem (5-6) solves elliptic problem (1) with a particular right hand side $f^{\text{opt}}(\mathbf{x})$, while $u^{\text{ite}}(\mathbf{x}, \omega)$ with $f^{\text{ite}}(\mathbf{x})$ during the iterations,
- it is cheaper/easier to determine the parameters α, β, γ numerically/theoretically for elliptic problem than the optimal control problem.

Inspiration: use the parameters of state constraint for the optimal control problem.

The parameters of state problem are numerically determined with various RHSs:

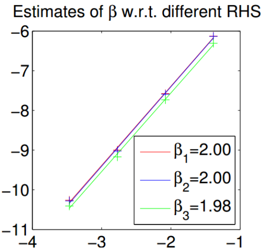
$$\text{Smooth functions}: \quad f(\mathbf{x}) = \begin{cases} \sin(x)\cos(y) & \text{in } D = [0,1]^2, \\ x^2 + y^2 & \text{in } D = [0,1]^2, \\ e^{x+y} & \text{in } D = [0,1]^2. \end{cases} \tag{10}$$

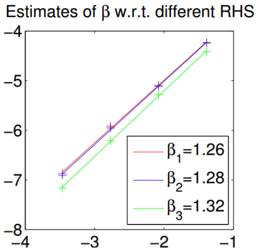
$$\text{Non-smooth functions}: \quad f(\mathbf{x}) = \sigma \sum_{k=1}^{K} \frac{1}{\sqrt{V_k}} \chi_k(\mathbf{x}) X_k(\omega) \quad \text{in } D = [0,1]^2. \tag{11}$$

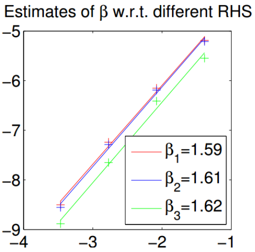
[§] J. Charrier et al, Finite element analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods, SIAM J.N.A., 2013.

β almost doesn't change w.r.t. smooth RHSs (10) when the same coefficient is provided.

	$f(\mathbf{x})$ in certain norm			β w.r.t. different norm		
	$\|\cdot\|_{L^2}$	$\lvert\cdot\rvert_{H^1}$	$\|\cdot\|_{L^\infty}$	$\|\cdot\|_{L^2}$	$\lvert\cdot\rvert_{H^1}$	$\|\cdot\|_{L^\infty}$
$\sin(x)\cos(y)$	0.45	0.78	0.84	2.00	1.26	1.59
x^2+y^2	0.79	1.63	1.99	2.00	1.28	1.61
e^{x+y}	3.19	4.52	7.34	1.98	1.32	1.62



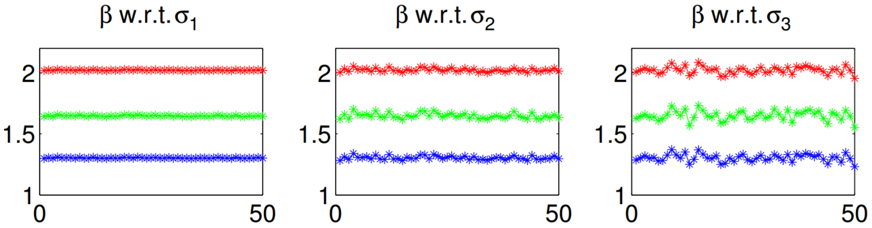




Sensitive Analysis - Non-smooth Functions

For discontinuous RHSs (11) with different amplitudes σ and random inputs $\{X_k(\omega)\}_{k=1}^K$, β almost doesn't change as well when the same coefficient is provided.

	$\mathbb{E}[\beta]$ $\|\cdot\|_{L^2}$	$\mathbb{E}[\beta]$ $\|\cdot\|_{H^1}$	$\mathbb{E}[\beta]$ $\|\cdot\|_{L^\infty}$	$\mathbb{V}[\beta]$ $\|\cdot\|_{L^2}$	$\mathbb{V}[\beta]$ $\|\cdot\|_{H^1}$	$\mathbb{V}[\beta]$ $\|\cdot\|_{L^\infty}$
$\sigma_1 = 3.2\sqrt{2}$	2.02	1.30	1.65	7.37×10^{-6}	8.46×10^{-6}	1.89×10^{-5}
$\sigma_2 = 16\sqrt{2}$	2.02	1.30	1.65	1.83×10^{-4}	2.12×10^{-4}	4.56×10^{-4}
$\sigma_3 = 32\sqrt{2}$	2.02	1.30	1.65	8.20×10^{-4}	9.18×10^{-4}	1.53×10^{-3}



Sensitive Analysis - Non-smooth Functions

Furthermore, some sensitive analysis results for the parameter γ are given as follows.

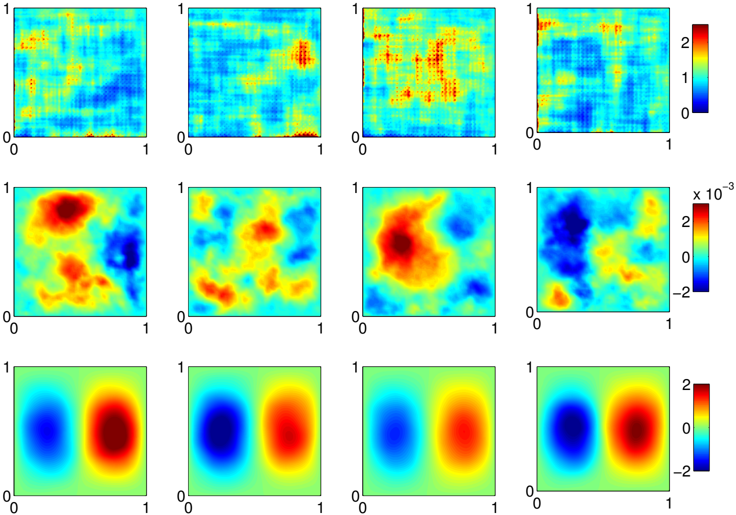
	$\mathbb{E}[\gamma]$		$\mathbb{V}[\gamma]$	
	CPU time	Tic-Toc time	CPU time	Tic-Toc time
$\sigma_1 = 3.2\sqrt{2}$	1.95	1.97	1.29×10^{-2}	3.56×10^{-3}
$\sigma_2 = 16\sqrt{2}$	1.97	1.97	1.06×10^{-2}	3.26×10^{-3}
$\sigma_3 = 32\sqrt{2}$	1.91	1.96	1.39×10^{-2}	2.46×10^{-3}

Observation: the value of γ roughly equals to the dimension of problem in our case.

Conclusion: given that $\mathbb{V}[\beta]$ and $\mathbb{V}[\gamma]$ are small, the values of β and γ would be very stable if $f(\mathbf{x})$ varies in a certain range for both smooth and non-smooth functions.

Application: in surrogate model problem (5-6), the RHSs of both state and adjoint equations vary along with the iterations. Then by our conclusion, the parameters β and γ of our surrogate model problem are stable when the same coefficient is provided and can be numerically obtained via sensitive analysis.

Realizations of $a(\mathbf{x},\omega)$, $u^{\text{ini}}(\mathbf{x},\omega)$ and $u^{\text{opt}}(\mathbf{x},\omega)$



Opt 1: MLMC vs SLMC

WLOG, given the total cost $T_c^* = 1.7 \times 10^2$ seconds of tic-toc time, we have

Sample sizes w.r.t. Opt 1

	N_0^{ML}	N_1^{ML}	N_2^{ML}	N_3^{ML}	N_3^{SL}
$\|\cdot\|_{L^2}$	480	130	50	20	40
$\vert \cdot \vert_{H^1}$	400	130	60	25	40
$\|\cdot\|_{L^\infty}$	440	140	60	20	40

Optimal errors w.r.t. Opt 1

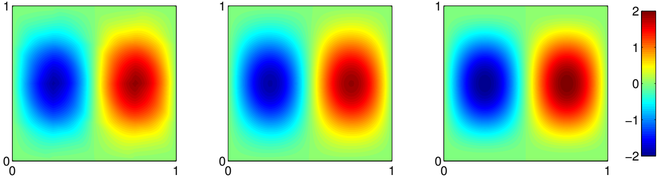
	$e_3^{ML} - C_F h_3^\alpha$	$e_3^{SL} - C_F h_3^\alpha$	ratio
$\|\cdot\|_{L^2} (\times 10^{-4})$	1.63	6.96	23.45%
$\vert \cdot \vert_{H^1} (\times 10^{-3})$	1.93	5.42	35.69%
$\|\cdot\|_{L^\infty} (\times 10^{-4})$	7.45	24.31	30.67%

Opt 1: MLMC vs SLMC - Continued

Optimal control results w.r.t. Opt 1

	$\widehat{\mathcal{J}}_\beta^{\text{ini}}$	$\widehat{\mathcal{J}}_\beta^{\text{opt}}$			ratio		
		$\Vert\cdot\Vert_{L^2}$	$\vert\cdot\vert_{H^1}$	$\Vert\cdot\Vert_{L^\infty}$	$\Vert\cdot\Vert_{L^2}$	$\vert\cdot\vert_{H^1}$	$\Vert\cdot\Vert_{L^\infty}$
MLMC($\times 10^{-2}$)	23.06	1.18	1.17	1.18	5.13%	5.08%	5.11%
SLMC($\times 10^{-2}$)	23.06	0.98	0.98	0.98	4.26%	4.26%	4.26%

$\mathbb{E}[u^{\text{opt}}_{\text{MLMC}}]$, $\mathbb{E}[u^{\text{opt}}_{\text{SLMC}}]$ and target function $U(\mathbf{x})$ w.r.t. $\Vert\cdot\Vert_{L^2}$ norm



Opt 2: MLMC vs SLMC

WLOG, given the objective accuracy $e_3^* - \mathcal{O}(h_3^\alpha) = 8.4 \times 10^{-4}$, we have

Sample sizes w.r.t. Opt 2

	N_0^{ML}	N_1^{ML}	N_2^{ML}	N_3^{ML}	N_3^{SL}
$\|\cdot\|_{L^2}$	120	30	10	5	35
$\lvert\cdot\rvert_{H^1}$	1200	360	120	40	260
$\|\cdot\|_{L^\infty}$	500	120	40	15	115

Optimal workload w.r.t. Opt 2

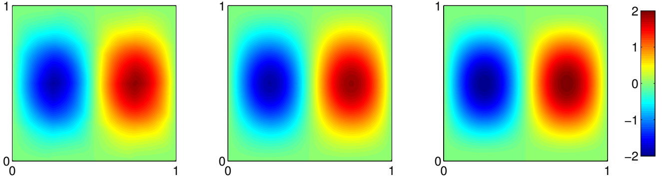
	T_c^{ML}	T_c^{SL}	ratio
$\|\cdot\|_{L^2}(\times 10^1)$	6.19	16.42	37.71%
$\lvert\cdot\rvert_{H^1}(\times 10^2)$	7.54	12.51	60.24%
$\|\cdot\|_{L^\infty}(\times 10^2)$	2.77	5.60	49.34%

Opt 2: MLMC vs SLMC - $\|\cdot\|_{L^2}$ Norm

Optimal control results w.r.t. Opt 1

	$\widehat{\mathcal{J}}_\beta^{\text{ini}}$	$\widehat{\mathcal{J}}_\beta^{\text{opt}}$ $\|\cdot\|_{L^2}$	$\widehat{\mathcal{J}}_\beta^{\text{opt}}$ $\lvert\cdot\rvert_{H^1}$	$\widehat{\mathcal{J}}_\beta^{\text{opt}}$ $\|\cdot\|_{L^\infty}$	ratio $\|\cdot\|_{L^2}$	ratio $\lvert\cdot\rvert_{H^1}$	ratio $\|\cdot\|_{L^\infty}$
MLMC($\times 10^{-2}$)	23.06	1.18	1.21	1.20	5.12%	5.25%	5.21%
SLMC($\times 10^{-2}$)	23.06	0.96	0.98	0.97	4.15%	4.23%	4.22%

$\mathbb{E}[u^{\text{opt}}_{\text{MLMC}}]$, $\mathbb{E}[u^{\text{opt}}_{\text{SLMC}}]$ and target function $U(\mathbf{x})$ w.r.t. $\|\cdot\|_{L^2}$ norm



Concluding Remarks

- MLMC inherits the advantage of Monte Carlo method as being insensitive to the dimension of probability space.
- MLMC performs most simulations with low accuracy at a low cost, with relatively few simulations being performed at high accuracy and a high cost.
- By applying a grid-based representation to the stochastic coefficient, the dimensionality of random fields in our MLMC scheme is also a multilevel type, which further reduces the computational cost and avoids the truncation error in KLE.
- The stability of parameters is verified via sensitivity test, which indicates the applicability of the MLMC formulas for our stochastic optimal control problem.
- The idea using the parameters of state constraint for the optimization problem is shown to be an economical and reliable approach. This can be a contribution that is potentially applicable to other optimal control problems with different objective functionals and state equations.
- Theoretical verification of MLMC assumption, preconditioned conjugate gradient method and optimal choices of parameters will be carried out in the future.

Thank you