Scalable Hierarchical Sampling of Gaussian Random Fields for Large-Scale Multilevel Monte Carlo Simulations

Quantification of Uncertainty: Improving Efficiency and Technology

LLNL-PRES-734783

Sarah Osborn[1]
Panayot Vassilevski[1]
Umberto Villa[2]

[1] Center for Applied Scientific Computing, Lawrence Livermore National Laboratory
[2] Institute for Computational Engineering and Sciences (ICES), University of Texas at Austin

July 19, 2017

LLNL-PRES-XXXXXX
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Forward Propagation Uncertainty Quantification

- When modeling some physical phenomena, inputs are often subject to *uncertainty.*

- **Example**: Uncertainty in Groundwater Flow with Darcy's Law
  - Permeability coefficient k is subject to uncertainty.
  - Model k as a spatially correlated log-normal random field.
  - $k(x, \omega) = \exp[\theta(x, \omega)]$ where θ is a Gaussian random field with known mean and covariance.

- **Goal**: Given prior assumptions about uncertainty in input data, quantify uncertainty in the solution for *large-scale simulations* using Monte Carlo sampling methods.

Key Computational Challenges for Large-Scale Monte Carlo Sampling Methods

Many samples are necessary with a *fine* spatial discretization.

→

Multilevel Monte Carlo

- Use specialized element-based agglomeration technique to construct hierarchy.

Scalable generation of random input coefficient realizations

→

SPDE Sampling Technique

- Solve a stochastic PDE (SPDE) with mixed finite element method.
- Requires solution of *saddle point* problem with random right hand side.

Efficient solution of forward problem

→

Specialized preconditioners

- Discretization leads to matrices with *saddle point* structure.
- Employ methods from element-based multigrid.

Key Computational Challenges for Large-Scale Monte Carlo Sampling Methods

Many samples are necessary with a *fine* spatial discretization.

Multilevel Monte Carlo

- Use specialized element-based agglomeration technique to construct hierarchy.

Scalable generation of random input coefficient realizations

SPDE Sampling Technique

- Solve a stochastic PDE (SPDE) with mixed finite element method.
- Requires solution of *saddle point* problem with random right hand side.

Efficient solution of forward problem

Specialized preconditioners

- Discretization leads to matrices with *saddle point* structure.
- Employ methods from element-based multigrid.

Monte Carlo Method

- **Goal**: Estimate $\mathbb{E}[Q]$, the expected value of a quantity of interest $Q(\boldsymbol{X}(x,\omega))$ where $\boldsymbol{X}(x,\omega)$ is the solution of a PDE with random field coefficient.

$$\mathbb{E}[Q] \approx \widehat{Q}_h^{MC} = \frac{1}{N}\sum_{i=0}^{N} Q_h(\omega_i)$$

where $Q_h(\omega_i)$ is the i-th sample of Q approximated with spatial discretization h.

Mean Square Error (MSE) of method:

$$\underbrace{\frac{1}{N}\mathbb{V}[Q_h]}_{\text{Estimator Variance}} + \underbrace{\left(\mathbb{E}[Q - Q_h]\right)^2}_{\text{Bias: Discretization Error}}$$

Multilevel Monte Carlo Method

- This variance reduction technique uses a sequence of spatial approximations $Q_\ell, \ell = L, \dots, 1$ which approximate $Q_0 = Q_h$ with increasing accuracy (and cost).
- Linearity of expectation implies

$$\mathbb{E}[Q] \approx \mathbb{E}[Q_h] = \mathbb{E}[Q_L] + \sum_{\ell=0}^{L-1} \mathbb{E}[Q_\ell - Q_{\ell+1}].$$

The **multilevel MC estimator** is

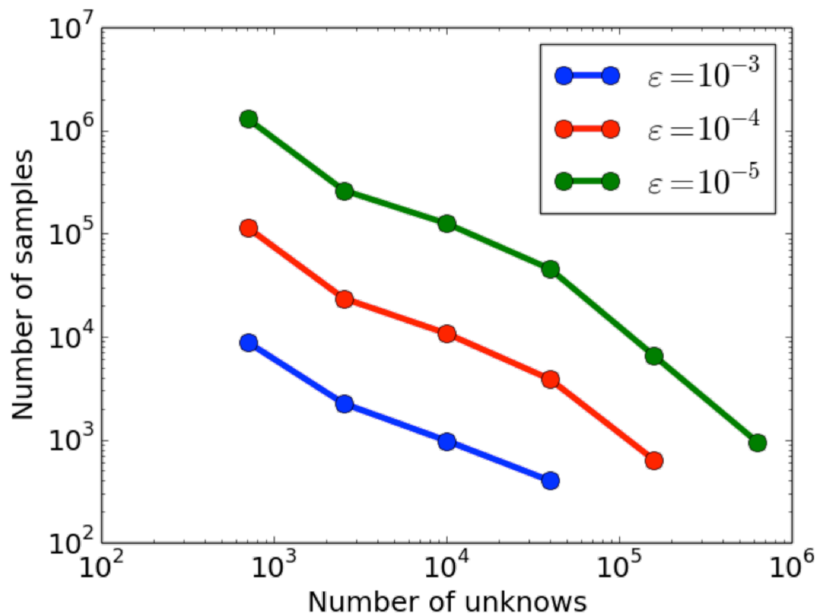
$$\widehat{Q}_h^{MLMC} = \frac{1}{N_L} \sum_{i=0}^{N_L} Q_L(\omega_i) + \sum_{i=0}^{L-1} \left[\frac{1}{N_\ell} \sum_{i=0}^{N_\ell} \left(Q_\ell(\omega_i) - Q_{\ell+1}(\omega_i)\right) \right].$$

M. **Giles**. Oper. Res. (2008)
K. **Cliffe**, M. **Giles**, R. **Scheichl**, and A. **Teckentrup**. Comput. Vis. Sci., (2011)

Multilevel Acceleration of Monte Carlo Method

- The MSE of the MLMC Estimator is

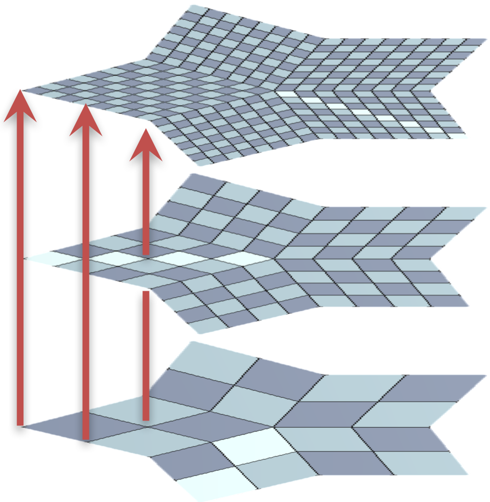
$$\underbrace{\frac{1}{N_L}\mathbb{V}[Q_L]}_{\text{Fixed cost independent of h}} + \underbrace{\sum_{\ell=1}^{L-1}\frac{1}{N_\ell}\mathbb{V}[Q_\ell - Q_{\ell+1}]}_{\mathbb{V}[Q_\ell - Q_{\ell+1}] \lll \mathbb{V}[Q_\ell]} + \underbrace{(\mathbb{E}[Q-Q_h])^2}_{\text{Discretization error}}$$

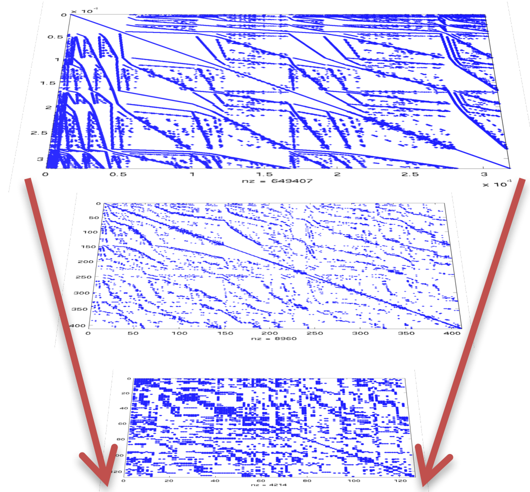


For a desired tolerance, the number of samples on each level is chosen to minimize the total computational cost.

Generation of Hierarchy of Spatial Discretizations with Element-Based Multigrid (AMGe)

Recall pros/cons of multigrid (MG) methods:



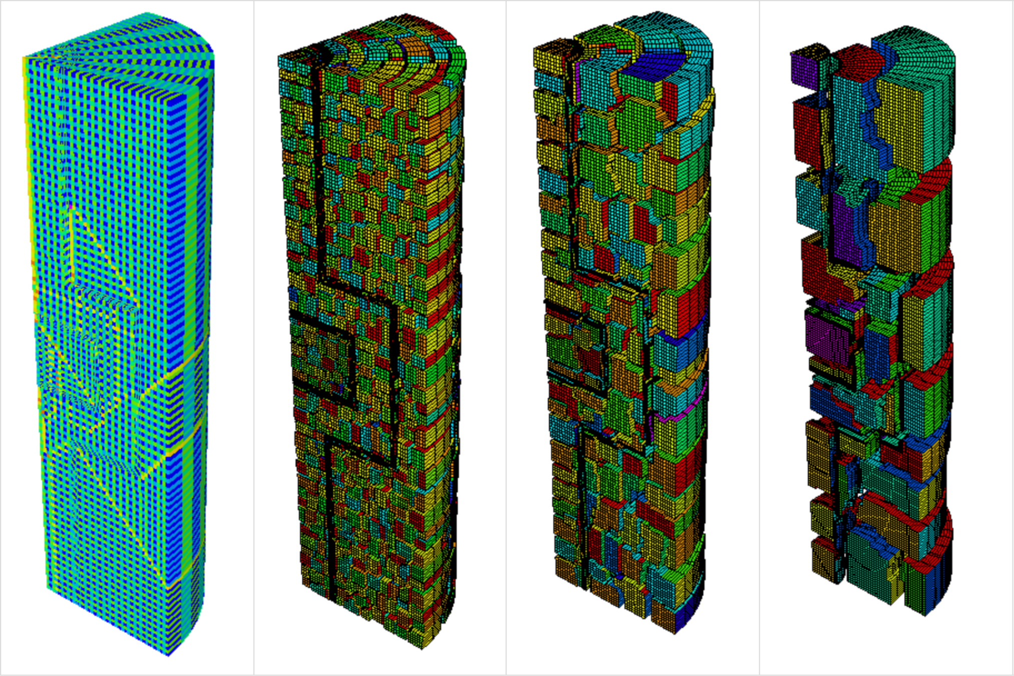


- Geometric Multigrid (GMG)
  - Scalable for many regular/semi-structured grid problem
  - Requires a nested hierarchy of grids
  - Uses information from discretization
  - Infeasible to implement for arbitrary unstructured-grid problems
- Algebraic Multigrid (AMG)
  - Optimal and effective solver for many PDEs on arbitrary grids
  - Requires only the fine-grid matrix; no spatial mesh needed
  - Closer to a black-box method

Element-based Multigrid (AMGe)

AMGe methods aim to leverage the advantages of the two approaches and to mitigate their shortcomings.

- GMG with nonstandard elements (agglomerates of fine-grid ones) and operator-dependent coarse finite element spaces.
- By using some "extra" information, AMGe can handle effectively a broader class of problems than classical AMG.
- **Coarse spaces have *guaranteed approximation properties.***



Hierarchy of agglomerated meshes

Coarsening de Rham Complexes on Agglomerated Elements

The de Rham complex plays an important role in analysis and discretization of PDEs.

$$
\begin{array}{ccccccc}
H^1 & \xrightarrow{\nabla} & H(\mathrm{curl}) & \xrightarrow{\nabla\times} & H(\mathrm{div}) & \xrightarrow{\nabla\cdot} & L^2 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
S_h & \xrightarrow{\nabla} & Q_h & \xrightarrow{\nabla\times} & R_h & \xrightarrow{\nabla\cdot} & W_h \\
\downarrow \Pi^S & & \downarrow \Pi^Q & & \downarrow \Pi^R & & \downarrow \Pi^W \\
S_H & \xrightarrow{\nabla} & Q_H & \xrightarrow{\nabla\times} & R_H & \xrightarrow{\nabla\cdot} & W_H
\end{array}
$$

- Generate a coarse sequence such that
  - The sequence is exact.
  - The spaces are conforming.
  - The commutativity property is preserved.
  - ***The approximation properties of the original spaces are preserved.***

J. **Pasciak**, P. **Vassilevski**. SISC. (2008)
I. **Lashuk**, P. **Vassilevski**. CMAM. (2011)

One hierarchy, many uses…..

The hierarchy of de Rham sequences with operator-dependent coarse spaces with approximation properties can be used for

- Robust multilevel preconditioners
- Discretization on a hierarchy of levels
  - Numerical upscaling
  - **Multilevel Monte Carlo simulations**
  - **Scalable Generation of Gaussian Random Fields**

ParELAG

A parallel distributed memory C++ library for an AMGe framework to coarsen a wide class of PDEs on general unstructured meshes developed at LLNL.

https://github.com/LLNL/parelag

Key Computational Challenges for Large-Scale Monte Carlo Sampling Methods

Many samples are necessary with a *fine* spatial discretization.

→

Multilevel Monte Carlo

- Use specialized element-based agglomeration technique to construct hierarchy.

Scalable generation of random input coefficient realizations

→

SPDE Sampling Technique

- Solve a stochastic PDE (SPDE) with mixed finite element method.
- Requires solution of *saddle point* problem with random right hand side.

Efficient solution of forward problem

→

Specialized preconditioners

- Discretization leads to matrices with *saddle point* structure.
- Employ methods from element-based multigrid.

Scalable Sampling of a Gaussian Random Field

Challenge: How to generate realizations of a Gaussian field (GF) on a hierarchy of spatial discretizations??

We consider a stationary isotrophic field with Matérn covariance function

$$\text{cov}(x,y) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa\,\|y-x\|)^\nu K_\nu(\kappa\,\|y-x\|) \text{ where } x,y \in \mathbb{R}^d.$$

Karhunen-Loève Expansion:

- Dense eigenvalue computation → Bottleneck
- State of the art methods exploit **Fast Multipole Methods** and **randomized eigensolvers** to alleviate this issue.

Scalable Sampling of a Gaussian Random Field

Gaussian Markov random field representation:

GFs with Matérn covariance functions are solutions of the *stochastic* PDE

$$(\kappa^2 - \Delta)^{\alpha/2}\theta(x,\omega) = g\mathcal{W}(x,\omega), \quad x \in \mathbb{R}^d, \alpha = \nu + \frac{d}{2}$$

- $\mathcal{W}(x,\omega)$: spatial Gaussian white noise with unit variance
- g: scaling factor to impose unit marginal variance
- $\kappa \in \mathbb{R}$: inversely proportional to correlation length

Special case:

In 3D, realizations of a Gaussian random field with exponential covariance function are solutions of the **stochastic reaction diffusion problem**.

F. **Lindgren**, H. **Rue,** J. **Lindstrom.** *J R Stat Soc Series B Stat Methodol*. (2011)

Stochastic PDE (SPDE) Sampler

Let $\nu = 1$ (2D) or $\nu = \frac{1}{2}$ (3D), then the realizations of the GF solve

$$(\kappa^2 - \Delta)\theta(x,\omega) = g\mathcal{W}(x,\omega), \quad x \in \mathbb{R}^d$$

Using the mixed finite element method, let

$\Theta_h \subset L^2(D)$ piecewise constant functions

$R_h \subset H(\mathrm{div}, D)$ lowest-order Raviart-Thomas elements

Find $(\mathbf{u}_h, \theta_h) \in (R_h, \Theta_h)$ such that

$$\begin{cases} (\mathbf{u}_h, \mathbf{v}_h) + (\theta_h, \mathrm{div}\, \mathbf{v}_h) = 0 & \forall \mathbf{v}_h \in R_h \\ (\mathrm{div}\, \mathbf{u}_h, q_h) - \kappa^2(\theta_h, q_h) = g(\mathcal{W}(\omega), q_h) & \forall q_h \in \Theta_h \end{cases}$$

with boundary conditions $\mathbf{u}_h \cdot \mathbf{n} = 0.$

Stochastic PDE (SPDE) Sampler

Noting that $\int_{D_i} \mathcal{W}(\omega) \sim \mathcal{N}(0, |D_i|)$ we obtain

$$\begin{bmatrix} M_h & B_h^T \\ B_h & -\kappa^2 W_h \end{bmatrix} \begin{bmatrix} u_h \\ \theta_h \end{bmatrix} = \begin{bmatrix} 0 \\ -g W_h^{\frac{1}{2}} \xi \end{bmatrix}, \quad \xi \sim \mathcal{N}(0, I)$$

where

- M_h is the mass matrix for the space R_h
- W_h is the (diagonal) mass matrix for space Θ_h
- B_h stems from the divergence constraint.

Able to leverage existing scalable solvers and preconditioners!

Hierarchical SPDE Sampler

- For MLMC, the same realization $\theta(\omega_i)$ must be computed at different spatial resolutions $\theta_h(\omega_i)$ (**fine**) and $\theta_H(\omega_i)$ (**coarse**).
- Recall the AMGe coarse spaces:

$$\Theta_H \subset \Theta_h \subset L^2(D) \quad \text{and} \quad R_H \subset R_h \subset H(\mathrm{div}, D)$$

- Define interpolation operators as

$$P_\theta : \Theta_H \to \Theta_h \text{ and } P_{\mathbf{u}} : R_H \to R_h$$

- Define the block interpolation operator as

$$\mathcal{P} = \begin{bmatrix} P_{\mathbf{u}} & 0 \\ 0 & P_\theta \end{bmatrix} \text{ so that } \mathcal{A}_H = \mathcal{P}^T \mathcal{A}_h \mathcal{P}.$$

Hierarchical SPDE Sampler

Then the Gaussian field θ_h admits the two-level decomposition

$$\theta_h(\omega) = P_\theta \theta_H(\omega) + \delta\theta_h(\omega),$$

where θ_H is a ***coarse representation of a Gaussian field from the same distribution***, and

$$\begin{bmatrix} \mathcal{A}_h & \mathcal{A}_h \mathcal{P} \\ \mathcal{P}^T \mathcal{A}_h & 0 \end{bmatrix} \begin{bmatrix} \delta\mathcal{U}_h \\ \mathcal{U}_H(\omega) \end{bmatrix} = \begin{bmatrix} \mathcal{F}_h \\ 0 \end{bmatrix},$$

where

$$\delta U_h = \begin{bmatrix} \delta \mathbf{u}_h \\ \delta\theta_h(\omega) \end{bmatrix}, \; U_H = \begin{bmatrix} \mathbf{u}_H \\ \theta_H(\omega) \end{bmatrix}, \text{ and } \mathcal{F}_h = \begin{bmatrix} 0 \\ -g W_h^{1/2} \xi_h(\omega) \end{bmatrix}.$$

Hierarchical SPDE Sampler: Numerical Solution

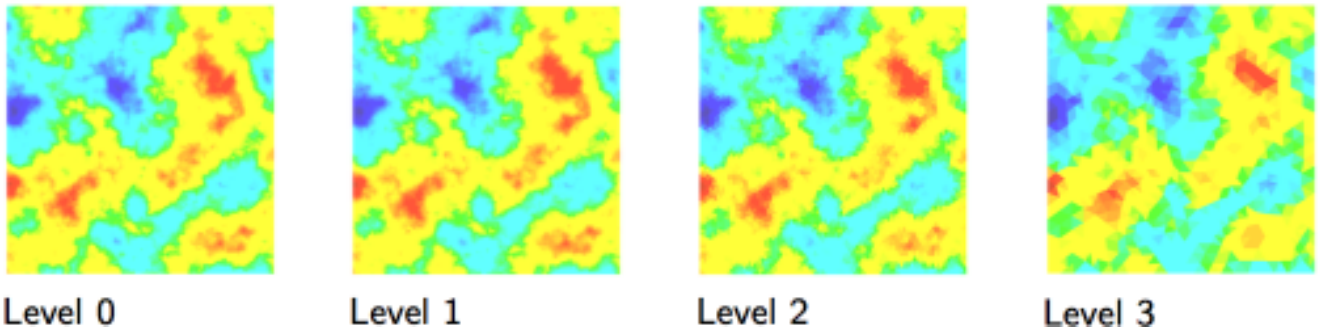
- Given $\xi_h(\omega_i)$, solve the saddle point system

$$\mathcal{A}_H \begin{bmatrix} \mathbf{u}_H \\ \theta_H(\omega_i) \end{bmatrix} = \mathcal{P}^T \begin{bmatrix} 0 \\ -gW_h^{1/2}\xi_h \end{bmatrix}, \xi_h \sim \mathcal{N}(0, I)$$

to generate $\theta_H(\omega_i)$ (coarse representation of $\theta_h(\omega_i)$ on Θ_H).

- Then solve $\mathcal{A}_h U_h = F_h$ with $\mathcal{P} U_H$ as the initial guess.

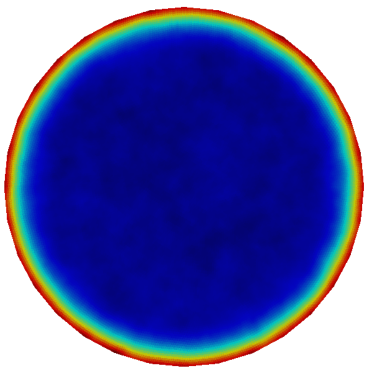
Sample realizations of Gaussian random field



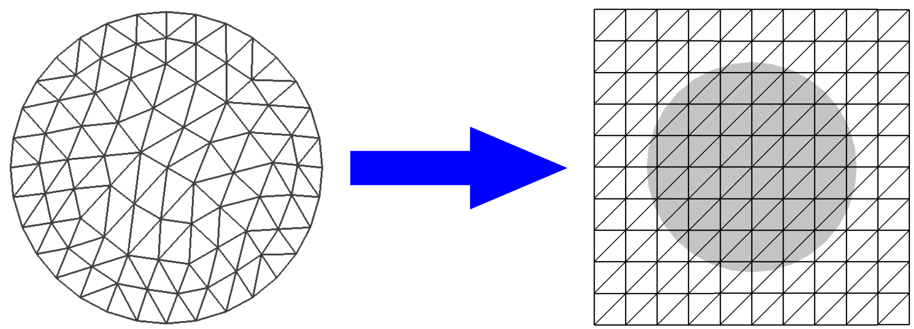
S.O., U. **Villa**, P. **Vassilevski**. *A multilevel, hierarchical sampling technique for spatially correlated random fields*. To appear SIAM SISC (2017).

Mesh Embedding with Non-Matching Meshes to Mitigate Artificial Boundary Effects

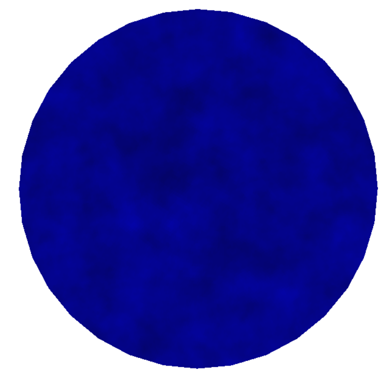
Sample marginal variance



Embed the original (unstructured) mesh in a regular, structured mesh.



Sample marginal variance with mesh embedding



- Solve SPDE on enlarged (structured) grid.
- Transfer the piecewise-constant solution to the original finite element space in parallel.
  - Meshes can be arbitrarily distributed!

S.O., P. **Zulian**,T. **Benson**, U. **Villa**, R. **Krause**, P. **Vassilevski**. *Scalable hierarchical PDE sampler for generating spatially correlated random fields using non-matching meshes*. Submitted (2017)

Model Problem: Uncertainty in Subsurface Flow

We solve the mixed Darcy equations

$$\begin{cases} k^{-1}\mathbf{q} + \nabla p = 0 & \text{in } D \\ \nabla \cdot \mathbf{q} = 0 & \text{in } D, \end{cases} \quad \Rightarrow \quad \begin{bmatrix} M_{k,h} & B_h^T \\ B_h & 0 \end{bmatrix} \begin{bmatrix} \mathbf{q}_h \\ p_h \end{bmatrix} = \begin{bmatrix} f_h \\ 0 \end{bmatrix}$$

where k is subject to uncertainty

with boundary conditions

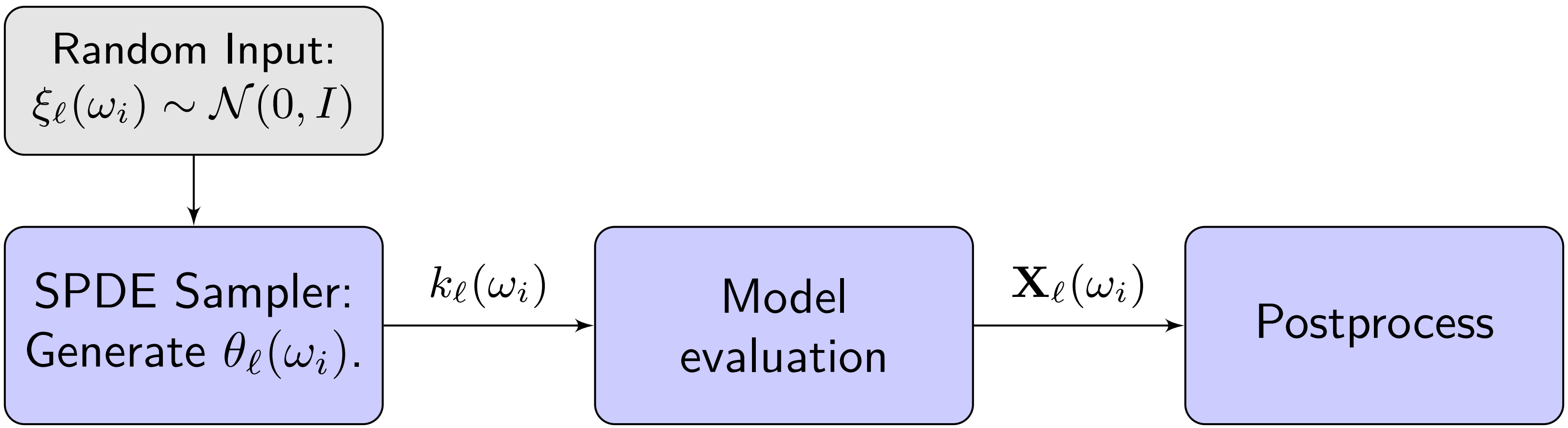
$\mathbf{q} \cdot \mathbf{n} = 0$ on Γ_N and $p = p_D$ on Γ_D.

Model k as a log-normal random field $k(x, \omega) = \exp[\theta(x, \omega)]$ where θ where is a Gaussian field with Matérn covariance function.

Multilevel Monte Carlo Simulation Workflow

MLMC Estimator: $\widehat{Q}_h^{MLMC} = \sum_{\ell=0}^{L} \widehat{(Q_\ell - Q_{\ell+1})}^{MC}$ where $\widehat{Q}_h^{MC} = \frac{1}{N}\sum_{i=1}^{N} Q_h(\omega_i)$

To generate a sample on level ℓ:



Solve saddle point problem on level ℓ of structured hierarchy. Compute $k_\ell(\omega_i) = \exp[\theta_\ell(\omega_i)]$ and transfer to original FE space.

Solve forward model problem on level ℓ of original, unstructured hierarchy.

Compute quantity of interest $Q_\ell(\mathbf{X}_\ell(\omega_i))$.

Key Computational Challenges for Large-Scale Monte Carlo Sampling Methods

Many samples are necessary with a *fine* spatial discretization.

→

Multilevel Monte Carlo

- Use specialized element-based agglomeration technique to construct hierarchy.

Scalable generation of random input coefficient realizations

→

SPDE Sampling Technique

- Solve a stochastic PDE (SPDE) with mixed finite element method.
- Requires solution of *saddle point* problem with random right hand side.

Efficient solution of forward problem

→

Specialized preconditioners

- Discretization leads to matrices with *saddle point* structure.
- Employ methods from element-based multigrid.

Efficient Solvers for Saddle Point Problems

- We need to solve a large, sparse saddle point system of the form:

$$\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

(where C=0 for mixed Darcy eqns)

- Possible preconditioning strategies:
  - Block factorization preconditioners:
    - Build MG-based approximations for A^{-1} and inverse of approximate Schur-complement where $S = -C - B\mathrm{diag}(A)^{-1}B^T$.
  - Monolithic AMGe preconditioners
    - Treat whole system simultaneously with one MG method.
    - Blocked grid transfers from de Rham sequence.

Numerical Results: Implementation and Solver Specifics for SPDE Sampler and Forward Problem

ParELAG

MFEM: *scalable* C++ library for finite element methods

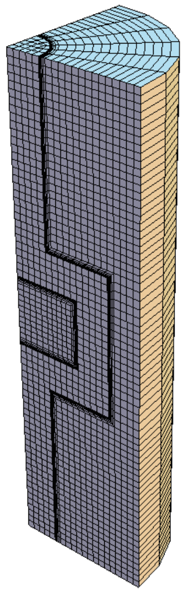
hypre

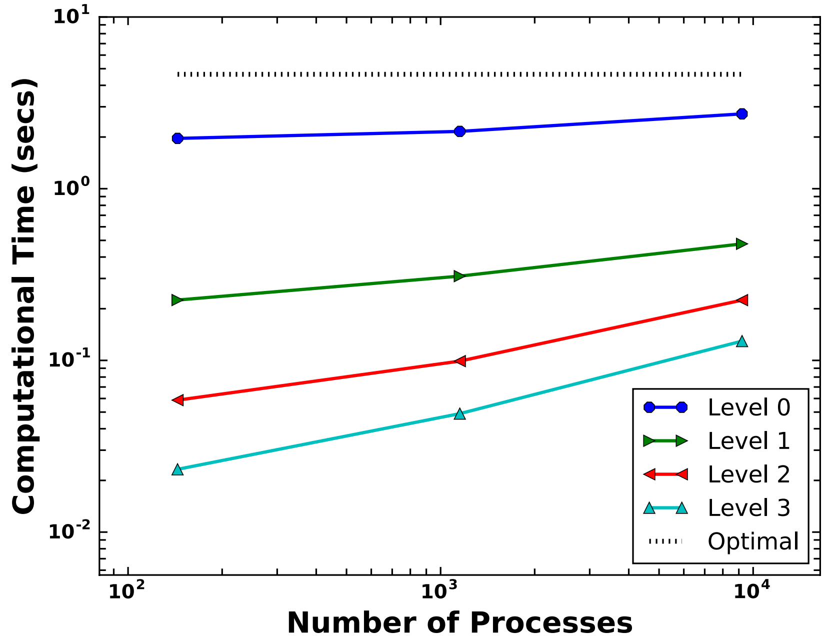
Scalable linear solvers and multigrid methods

- Solve saddle point systems with preconditioned GMRES:
  - Monolithic AMGe:
    - Block LDU smoother using a single sweep of point Gauss-Seidel to approximate A^{-1}.
    - Blocked grid transfers from hierarchy of de Rham sequence.
  - Block + AMGe:
    - A^{-1} approximated by a single AMGe V-cycle using a sweep of point Gauss-Seidel as a smoother.
  - Block + GS:
    - A^{-1} approximated by a single sweep of point Gauss-Seidel

S^{-1} is approximated by a single BoomerAMG V-cycle for each preconditioner.

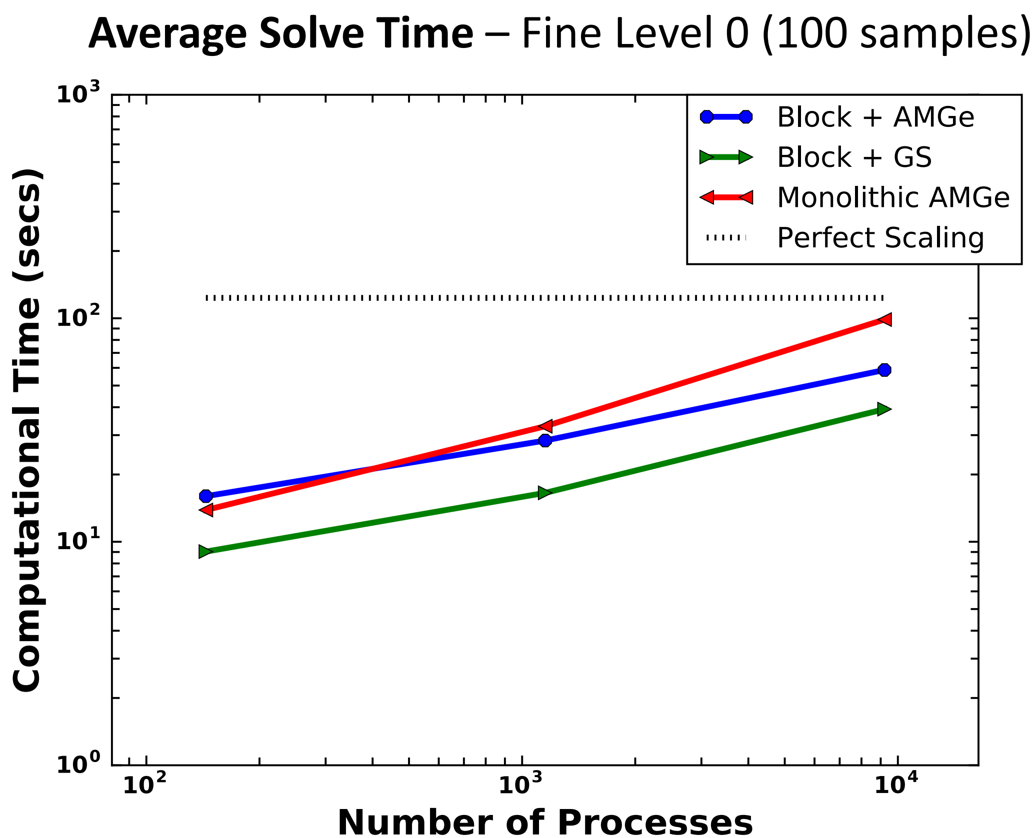
Weak Scaling of SPDE Sampler: Crooked Pipe Problem





- Finite element level (Level = 0) has ≈51K stochastic dofs per process, largest problem has approximately $4.7\text{x}10^{8}$ stochastic degrees of freedom.
- The saddle point system is solved with GMRES preconditioned with '**Monolithic AMGe**'.

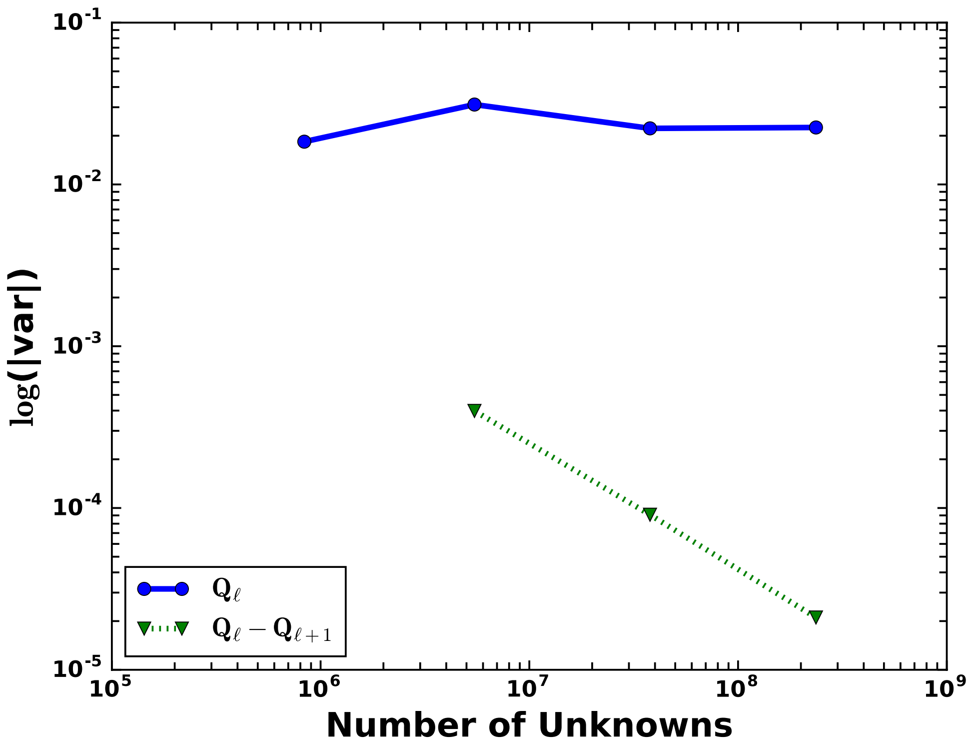
Weak Scaling of Mixed Darcy Equations with Random Permeability: Crooked Pipe Problem



Finite element level (Level = 0) has ≈209K velocity/pressure dofs per process, largest problem has ≈ $1.9\text{x}10^9$ dofs.

Multilevel Variance Reduction: Crooked Pipe Problem

MLMC Simulation with hierarchical SPDE sampler with non-matching mesh embedding



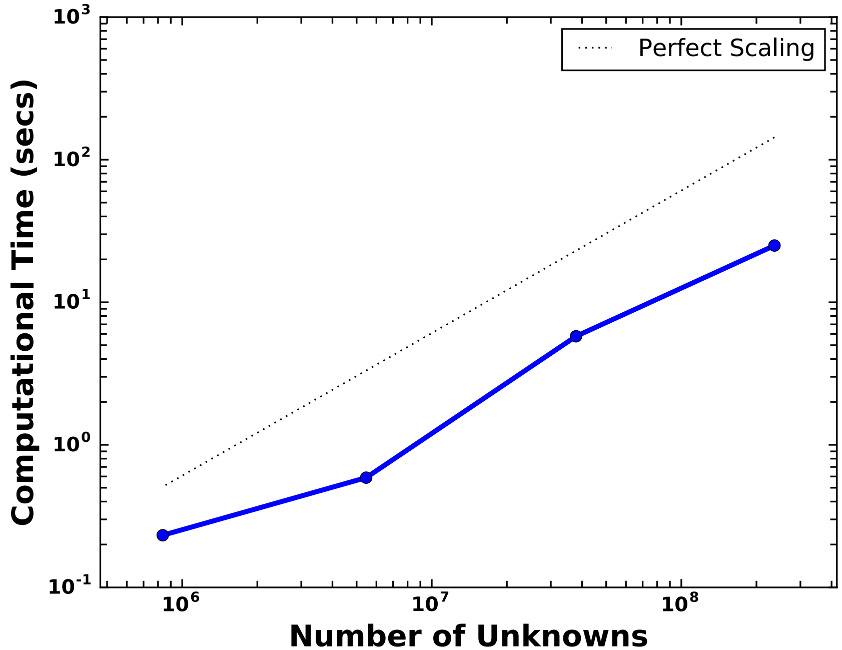
The QoI is the effective permeability given by

$$\frac{1}{|\Gamma_{out}|} \int_{\Gamma_{out}} \mathbf{q}(\cdot, \omega) \cdot \mathbf{n} \, dS.$$

MLMC Performance: Crooked Pipe Problem

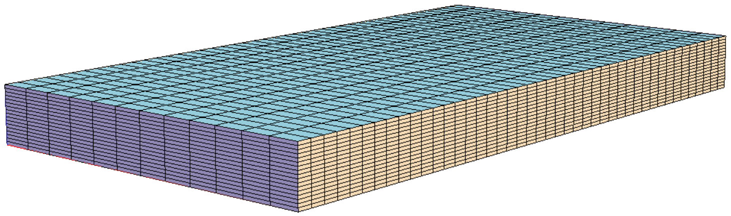
- MSE: $\epsilon^2 = 2.5e^{-5}$
- 240M velocity/pressure unknowns on fine level
- 59M stochastic dimensions
- 1.2K processors/sample generation
- Preconditioner:
  - Sampler: Monolithic AMGe
  - Darcy: Block + GS

Average time to compute a sample $Q_\ell(\omega_i) - Q_{\ell+1}(\omega_i)$

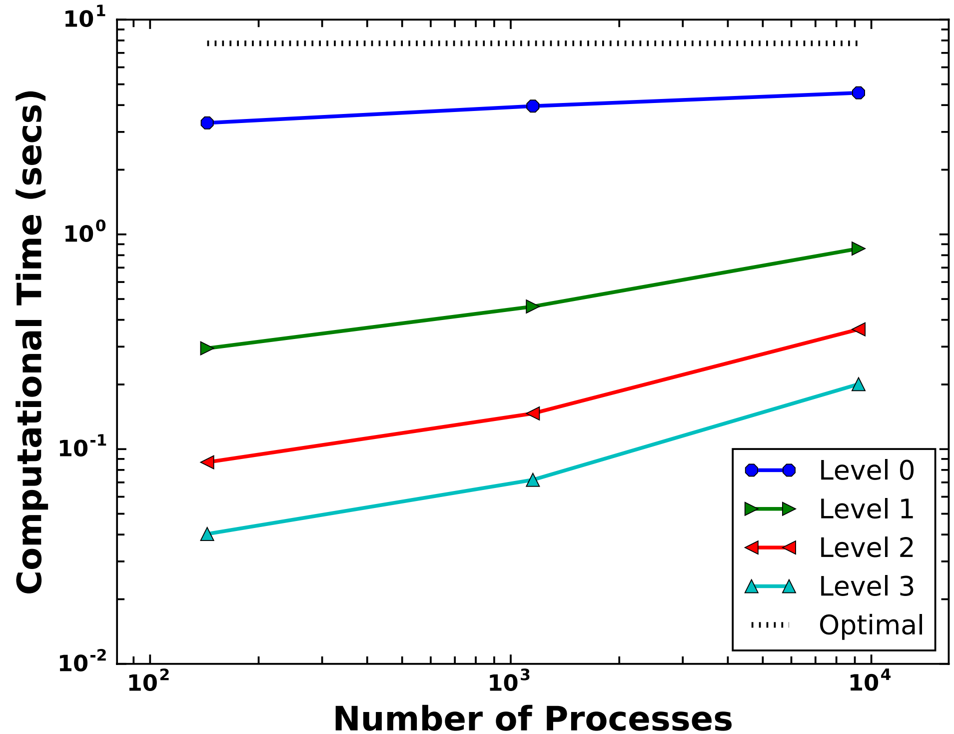


	MC (estimated)	MLMC
N_0	1799	12
Total samples	1799	3147
Wall Time	12.2 hours	0.4 hours

Weak Scaling of SPDE Sampler: SPE10 Problem



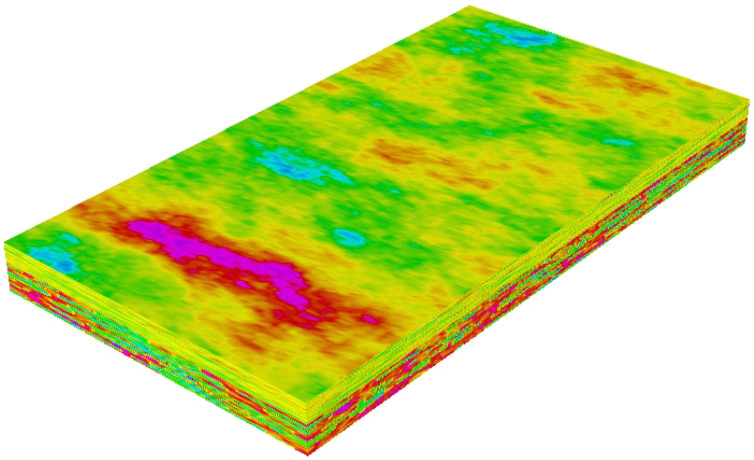
SPE10 model:
1200x2200x170(ft) regular cartesian grid (highly stretched elements)



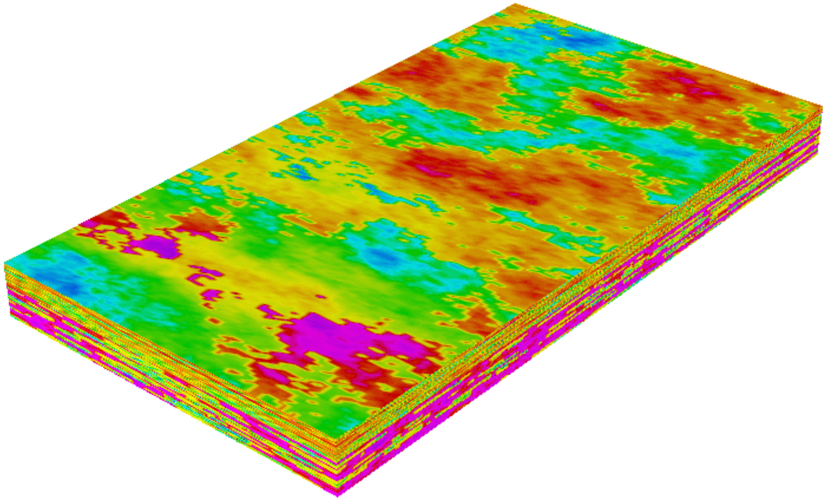
- Finite element level (Level = 0) has ≈ 32K stochastic dofs per process, largest problem has approximately $2.9\text{x}10^{8}$ stochastic dofs.
- Solver: GMRES preconditioned with '**Monolithic AMGe**'

MLMC for SPE10 Problem

Random permeability coefficient $k(x,\omega)$ is modeled as log-normal random field where $\exp[\log[k_{SPE10}\ (x)] + \theta(x,\omega)]$.



x/y component

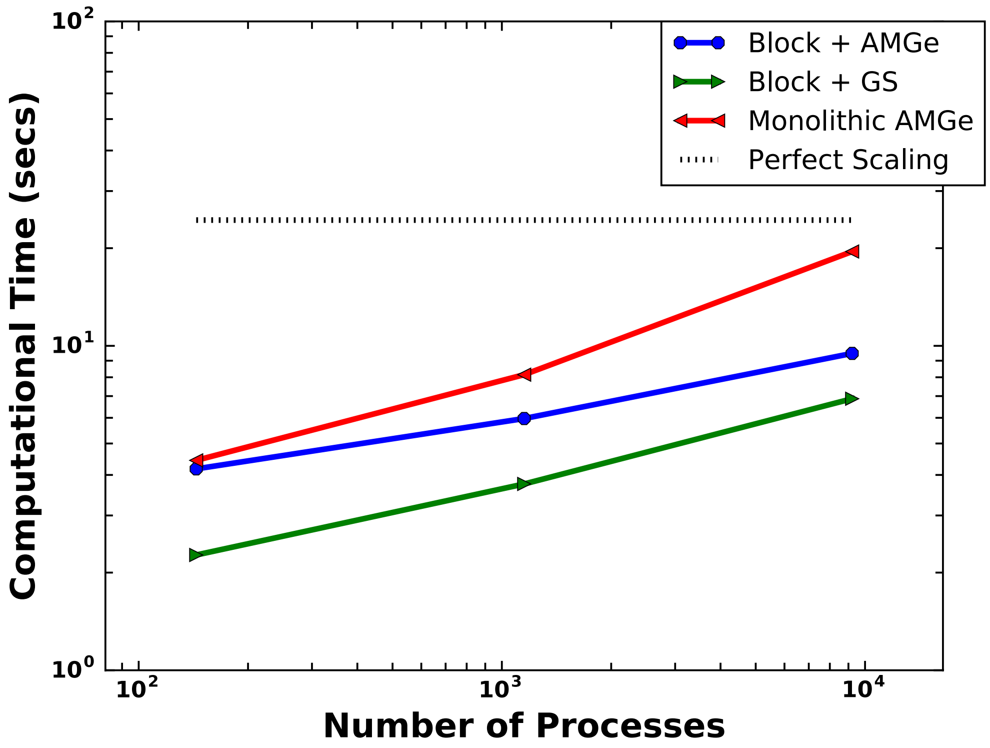


z component

Logarithmic plots of relative permeability coefficient from SPE10 dataset which has large jumps between the mesh elements.

Weak Scaling of Mixed Darcy Equations with Random Permeability: SPE10 Problem

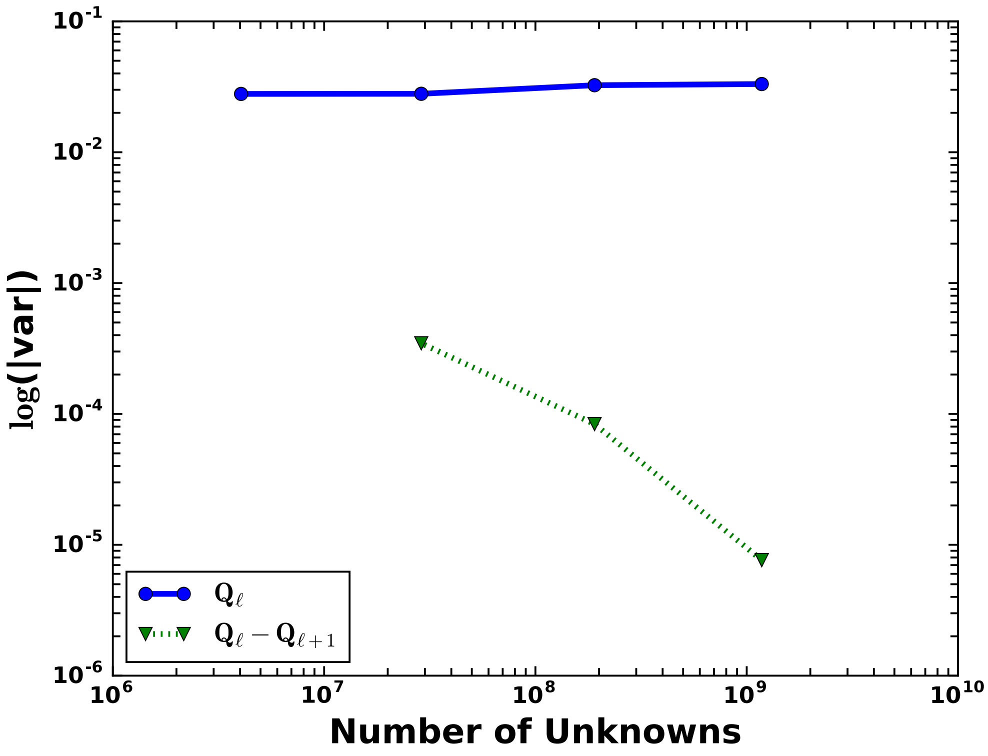
Average Solve Time – Fine Level 0 (100 samples)



Finite element level (Level = 0) has ≈130K velocity/pressure dofs per process, largest problem has ≈ $1.2\text{x}10^9$ dofs.

Multilevel Variance Reduction: SPE10 Problem

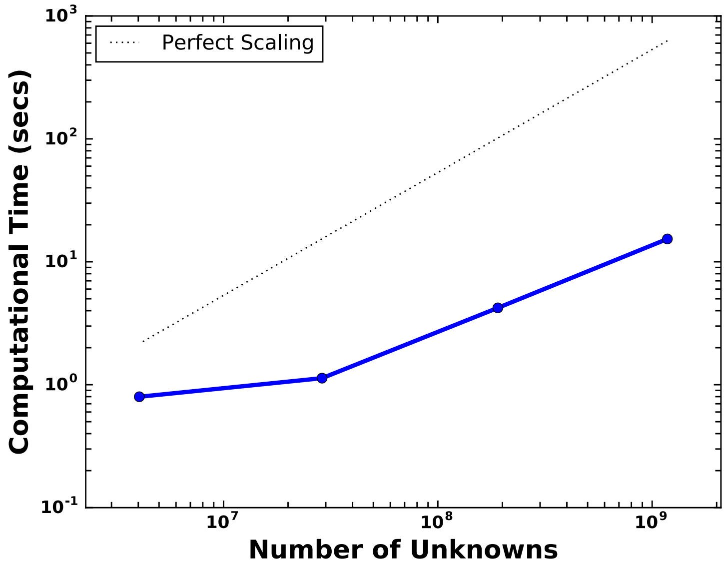
MLMC Simulation with hierarchical SPDE sampler with non-matching mesh embedding



The QoI is $p(x^*)$ for $x^* = (600,1100,85)$

MLMC Performance: SPE10 Problem

- MSE $\epsilon^2 = 6.25e^{-6}$
- 1.2B velocity/pressure unknowns on fine level
- 443M stochastic dimensions



Average time to compute a sample $Q_\ell(\omega_i) - Q_{\ell+1}(\omega_i)$.

- 9K processors/sample generation
- Preconditioner:
  - Sampler: Monolithic AMGe
  - Darcy: Block + GS

	MC (estimated)	MLMC
N_0	10623	42
Total samples	10623	13690
Wall Time	41.9 hours	3.9 hours

MLMC with SPDE sampling makes large-scale Monte Carlo simulations feasible!!

Concluding Remarks

- Scalable sampling of Gaussian random fields is necessary for large-scale uncertainty quantification simulations.
  - **Proposed Solution**: Hierarchical SPDE sampler
  - Sampling strategy is based on solving a mixed discretization of stochastic PDE.
  - Use mesh embedding on non-matching meshes to mitigate artificial boundary effects with scalable transfer of data between meshes.

- Successfully applied the new sampling technique to *large-scale* MLMC simulations of subsurface flow problems.
  - Constructed hierarchy of coarse spaces using specialized element-based agglomeration techniques.
  - Able to leverage specialized preconditioners for saddle point problems.

- **Future Work/Remarks**:
  - Only leveraging parallelism in spatial dimension.
  - Further parallelism possible within and across levels as investigated by B. B. **Gmeiner**, D. **Drzisga**, U. **Rude**, R. **Scheichl**, B. **Wohlmuth** (2016).

Lawrence Livermore National Laboratory