Multi Level Monte Carlo methods with Control Variate for elliptic SPDEs

Fabio Nobile[1,2], Erik von Schwerin[3], Raul Tempone[4], Francesco Tesei[1]

[1]CSQI-MATHICSE, EPFL, Lausanne, Switzerland, [2]MOX - Politecnico di Milano, Italy, [3]KTH - Royal Institute of Technology, Stockholm, Sweden, [4] KAUST - Thuwal, Kingdom of Saudi Arabia

Part I: Goal and Motivations

Goal: let $Q(u)$ be the quantity of interest (QoI) related to the solution u of an elliptic PDE with random diffusion coefficient; we aim to compute it efficiently in the case in which deterministic methods are not effective e.g. when the coefficient has a low regularity.
Idea: combine the Stochastic Collocation (SC), which is very effective in the case of smooth input random field, with the Multi Level Monte Carlo (MLMC) in order to reduce even more the computational cost needed to get a prescribed mean square error (MSE) in the case of rough coefficient, e.g. with exponential covariance.
How: by using the solution of an auxiliary problem with regularized coefficient as control variate, whose mean can be well approximated by a SC scheme.

Mathematical Model

The model considered is the Darcy model, widely used to describe for instance the groundwater flow problem. The mathematical formulation is:

$$\text{find } u(x,\omega) : \overline{D} \times \Omega \to \mathbb{R} \text{ such that almost everywhere in } \omega \in \Omega \quad \text{holds :}$$

$$\mathrm{div}(-a(x,\omega)\nabla u(x,\omega)) = f(x) \text{ in } D \quad + \quad \text{BC on } \partial D, \tag{1}$$

with $a(x,\omega) = e^{\gamma(x,\omega)}$ lognormal stationary random field s.t. $\mathbb{E}[\gamma](x) = \mu(x) = \mu$,

$$\text{Matérn Family: } \mathrm{cov}_\gamma(x_1,x_2) = \frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}\left(\sqrt{2\nu}\frac{\|x_1-x_2\|}{L_c}\right)^\nu K_\nu\left(\sqrt{2\nu}\frac{\|x_1-x_2\|}{L_c}\right)$$

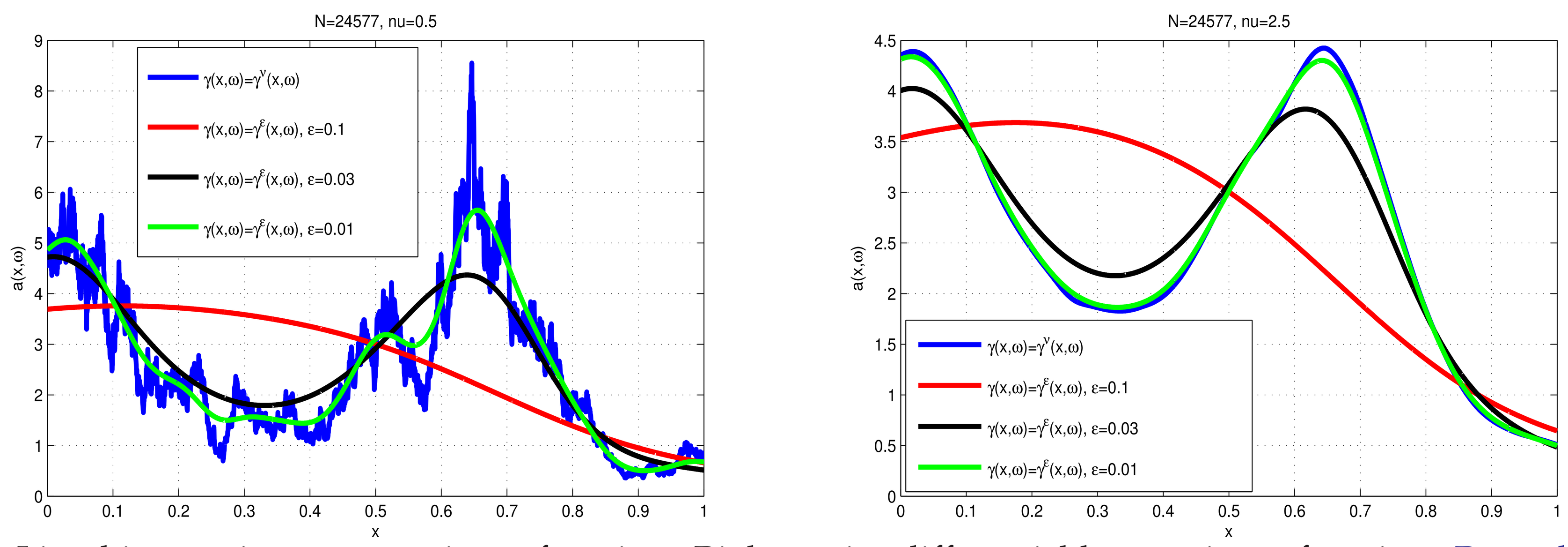
Γ is the gamma function, K_ν is the modified Bessel function of the second kind.
IMPORTANT: such input random field has a covariance function which is $\mathcal{C}^s$ and, in turn, its realizations are Hölder continuous, namely $\mathcal{C}^{s,\alpha}(D)$, with $2\nu = s+\alpha$, $s \in \mathbb{N}$, $\alpha \in (0,1]$. For $\nu = 0.5$ the covariance function is only Lipschitz continuous and the realizations $\mathcal{C}^{0,\alpha}(D)$ with $\alpha < 0.5$.

Control Variate

Let $Q = Q(u)$ be the QoI obtained starting from the solution of problem (1) having a random field γ with Matérn covariance family with parameter ν; this is the quantity that we want to estimate.
Let $Q^\epsilon = Q(u^\epsilon)$ be the QoI related to the solution of problem (1) having a regularized version of the input random field, obtained via convolution with a Gaussian kernel, namely γ^ϵ.

$$\gamma^\epsilon = \gamma^\nu * \phi_\epsilon(x), \text{ where } \phi_\epsilon(x) = \frac{1}{(2\pi\epsilon^2)^{\frac{d}{2}}} e^{-\frac{\|x\|^2}{2\epsilon^2}}$$

$$\text{Control Variate QoI :} \quad Q^{CV} = Q - (Q^\epsilon - \mathbb{E}[Q^\epsilon])$$



Left: Lipschitz continuous covariance function. Right: twice differentiable covariance function. Remark 1: the smaller ϵ is the more the two random fields γ^ν and γ^ϵ are correlated. Moreover if ϵ tends to 0 the random field γ^ϵ tends to γ^ν.
Remark 2: as the two random fields γ^ν and γ^ϵ get highly correlated the variance of the control variate $\mathbb{V}\mathrm{ar}(Q^{CV})$ gets smaller, namely $\mathbb{V}\mathrm{ar}[Q^{CV}] << \mathbb{V}\mathrm{ar}[Q^\nu]$.

Part II: MLMC with Control Variate (MLCV)

$\mathcal{T}_{h_0},\dots,\mathcal{T}_{h_L}$: sequence of increasingly fine triangulations;
$Y_{h_l}^{CV} = Q_{h_l}^{CV} - Q_{h_{l-1}}^{CV}$: difference of the QoI between two consecutive grids.

$$\text{Telescopic sum + Linearity of expectation:} \quad \mathbb{E}[Q_{h_L}^{CV}] = \sum_{l=0}^{L}\mathbb{E}[Y_{h_l}^{CV}], \quad Q_{h_{-1}}^{CV} = 0$$

$$\text{Estimator: } \hat{Q}^{MLCV}_{h_L,\{M_l\}} = \sum_{l=0}^{L}\frac{1}{M_l}\sum_{i=1}^{M_l}\left(Q_{h_l}^i - Q_{h_{l-1}}^i - (Q_{h_l}^{\epsilon,i} - Q_{h_{l-1}}^{\epsilon,i})\right) + \mathbb{E}[Q_{h_L}^{\epsilon,SC}]$$

$$\text{MSE: } e(\hat{Q}^{MLCV}_{h_L,\{M_l\}})^2 \leqslant \sum_{l=0}^{L}\frac{\mathbb{V}\mathrm{ar}(Y_{h_l}^{CV})}{M_l} + 2\mathbb{E}[Q_{h_L}^\epsilon - Q_{h_L}^{\epsilon,SC}]^2 + 2\mathbb{E}[Q_{h_L} - Q]^2$$
$$= (i) + (ii) + (iii)$$

$\mathbb{E}[Q_{h_L}^{\epsilon,SC}]$: mean of the QoI $Q_{h_L}^\epsilon$ computed with a SC scheme on sparse grids.
M_l: number of samples on each level. A good choice of M_l, for $l = 0,\dots,L$, represents a crucial issue for the effectiveness of the method.
(i): represents the variance of the estimator, i.e. the statistical error: it is expected to be significantly smaller than the variance of standard MLMC.
(ii): represents the error due to the approximation of the mean of the control variate $\mathbb{E}[Q^\epsilon]$ by a SC scheme;
(iii): represent the error due to the finite element (FE) discretization;
Remark 1: on each level the variance reduction is due to the difference of the QoI between consecutive grids as well as to the presence of the control variate;
Remark 2: if ϵ tends to 0 the statistical error (i) vanishes; on the other hand keeping small the SC error (ii) becomes too costly.

MLCV Algorithm

1. Given a prescribed tol select h_L in such a way to have $(iii) \leqslant tol^2$;
2. set $h_0 = O(|D|)$ and evaluate $(iv) = \mathbb{V}\mathrm{ar}(Y_{h_l}^{CV})$ and $(v) = \mathbb{V}\mathrm{ar}(Q_{h_l}^{CV})$; we can select among two basic strategies:
   Strategy 1: if $(iv) < (v)\ \forall\, l$ apply the MLCV scheme starting from level 0;
   Strategy 2: if $(iv) \approx (v)$ for $l = 0,\dots,l_0$ set $h_0 = h_{l_0}$; and use the control variate only on level l_0 and a standard MLMC on subsequent levels, namely

$$\hat{Q}^{MLCV}_{h_L,\{M_l\}} = \frac{1}{M_{l_0}}\sum_{i=1}^{M_{l_0}}\left(Q_{h_{l_0}}^i - Q_{h_{l_0}}^{\epsilon,i}\right) + \sum_{l=l_0+1}^{L}\frac{1}{M_l}\sum_{i=1}^{M_l}\left(Q_{h_l}^i - Q_{h_{l-1}}^i\right) + \mathbb{E}[Q_{h_{l_0}}^{\epsilon,SC}]$$

3. according to the strategy selected compute the number of samples M_l for $l = 0,\dots,L$ and the number of knots of the sparse grid M_{SG} by solving an optimization problem in such a way to have $(i) + (ii) \leqslant tol^2$

Computing M_l and M_{SG}: optimization problem

- ϵ is considered fixed
- C_l is the computational cost needed to solve one deterministic system on the grid of mesh size h_l; cost model: $C_l = 2C_{l-1} = \dots = 2^l C_0 = \gamma 2^l h_0^{-1}$.
  For instance see [Cliffe, Giles, Scheichl, Teckentrup, '11]

Computational cost

- strategy 1: $C(M_l, M_{SG}) = 2M_0C_0 + 2\sum_{l=1}^{L} M_l(C_l + C_{l-1}) + M_{SG}C_L$
- strategy 2: $C(M_l, M_{SG}) = 2M_{l_0}C_{l_0} + \sum_{l=l_0+1}^{L} M_l(C_l + C_{l-1}) + M_{SG}C_{l_0}$

Associated Error fitted model

- strategy 1: $e(M_l, M_{SG}) = \sum_{l=0}^{L}\frac{\min\left\{c_1h_l^{2\min\{\nu,p\}}, c_2h_l^{4\min\{\nu,p\}}\right\}}{M_l} + c_3M_{SG}^\alpha$
- strategy 2: $e(M_l, M_{SG}) = \sum_{l=l_0}^{L}\frac{c_1h_l^{4\min\{\nu,p\}}}{M_l} + c_2M_{SG}^\alpha$

Perform a Lagrange optimization by considering

$$\mathcal{L}(M_l, M_{SG}, \lambda) = C(M_l, M_{SG}) - \lambda(e(M_l, M_{SG}) - tol^2)$$

Part III: Numerics

Numerical test 1D: consider the Darcy equation (1) in $(0,1)$ with boundary conditions $u(0,\omega) = 1$, $u(1,\omega) = 0$; external force $f(x) = 0$. We are interested in evaluating the QoI $Q(u^\nu) = \int_D u^\nu(x,\omega)dx$.
The random field $a(x,\omega)$ has been approximated with $N = 2K+1$ random variables $y_n(\omega) \sim N(0,1)$ iid by using a Fourier expansion computed on the interval $(-T,T)$ that contains $(-1,1)$

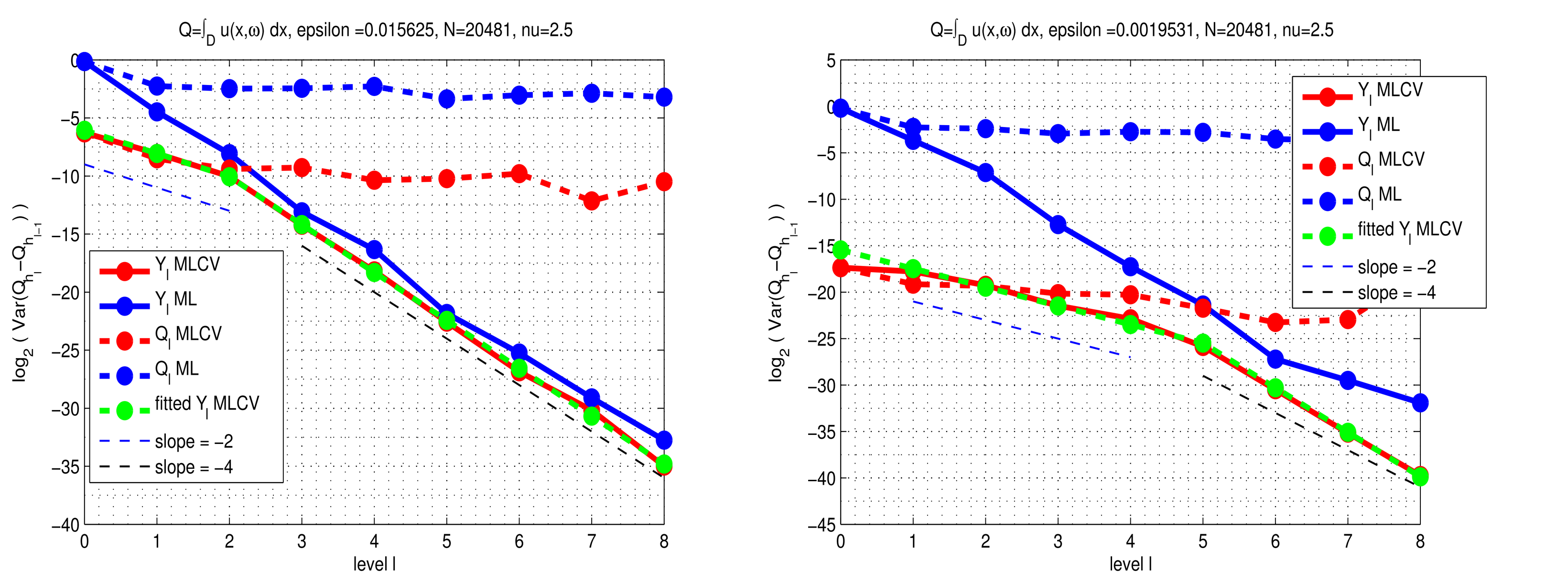
$$\gamma(x,\omega) \approx \gamma_N(x,y_1,\dots,y_N) = \sigma a_0 y_0 + \sigma\sum_{n=1}^{K} a_n\left(y_{2n}\cos(\pi n x/T) + y_{2n-1}\sin(\pi n x/T)\right),$$

Since the $\{a_n\}$ are the Fourier coefficients of the covariance function, they have been computed via FFT. To do this we chose to consider an interval $(-T,T)$ in order to make negligible the errors due to the fact that we are representing periodically a covariance function that is not periodic.
Remark: the number N of r.v. kept in the truncation γ_N of the random field γ has to be chosen in such a way that the maximal frequency remains bigger than h_L^{-1}.

Variance Reduction: MLCV vs. MLMC

Error estimate (preliminary result)

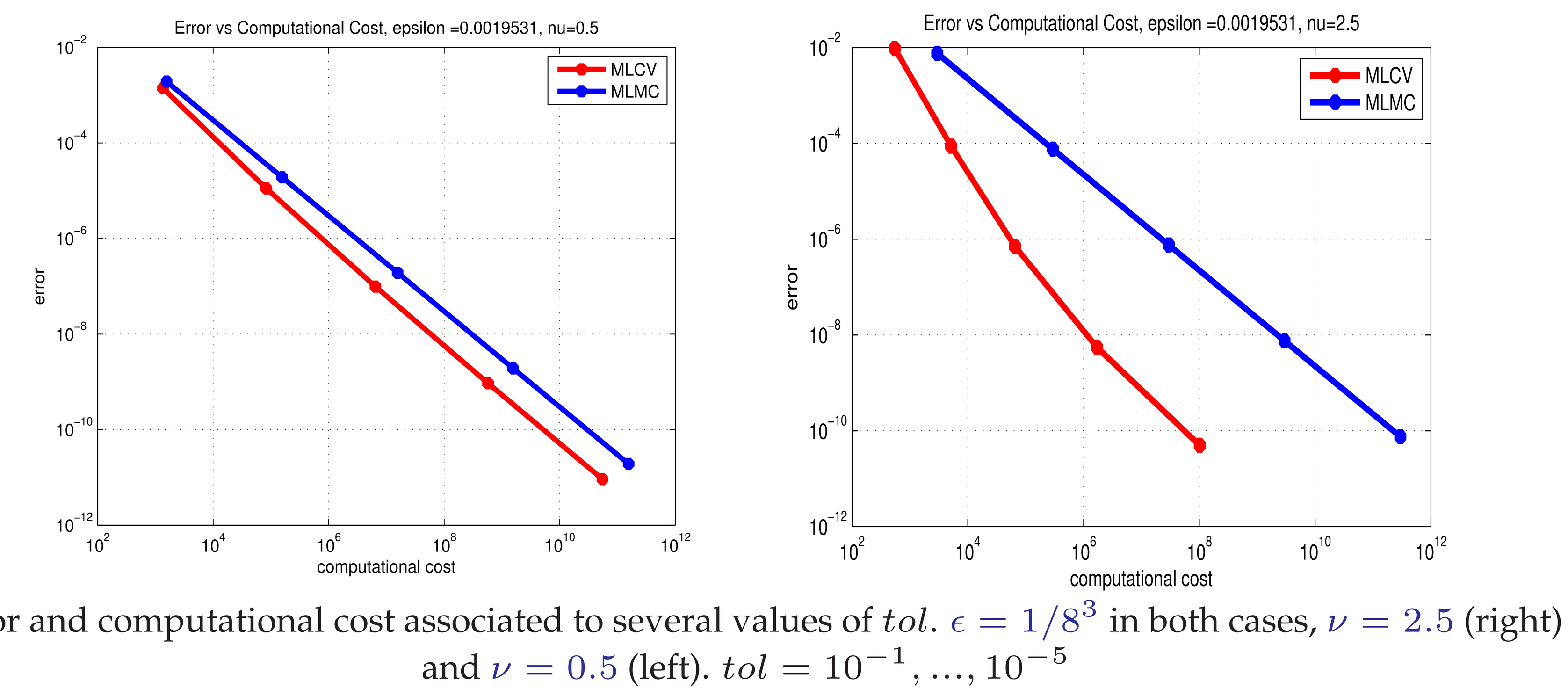
$$\mathbb{E}\left[(Q - Q_{h_l} - (Q^\epsilon - Q_{h_l}^\epsilon))^2\right]^{\frac{1}{2}} \leq \min\left\{c_1 h_l^{\min\{\nu,p\}}\epsilon^{\min\{\nu,2\}}, c_2h_l^{2\min\{\nu,p\}}\epsilon^{\min\{(\nu-p)_+,2\}}\right\}$$



Variance of the difference of the QoI between consecutive grids. The dashed lines represent the slopes h_l^2 and h_l^4. $\nu = 2.5$. Left: $\epsilon = 1/8^2$, right: $\epsilon = 1/8^3$.

Error vs Cost: MLCV vs MLMC

error = statistical error + SC error



Error and computational cost associated to several values of tol. $\epsilon = 1/8^3$ in both cases, $\nu = 2.5$ (right) and $\nu = 0.5$ (left). $tol = 10^{-1},\dots,10^{-5}$

Perspectives and Future Work

- Perform an analysis in order to distribute as conveniently as possible the computational cost needed to get a prescribed MSE;
- Need to find the optimal tradeoff between CV and SC: try to include the choice of ϵ in the optimization procedure.

References

[1] A.L. Teckentrup, R. Scheichl, M.B. Giles, E. Ullmann. *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*, preprint, 2012.

[2] A. Barth, C. Schwab, N. Zollinger, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*. Numerische Mathematik, 119(1):123-161, 2011

[3] I. Babuška, F. Nobile, R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM REVIEW Vol. 52, No. 2, pp. 317-355

[4] J. Charrier, R. Scheichl, A. Teckentrup, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods* , 2011