A Multigrid Multilevel Quasi-Monte Carlo Method with Sample Reuse

Pieterjan Robbe
Dirk Nuyens
Stefan Vandewalle

MCQMC • July 6, 2018

KU LEUVEN

Department of Computer Science
NUMA research group
Celestijnenlaan 200A bus 2402
3001 Leuven, Belgium

Contents

PART 1: MG-MLMC **without** sample reuse

PART 2: MG-MLMC **with** sample reuse

PART 3: Results for the elliptic model problem

Model parametric elliptic PDE

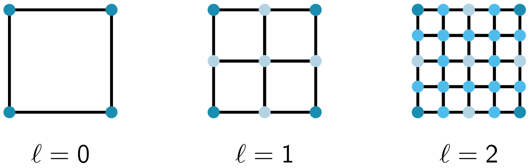
$$-\nabla \cdot a(\boldsymbol{x}, \boldsymbol{y}) \nabla u(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x})$$

- Early work
  - [Ghanem, Spanos, 1997]
  - [Babuska, Tempone, Zouraris, 2004]
  - [Babuska, Nobile, Tempone, 2007]
  - and many others
- Parametric PDE setting in
  - [Cohen, DeVore, Schwab, 2011]
- Recent interest from multilevel/QMC community
  - [Graham, Kuo, Nuyens, Scheichl, Sloan, 2011]
  - [Cliffe, Giles, Scheichl, Teckentrup, 2011]
  - [Kuo, Schwab, Sloan, 2012]
  - [Kuo, Nuyens, 2016]
  - and many others

PART 1

MG-MLMC **without** sample reuse

A hierarchy of coarser grids



- Goal: compute statistics of **quantity of interest**

$$F(\boldsymbol{y}) := F(u(\boldsymbol{x}, \boldsymbol{y}))$$

- Solution of the PDE (and hence quantity of interest F) is approximated numerically
- Suppose we have a **hierarchy of approximations** F_ℓ, $\ell = 0, \ldots, L$ and $F_\ell \to F$ as $\ell \to \infty$
- Do not compute $\mathrm{E}[F_L]$ by sampling from F_L, but by sampling from the whole hierarchy F_ℓ, $\ell = 0, \ldots, L$

Multilevel Monte Carlo (MLMC)

- Basis is the **telescoping sum**

$$\begin{aligned}\mathsf{E}[F_L] &= \mathsf{E}[F_0] + \mathsf{E}[F_1 - F_0] + \ldots + \mathsf{E}[F_L - F_{L-1}] \\ &= \mathsf{E}[F_0] + \sum_{\ell=1}^{L} \mathsf{E}[F_\ell - F_{\ell-1}]\end{aligned}$$

- MLMC estimator uses Monte Carlo to estimate each term:

$$\mathcal{Q}_L^{\mathrm{MC}} = \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} \left(F_\ell(\boldsymbol{y}_\ell^{(n)}) - F_{\ell-1}(\boldsymbol{y}_\ell^{(n)}) \right) \quad (F_{-1} := 0)$$

- Crucially, use the **same random numbers** $\boldsymbol{y}_\ell^{(n)}$ in each sample

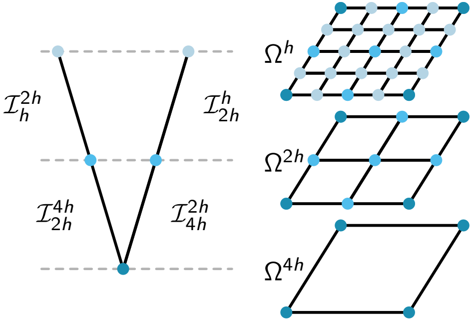
$$\begin{aligned}V_\ell := \mathsf{V}[F_\ell - F_{\ell-1}] &= \mathsf{V}[F_\ell] + \mathsf{V}[F_{\ell-1}] - 2\mathrm{cov}(F_\ell, F_{\ell-1}) \\ &\ll \mathsf{V}[F_\ell] + \mathsf{V}[F_{\ell-1}]\end{aligned}$$

- Most samples will be taken on the coarse grid, where samples are cheap, and only few samples are needed on the finest grid

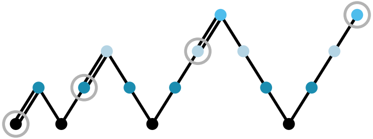
Full Multigrid (FMG)

- **Full Multigrid** can compute a solution to discretization accuracy in $\mathcal{O}(M)$ time, where M is the number of DOF
- FMG also computes free solutions on coarser grids

V-cycle



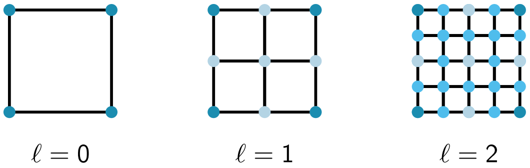
FMG-cycle

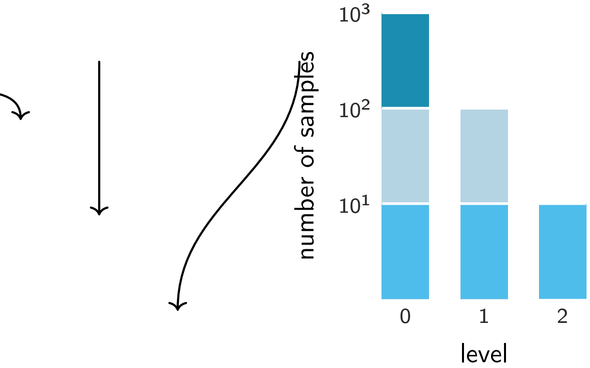


PART 2

MG-MLMC **with** sample reuse

Recycling samples





Multigrid Multilevel Monte Carlo (MG-MLMC)

- Idea is to **recycle** the coarse solutions from the FMG method as coarse samples in the MLMC method
- MG-MLMC estimator [Kumar, Oosterlee, Dwight, 2017]

$$\mathcal{Q}^{\mathrm{MC}}_{L,\mathrm{reuse}} := \sum_{\ell=0}^{L} \left(\frac{1}{\sum_{i=\ell}^{L} N_i} \right) \sum_{k=\ell}^{L} \sum_{n=1}^{N_k} \left(F_\ell(\boldsymbol{y}_k^{(n)}) - F_{\ell-1}(\boldsymbol{y}_k^{(n)}) \right)$$

- For example, for $L = 2$ we find

$$\begin{aligned}
\mathcal{Q}^{\mathrm{MC}}_{2,\mathrm{reuse}} = {} & \\
& \frac{1}{N_0 + N_1 + N_2} \left(\sum_{n=1}^{N_0} F_0(\boldsymbol{y}_0^{(n)}) + \sum_{n=1}^{N_1} F_0(\boldsymbol{y}_1^{(n)}) + \sum_{n=1}^{N_2} F_0(\boldsymbol{y}_2^{(n)}) \right) \\
& + \frac{1}{N_1 + N_2} \left(\sum_{n=1}^{N_1} \left(F_1(\boldsymbol{y}_1^{(n)}) - F_0(\boldsymbol{y}_1^{(n)}) \right) + \sum_{n=1}^{N_2} \left(F_1(\boldsymbol{y}_2^{(n)}) - F_0(\boldsymbol{y}_2^{(n)}) \right) \right) \\
& + \frac{1}{N_2} \left(\sum_{n=1}^{N_2} \left(F_2(\boldsymbol{y}_2^{(n)}) - F_1(\boldsymbol{y}_2^{(n)}) \right) \right)
\end{aligned}$$

Multigrid Multilevel Monte Carlo (MG-MLMC)

- Variance of the MG-MLMC estimator is

$$\mathrm{V}[\mathcal{Q}^{\mathrm{MC}}_{L,\mathrm{reuse}}] = \sum_{\ell=0}^{L} \left(\frac{V_\ell}{\sum_{i=\ell}^{L} N_i} \right) + 2 \sum_{0 \le \ell < \tau \le L} \rho_{\ell\tau} \sqrt{\left(\frac{V_\ell}{\sum_{i=\ell}^{L} N_i} \right) \left(\frac{V_\tau}{\sum_{i=\tau}^{L} N_i} \right)}$$

- 3 approaches to obtain a variance estimate:

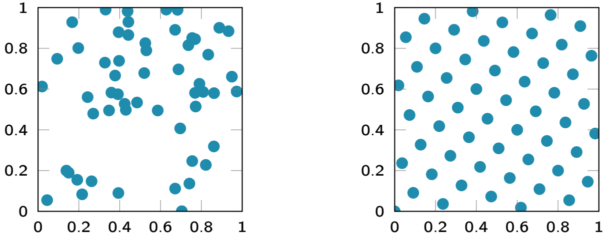
1. Assume the correlation coefficients $\rho_{\ell\tau} \approx 1$ and **analytically** solve optimization problem to compute N_ℓ
2. Use the **debiasing technique** from [Rhee, Glynn, 2015]: randomization of the final level L
3. Randomly shifted lattice rules from **Quasi-Monte Carlo** literature

Quasi-Monte Carlo (QMC)

- A **Quasi-Monte Carlo** method uses well-chosen sample points, as opposed to the random points with Monte Carlo
- A popular choice are **rank-1 lattice rules**

$$\boldsymbol{t}^{(n)} := \frac{n\boldsymbol{z} \bmod N}{N} = \left\{ \frac{n\boldsymbol{z}}{N} \right\}$$

where $\boldsymbol{z} \in \mathbb{Z}_N^s$ is a generating vector and $\{\cdot\}$ denotes mod 1
- Can potentially obtain $\mathcal{O}(1/N)$ convergence, if integrand is *sufficiently smooth and decaying importance of dimensions*[1,2]



[1] more details to be found in standard works such as [Dick, Kuo, Sloan, 2013]
[2] see also [Kuo, Nuyens, 2016] for the lognormal case

Obtaining a variance estimate

- Lattice points are chosen deterministically, hence correlated
- Solution is **random shifting**:

$$\bar{\mathcal{Q}}^{\mathrm{QMC}}_{L,P,\mathrm{reuse}} := \frac{1}{P}\sum_{p=1}^{P}\sum_{\ell=0}^{L}\left(\frac{1}{\sum_{i=\ell}^{L} N_i}\right)\sum_{k=\ell}^{L}\sum_{n=1}^{N_k}\left(F_\ell(\boldsymbol{y}^{(n)}_{k,p}) - F_{\ell-1}(\boldsymbol{y}^{(n)}_{k,p})\right)$$

where $\boldsymbol{y}^{(n)}_{k,p} := \Phi^{-1}\left(\left\{\boldsymbol{t}^{(k)}_{\ell} + \boldsymbol{u}^{(p)}_{k}\right\}\right)$ and $\boldsymbol{u}^{(p)}_{k} \sim \boldsymbol{U}(0,1)$, i.i.d

- Sample variance is used as an estimate for the variance

$$\mathbb{V}[\bar{\mathcal{Q}}^{\mathrm{QMC}}_{L,P,\mathrm{reuse}}] \approx \frac{1}{P(P-1)}\sum_{p=1}^{P}\left(\mathcal{Q}^{\mathrm{QMC}}_{L,p,\mathrm{reuse}} - \bar{\mathcal{Q}}^{\mathrm{QMC}}_{L,P,\mathrm{reuse}}\right)^2$$

Cost analysis

Proposition

Let β be the rate of the decrease in variance and γ the rate of the increase in cost per level ℓ, and $1/\lambda$ is the convergence rate of the randomized quasi-Monte Carlo method, or $\lambda = 1$ when using Monte Carlo.
Then, the cost reduction factor of the Multigrid Multilevel (Quasi-)Monte Carlo algorithm using a geometric refinement $h_\ell = \varrho\, h_{\ell+1}$ is given by

$$\frac{cost(\bar{\mathcal{Q}}^{QMC}_{L,P,reuse})}{cost(\bar{\mathcal{Q}}^{QMC}_{L,P})} = 1 - \varrho^{-(\beta+\gamma)\lambda/(\lambda+1)}$$

This means that the sample recycling is more efficient when

- The variance of the difference decays slowly (small β)
- The lattice rule has good performance (small λ)

PART 3

Results for the elliptic model problem

Gaussian random field generation

- $a(\boldsymbol{x}, \boldsymbol{y})$ is derived from a **Gaussian random field** $z(\boldsymbol{x}, \boldsymbol{y})$ with given mean $z_0(\boldsymbol{x})$ and covariance function, e.g.,

$$\mathrm{cov}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} r\right)^{\nu} K_{\nu}\left(\sqrt{2\nu} r\right), \quad r = \frac{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2}{\lambda_c}$$

- Samples can be generated using a **KL expansion**

$$z(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j \geq 1} y_j \sqrt{\theta_j} \psi_j(\boldsymbol{x})$$

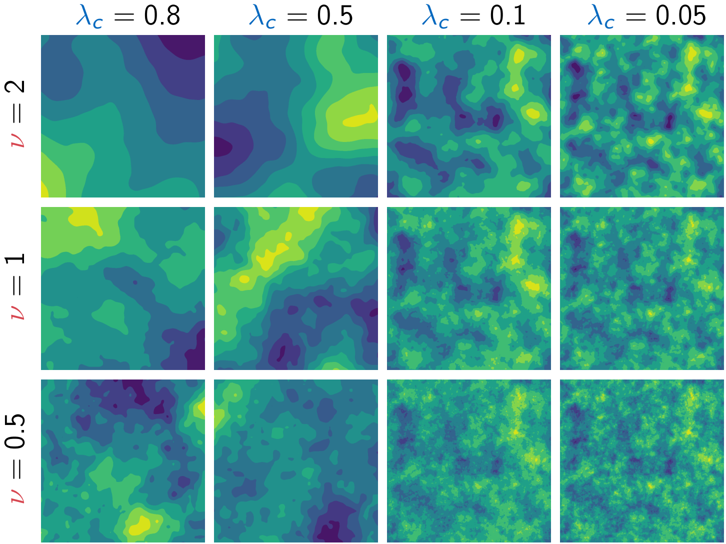
where the eigenvalues θ_j and eigenfunctions $\psi_j(\boldsymbol{x})$ satisfy

$$\int_D \mathrm{cov}(\boldsymbol{x}_1, \boldsymbol{x}_2) \psi_j(\boldsymbol{x}_2) \mathrm{d}\boldsymbol{x}_2 = \theta_j \psi_j(\boldsymbol{x}_1)$$

and the y_j are standard normal random numbers

- $a(\boldsymbol{x}, \boldsymbol{y}) = \exp(z(\boldsymbol{x}, \boldsymbol{y}))$ is known as the "*lognormal case*"

Example Gaussian random fields



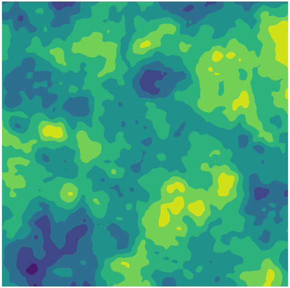
see GaussianRandomFields.jl

Numerical Results

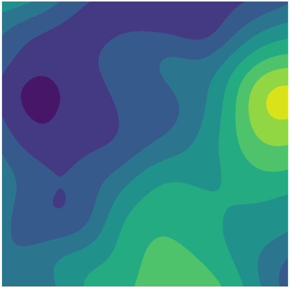
- PDE with **lognormal** diffusion coefficient defined on the unit cube $[0,1]^2$, homogenous Dirichlet boundary conditions, $f := 0$
  Quantity of interest is point evaluation at $\boldsymbol{x} = (0.5, 0.5)$
- GRF generated using a truncated KL expansion
  We consider $\lambda_c \in \{0.1, 0.3, 0.5\}$ and $\nu \in \{0.5, 1.0, 2.0\}$
- Discretized using finite differences on a rectangular grid
  Coarsest grid has 4^2 degrees of freedom
  Finest grid has 512^2 degrees of freedom
- Multigrid solver uses full weighting and bi-linear interpolation, 2 pre-smoothing and 1 post-smoothing steps and symmetric Gauss-Seidel as a smoother

Numerical Results

- Consider two different Gaussian random fields



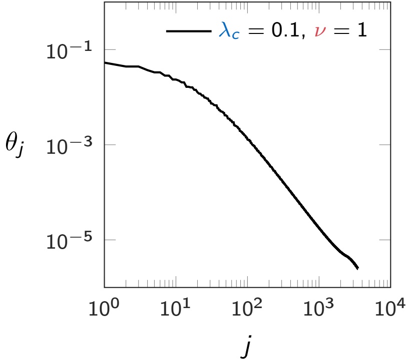
$\lambda_c = 0.1$, $\nu = 1$



$\lambda_c = 0.5$, $\nu = 2$

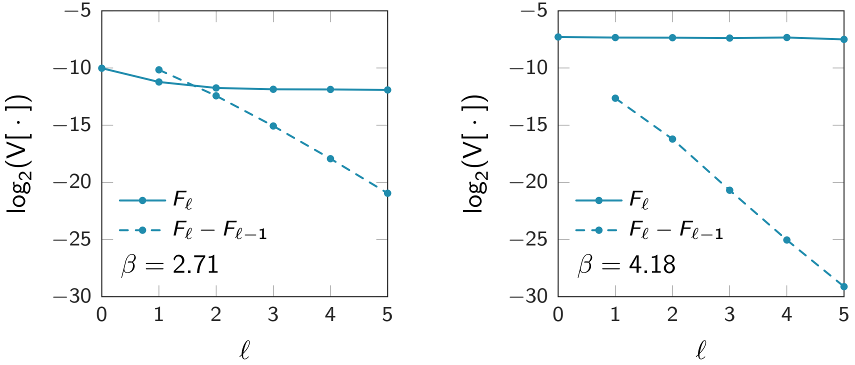
Numerical Results

- Decay of the eigenvalues in the KL expansion



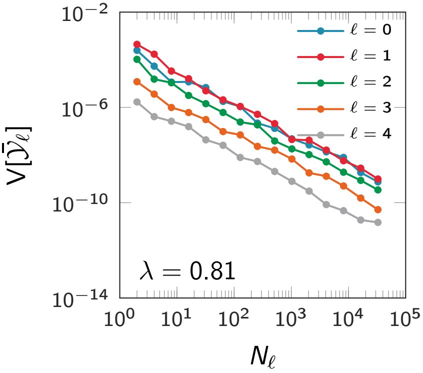
Numerical Results

- Rate of decrease in variance β



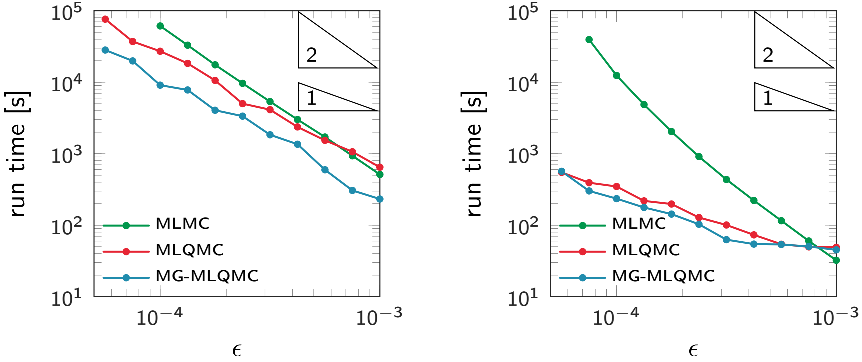
Numerical Results

- Convergence rate of the QMC sequence $1/\lambda$



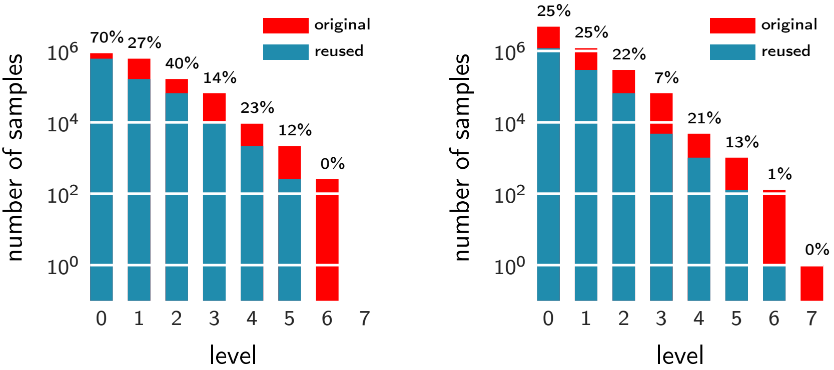
Numerical Results

- Run time comparsion with/without recycling



Numerical Results

- Sample distribution across the levels



Conclusions

- Multigrid + Multilevel (Quasi-)Monte Carlo is a **powerful** combination
- Recycling of coarse samples from FMG algorithm is possible without additional error by using **randomly shifted** lattice rules
- Sample recycling is more efficient when underlying problem is **nonsmooth**

Robbe, P., Nuyens, D., Vandewalle, S., Recycling Samples in the Multigrid Multilevel (Quasi-)Monte Carlo Method, ArXiv:1806.05619.

Note: recent work in [Detomasso, Dodwell, Scheichl, 2018]

Thank you for your attention