A Hierarchical Acceleration of a Stochastic Collocation Method for PDEs with Random Input Data

Guannan Zhang

Computer Science and Mathematics Division – Oak Ridge National Laboratory

Supported by DOE, ASCR via Householder Fellowship and EQUINOX

Collaborators

Max Gunzburger Florida State University

Clayton Webster, Oak Ridge National Laboratory

July 8, 2014

Outline

1. Stochastic problems: PDEs with random inputs
2. Stochastic non-intrusive methods (MC, NIPC, SC)
3. Piecewise adaptive sparse-grid approximation with hierarchical acceleration
4. Error estimate and complexity analyses
5. A numerical example
6. Concluding remarks

Partial differential equations with random input data

A simplified general (stationary) setting

Consider an operator $\mathcal{L}$, linear or nonlinear, on a domain $D \subset \mathbb{R}^d$, which depends on some coefficients $a(\omega, x)$ with $x \in D$, $\omega \in \Omega$ and $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space. The forcing $f = f(\omega, x)$ and the solution $u = u(\omega, x)$ are random fields s.t.

$$\mathcal{L}(a)(u) = f \quad \text{a.e. in } D$$

equipped with suitable boundary conditions.

A_1. the solution has realizations in the Banach space $W(D)$, i.e. $u(\cdot, \omega) \in W(D)$ almost surely

$$\|u(\cdot, \omega)\|_{W(D)} \leq C \|f(\cdot, \omega)\|_{W^*(D)}$$

A_2. the forcing term $f \in L^2_{\mathbb{P}}(\Omega; W^*(D))$ is such that the solution u is unique and bounded in $L^2_{\mathbb{P}}(\Omega; W(D))$

A_3. $\mathbb{P}\left\{a(\omega, x) \in (a_{\min}, a_{\max}), \ \ \forall x \in \overline{D}\right\} = 1$ with $0 < a_{\min} < a_{\max} < \infty$

Examples

Linear and nonlinear elliptic PDEs with random input data

Example: The linear elliptic problem

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

with $f(\omega,\cdot)$ square integrable with respect to $\mathbb{P}$, satisfies assumptions A_1, A_2 and A_3 with $W(D) = H_0^1(D)$

Example: The nonlinear elliptic problem

Similarly, for $k \in \mathbb{N}^+$,

$$\begin{cases} -\nabla \cdot (a(\omega,\cdot)\nabla u(\omega,\cdot)) + u(\omega,\cdot)|u(\omega,\cdot)|^k &= f(\omega,\cdot) \quad \text{in } \Omega \times D, \\ u(\omega,\cdot) &= 0 \quad \text{on } \Omega \times \partial D, \end{cases}$$

satisfies assumptions A_1, A_2 and A_3 with $W(D) = H_0^1(D) \cap L^{k+2}(D)$

Parameterization of random fields

Finite dimensional noise assumption

WLOG assume the random fields $a(\omega, x)$ and $f(\omega, x)$ depends on a finite number of random variables $\boldsymbol{y}(\omega) = [y_1(\omega), \ldots, y_N(\omega)] : \Omega \to \mathbb{R}^N$ s.t.

$$a(\omega, x) = a(\boldsymbol{y}(\omega), x), \quad f(\omega, x) = f(\boldsymbol{y}(\omega), x)$$

- Piecewise constant approximations: let $\{D_n\}_{n=1}^N$ be a partition of D then define $a(\omega, x) = \sum_{n=1}^N \sigma_n y_n(\omega) \chi_{D_n}(x)$
- Truncated infinite dimensional expansions: e.g. expand $a(\omega, x)$ in a Karhunen-Loève expansion and retain the first N terms.

- Given $a(\boldsymbol{y}(\omega), x)$, $f(\boldsymbol{y}(\omega), x) \Rightarrow u(\omega, x) = u(y_1(\omega), \ldots, y_N(\omega), x)$ s.t.

$$\mathcal{L}(a(\boldsymbol{y}, x))(u(\boldsymbol{y}, x)) = f(\boldsymbol{y}, x) \quad \text{in } D \text{ a.s.}$$

- $\boldsymbol{y}(\omega)$ has a joint PDF $\rho : \Gamma \equiv \prod_{n=1}^N \Gamma_n \subset \mathbb{R}^N \to \mathbb{R}_+$, with $\rho \in L^\infty(\Gamma)$
- The probability space $(\Omega, \mathcal{F}, P)$ is mapped to $(\Gamma, \mathcal{B}(\Gamma), \rho(\boldsymbol{y})d\boldsymbol{y})$, where $\mathcal{B}(\Gamma)$ denotes the Borel σ-algebra on Γ and $\rho(\boldsymbol{y})d\boldsymbol{y}$ is the probability measure of $\boldsymbol{y}$

Stochastic non-intrusive methods

Multivariate approximations of $u(\boldsymbol{y}, x)$ in $\Gamma \times D$

The general procedure of a non-intrusive approach of parametric PDEs is

- Choose a set of sample points $\mathcal{H}_M(\Gamma) = \{\boldsymbol{y}_k \in \Gamma\}_{k=1}^M$
- Solve the parametric PDE system to obtain $u_{N_h}(\boldsymbol{y}_k)$ for $k = 1, \ldots, M$ by using deterministic solvers with mesh size h
- Construct a multivariate approximation with the sampled values:

$$u_{N_h,M}(\boldsymbol{y}, x) = \sum_{k=1}^{M} c_k(x) \cdot \psi_k(\boldsymbol{y}),$$

where, for $k = 1, \ldots, M$, $u_{N_h,M}(\boldsymbol{y}_k, x) = u_{N_h}(\boldsymbol{y}_k, x)$, $\psi_k(\boldsymbol{y}) \in \mathcal{P}(\Gamma)$ are basis functions which expand the polynomial space $\mathcal{P}(\Gamma)$, and the coefficients $c_k(x)$ are determined by the samples $\{u_{N_h}(\boldsymbol{y}_k, x)\}_{k=1}^M$

Quantity of interest, e.g. $\mathbb{E}[u](x)$

$$\mathbb{E}[u](x) \approx \int_\Gamma u_{N_h,M}(\boldsymbol{y}, x)\rho(\boldsymbol{y})d\boldsymbol{y} = \sum_{k=1}^{M} c_k(x) \int_\Gamma \psi_k(y)\rho(\boldsymbol{y})dy$$

Examples of non-intrusive methods

- Monte Carlo and quasi Monte Carlo methods
  - low regularity requirement of $u(\boldsymbol{y}, x)$
  - slow convergence, e.g. $\mathcal{O}(M^{1/2})$
- Non-intrusive polynomial chaos expansions
  - high regularity requirement of $u(\boldsymbol{y}, x)$
  - orthogonal polynomial basis
  - fast convergence rates
- Global (sparse-grid) stochastic collocation methods
  - high regularity requirement of $u(\boldsymbol{y}, x)$
  - fast convergence rates
  - dimensional adaptivity
- Piecewise hierarchical (sparse-grid) stochastic collocation methods
  - low regularity requirement of $u(\boldsymbol{y}, x)$
  - faster convergence than Monte Carlo methods
  - local adaptivity

Hierarchical stochastic collocation methods

Complexity reduction through solution acceleration

At level $L \in \mathbb{N}_+$, a hierarchical SC approximation is (informally) defined by:

$$u_{N_h,M_L} = \mathcal{I}_L[u_{N_h}] \equiv \mathcal{I}_{L-1}[u_{N_h}] + \Delta_L[u_{N_h}]$$

- $\mathcal{I}_{L-1}[u_{N_h}]$ is the hierarchical interpolant at level $L-1$, and $\Delta_L[u_{N_h}]$ is the corresponding hierarchical incremental interpolant at level L
- $\Delta_L[u_{N_h}]$ interpolates on the difference grid $\Delta\mathcal{H}_{M_L} = \mathcal{H}_{M_L} \backslash \mathcal{H}_{M_{L-1}}$, s.t. $\Delta_L[u_{N_h}] \to 0$ as $L \to \infty$

Relationship to multilevel methods (e.g., MLMC, MLSC): reduce complexity by balancing errors across a sequence of spatial approximations [TJWG2014]

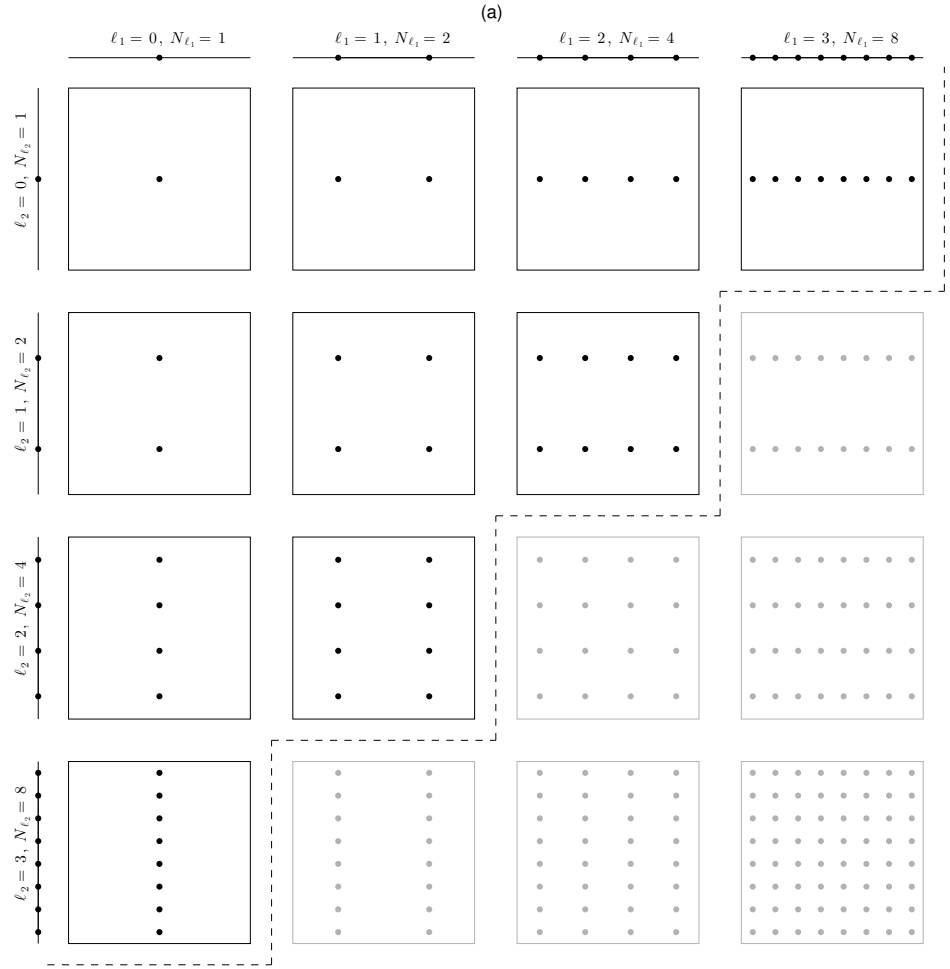
$$u_K^{(\mathrm{ML})} := \sum_{k=0}^{K} \mathcal{I}_{L_{K-k}} \left[u_{N_{h_k}} - u_{N_{h_{k-1}}} \right]$$

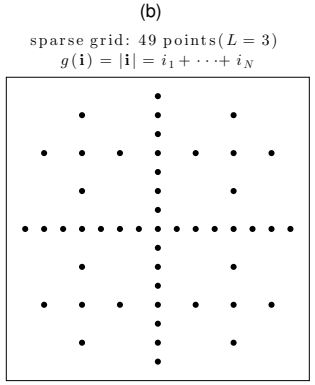
Local hierarchical sparse-grid approximation of $u(\boldsymbol{y}, x)$

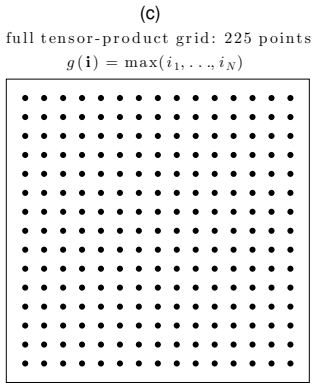
Sparse grid construction

The hierarchical sparse-grid interpolant $u_{N_h,M_L}(\boldsymbol{y}, x)$ is defined by

$$u_{N_h,M_L}(\boldsymbol{y}, x) = \sum_{l=0}^{L} \sum_{g(\mathbf{l})=l} \sum_{\mathbf{i}\in B_{\mathbf{l}}} c_{\mathbf{l},\mathbf{i}}(x) \cdot \psi_{\mathbf{l},\mathbf{i}}(\boldsymbol{y})$$







- $\psi_{\mathbf{l},\mathbf{i}}(\boldsymbol{y}) = \prod_{n=1}^{N} \psi_{l_n,i_n}(y_n)$
- $c_{\mathbf{l},\mathbf{i}}(x)$ is the surplus
- $B_{\mathbf{l}}$ is the index set of the "block" $\mathbf{l} = (l_1, \ldots, l_N)$
- L indicates the resolution level of $u_{N_h,M_L}(\boldsymbol{y}, x)$ in Γ
- Full tensor-product grid: $g(\mathbf{l}) = \max_{n=1,\ldots,N} l_n$
- Isotropic sparse grid: $g(\mathbf{l}) = l_1 + \cdots + l_N$

Construction of the surpluses $c_{\mathbf{l},\mathbf{i}}(x)$ in $u_{N_h,M_L}(\boldsymbol{y},x)$

- The surpluses are given in terms of the spatial finite element basis $\{\phi_j(x)\}_{j=1}^{N_h}$ by $c_{\mathbf{l},\mathbf{i}}(x) = \sum_{j=1}^{N_h} c_{j,\mathbf{l},\mathbf{i}} \cdot \phi_j(x)$ such that

$$u_{N_h,M_L}(\boldsymbol{y},x) = \sum_{j=1}^{N_h} \left(\sum_{l=0}^{L} \sum_{|\mathbf{l}|=l} \sum_{\mathbf{i}\in B_{\mathbf{l}}} c_{j,\mathbf{l},\mathbf{i}} \cdot \psi_{\mathbf{l},\mathbf{i}}(\boldsymbol{y}) \right) \phi_j(x)$$

- For $c_{\mathbf{l},\mathbf{i}}(x)$ with $\mathbf{l} = L$, the hierarchical interpolatory property provides us

$$c_{j,\mathbf{l},\mathbf{i}} = u_{N_h}(\boldsymbol{y}_{\mathbf{l},\mathbf{i}}, x_j) - u_{N_h,M_{L-1}}(\boldsymbol{y}_{\mathbf{l},\mathbf{i}}, x_j) \quad \text{for} \quad j = 1,\ldots,N_h$$

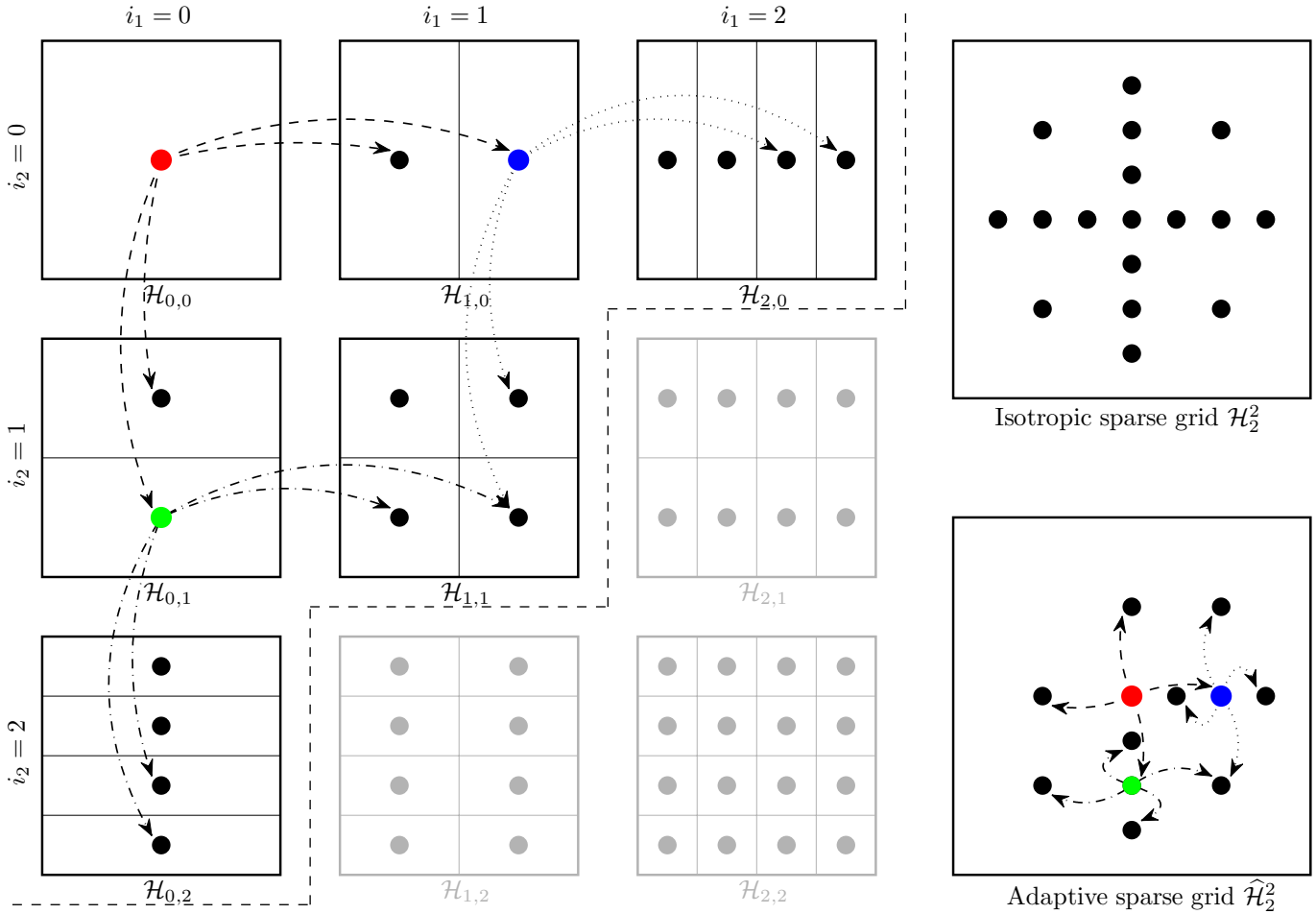
- When $u(\boldsymbol{y},x)$ has bounded mixed second-order weak derivatives with respect to $\boldsymbol{y}$, the surpluses can be bounded as

$$|c_{j,\mathbf{l},\mathbf{i}}| \le C 2^{-2\cdot|\mathbf{l}|} \quad \text{for} \quad \mathbf{i} \in B_{\mathbf{l}} \text{ and } j = 1,\ldots,N_h$$

- The hierarchical structure, especially the decay of surpluses, immediately leads to a strategy to do adaptivity

Two-dimensional adaptive sparse grid

Level $0, 1, 2$ sparse grids with $l_1 + l_2 \leq 2$



Each point that corresponds to a large surplus, e.g., the points in red, blue, or green, lead to 2 children points added in each direction resulting in the adaptive sparse grid

Computational cost of stochastic collocation methods

for complex PDE systems

- At each collocation point $\boldsymbol{y}_{\mathbf{l},\mathbf{i}}$, $u_{N_h}(x, \boldsymbol{y}_{\mathbf{l},\mathbf{i}})$ is approximated based on the solution from the selected linear system solver, i.e.

$$u_{N_h}(x, \boldsymbol{y}_{\mathbf{l},\mathbf{i}}) = \sum_{j=1}^{N_h} u_{j,\mathbf{l},\mathbf{i}} \phi_j(x) \approx \widetilde{u}_{N_h}(x, \boldsymbol{y}_{\mathbf{l},\mathbf{i}}) = \sum_{j=1}^{N_h} \widetilde{u}_{j,\mathbf{l},\mathbf{i}} \phi_j(x)$$

  where $\widetilde{\boldsymbol{u}}_{\mathbf{l},\mathbf{i}} = (\widetilde{u}_{1,\mathbf{l},\mathbf{i}}, \ldots, \widetilde{u}_{N_h,\mathbf{l},\mathbf{i}})^\top$ is the output of the solver.

- In the case of using conjugate gradient methods, the error $\boldsymbol{e}^k_{\mathbf{l},\mathbf{i}} = \boldsymbol{u}_{\mathbf{l},\mathbf{i}} - \boldsymbol{u}^k_{\mathbf{l},\mathbf{i}}$ is bounded by

$$\|\boldsymbol{e}^k_{\mathbf{l},\mathbf{i}}\|_{\boldsymbol{A}_{\mathbf{l},\mathbf{i}}} \leq 2 \left(\frac{\sqrt{\kappa_{\mathbf{l},\mathbf{i}}} - 1}{\sqrt{\kappa_{\mathbf{l},\mathbf{i}}} + 1} \right)^k \|\boldsymbol{e}^0_{\mathbf{l},\mathbf{i}}\|_{\boldsymbol{A}_{\mathbf{l},\mathbf{i}}}$$

- We describe the total computational cost for constructing $\widetilde{u}_{N_h,M_L} \approx u_{N_h,M_L}$ is represented by

$$\mathcal{C}_{\text{total}} = \sum_{l=0}^{L} \sum_{|\mathbf{l}|=l} \sum_{\mathbf{i} \in B_{\mathbf{l}}} \mathcal{M}_{\mathbf{l},\mathbf{i}}$$

  where $\mathcal{M}_{\mathbf{l},\mathbf{i}}$ is the number of iterations needed at the collocation point $\boldsymbol{y}_{\mathbf{l},\mathbf{i}}$.

Exploit the hierarchical structure to accelerate solutions

in the context of local hierarchical sparse-grid approximation

- The approximation $\widetilde{u}_{N_h,M_L}(x,\boldsymbol{y})$ can be represented in a hierarchical manner,

$$\widetilde{u}_{N_h,M_L}(x,\boldsymbol{y}) = \widetilde{u}_{N_h,M_{L-1}}(x,\boldsymbol{y}) + \sum_{g(\mathbf{l})=L} \sum_{\mathbf{i}\in B_{\mathbf{l}}} \widetilde{c}_{\mathbf{l},\mathbf{i}}(x)\cdot\psi_{\mathbf{l},\mathbf{i}}(\boldsymbol{y})$$

- At each collocation point $\boldsymbol{y}_{\mathbf{l},\mathbf{i}}$ on level L, $\boldsymbol{u}_{\mathbf{l},\mathbf{i}} = (u_{1,\mathbf{l},\mathbf{i}},\ldots,u_{N_j,\mathbf{l},\mathbf{i}})^\top$ can be represented by

$$u_{j,\mathbf{l},\mathbf{i}} = u_{N_h,M_{L-1}}(x_j,\boldsymbol{y}_{\mathbf{l},\mathbf{i}}) + c_{j,\mathbf{l},\mathbf{i}}, \quad \text{for} \quad j=1,\ldots,N_h$$

Key idea

Due to the decay of $|c_{j,\mathbf{l},\mathbf{i}}|$ as $|\mathbf{l}|\to\infty$, the initial guess for the CG solver is given by

$$\widetilde{\boldsymbol{u}}^0_{\mathbf{l},\mathbf{i}} = \left(\widetilde{u}_{N_h,M_{L-1}}(x_1,\boldsymbol{y}_{\mathbf{l},\mathbf{i}}),\ldots,\widetilde{u}_{N_h,M_{L-1}}(x_{N_h},\boldsymbol{y}_{\mathbf{l},\mathbf{i}})\right)^\top$$

where the error of such prediction is, for $j=1,\ldots,N_h$,

$$\left|\widetilde{u}^0_{j,\mathbf{l},\mathbf{i}} - u(x_j,\boldsymbol{y}_{\mathbf{l},\mathbf{i}})\right| \le \left|\widetilde{u}_{N_h,M_{L-1}}(x_j,\boldsymbol{y}_{\mathbf{l},\mathbf{i}}) - u_{N_h,M_{L-1}}(x_j,\boldsymbol{y}_{\mathbf{l},\mathbf{i}})\right| + c_{j,\mathbf{l}\,\mathbf{i}}$$

What about the "total" error?

when accounting for FEM, sparse-grid and solver errors

- The total error $e = u(x, \boldsymbol{y}) - \widetilde{u}_{N_h, M_L}(x, \boldsymbol{y})$ can be split into

$$\|e\| \leq \underbrace{\|u - u_{N_h}\|}_{e_1\,(\text{FEM error})} + \underbrace{\|u_{N_h} - u_{N_h, M_L}\|}_{e_2\,(\text{SG error})} + \underbrace{\|u_{N_h, M_L} - \widetilde{u}_{N_h, M_L}\|}_{e_3\,(\text{solver error})}$$

Lemma [GWZ14]: total error estimate

For the second-order elliptic PDE with homogeneous Dirichlet boundary conditions, the approximate solution $\widetilde{u}_{N_h, M_L}$ is constructed using the piecewise linear SG method, and the CG solver. Then the error $e = u - \widetilde{u}_{N_h, M_L}$ is bounded by

$$\|e\|_{L^2(D\times\Gamma)} \leq \underbrace{C_{\text{fem}} h^{r+1}}_{\text{bound of } e_1} + \underbrace{C_{\text{sg}} 2^{-2L} \sum_{k=0}^{N-1} \binom{L+N-1}{k}}_{\text{bound of } e_2} + \underbrace{2^N \binom{L+N}{N} e_{\text{cg}}}_{\text{bound of } e_3}$$

where $u \in H^{r+1}(D) \otimes L^2_\rho(\Gamma)$, the constant C_{fem} is independent of h and $\boldsymbol{y}$, the constant C_{sg} is independent of L and N, and e_{cg} is the maximum CG error.

Computational cost analyses

for ε-complexity

- The goal is to estimate the computational cost of constructing $\widetilde{u}_{N_h,M_L}$ with the prescribed accuracy $\|e\| = \|u_{N_h,M_L} - \widetilde{u}_{N_h,M_L}\| \leq \varepsilon$
- According to the error estimate, a sufficient condition of $\|e\| \leq \varepsilon$ is as follows:

$$\|e_1\| \leq C_{\text{fem}} \cdot h^{r+1} \leq \frac{\varepsilon}{3}$$

$$\|e_2\| \leq C_{\text{sg}} \cdot 2^{-2L} \sum_{k=0}^{N-1} \binom{L+N-1}{k} \leq \frac{\varepsilon}{3}$$

$$\|e_3\| \leq 2^N \binom{L+N}{N} e_{\text{cg}} \leq \frac{\varepsilon}{3}$$

- Let $\mathcal{C}_{\text{min}}$ represents the minimum cost, i.e. the minimum number of conjugate gradient iterations, to satisfy the above inequalities, and we are trying to estimate an upper bound of $\mathcal{C}_{\text{min}}$.
  - Estimate maximum h and minimum L to achieve the bounds for e_1 and e_2
  - Determine the necessary value of e_{cg} to achieve the bound for e_3
  - Estimate the upper bound for C_{min} based on e_{cg}

Computational cost analyses

for ε-complexity

- It is straightforward to obtain the upper bound of h, i.e.,

$$h \leq \left(\frac{\varepsilon}{3C_{\text{fem}}} \right)^{\frac{1}{r+1}}$$

- The upper bound of the minimum level L to achieve $e_2 \leq \frac{\varepsilon}{3}$ is given by

Lemma [GWZ14, WW95]

For $\varepsilon < 3C_{\text{sg}}$, the accuracy $\|e_2\| \leq \frac{\varepsilon}{3}$ can be achieved with level $L_{\min}$ bounded by

$$L_{\min} \leq L_k = \frac{t_k N}{2 \ln 2} + 1 \quad \text{with} \quad s = \frac{2e}{\ln 2} \left(\frac{3C_{\text{cg}}}{\varepsilon} \right)$$

where $\{t_k\}_{k=0}^{\infty}$ is a monotonically decreasing sequence defined by

$$t_k = \ln(t_{k-1} s) \quad \text{with} \quad t_0 = \frac{e}{e-1} \ln s.$$

- $\{L_k\}_{k=0}^{\infty}$ is also a monotonically decreasing sequence

Computational cost analyses

for ε-complexity

- The maximum CG error e_{cg} can be bounded by

$$e_{\text{cg}} = \max_{\mathbf{i}\in B_{\mathbf{l}},|\mathbf{l}|\leq L} \|\boldsymbol{e}_{\mathbf{l},\mathbf{i}}\|_2 \leq 2\sqrt{\kappa}\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^J \cdot \tau_0$$

with

$$\kappa = \max_{\mathbf{i}\in B_{\mathbf{l}},|\mathbf{l}|\leq L} \kappa_{\mathbf{l},\mathbf{i}} \leq \overline{\kappa}, \quad J = \max_{\mathbf{i}\in B_{\mathbf{l}},|\mathbf{l}|\leq L} J_{\mathbf{l},\mathbf{i}}, \quad \tau_0 = \max_{\mathbf{i}\in B_{\mathbf{l}},|\mathbf{l}|\leq L} \|e^0_{\mathbf{l},\mathbf{i}}\|$$

- κ is the maximum condition number of all the involved linear systems which can be bounded by a function of $1/h$ or $1/\varepsilon$.

- The minimum iteration number J to achieve $e_3 \leq \frac{\varepsilon}{3}$ for specific τ_0, ε, $\overline{\kappa}$, L and N is bounded by

$$J(\tau_0,\varepsilon,\overline{\kappa},L,N) \leq \frac{\frac{1}{2}\log_2(\overline{\kappa}) + \log_2\left[\frac{3\cdot 2^{N+1}\tau_0}{\varepsilon}\binom{L+N}{N}\right]}{\log_2\left(\frac{\sqrt{\overline{\kappa}}+1}{\sqrt{\overline{\kappa}}-1}\right)}$$

Computational cost analyses

without hierarchical acceleration

- Without hierarchical acceleration, $\tau_0 = \mathcal{O}(\|u\|_\infty)$, so that the minimum cost $\mathcal{C}_{\min}$ to achieve $\|e\| \leq \varepsilon$ can be bounded by

$$\mathcal{C}_{\min} \leq |\mathcal{H}_L(\Gamma)| \cdot J(\tau_0, \varepsilon, \overline{\kappa}, L_k, N)$$

whose estimate is given as follows:

Theorem [GWZ14], complexity without hierarchical acceleration

The minimum cost $\mathcal{C}_{\min}$ for building the standard piecewise linear SG approximation $\widetilde{u}_{N_h,M_L}(x,\boldsymbol{y})$ with the prescribed accuracy $\varepsilon > 0$ can be bounded by

$$\mathcal{C}_{\min} \leq \frac{\alpha_1}{N}\left[\alpha_2 + \alpha_3 \frac{\log_2\left(\frac{3C_{\mathrm{sg}}}{\varepsilon}\right)}{N}\right]^{\alpha_4 N} \left(\frac{3C_{\mathrm{sg}}}{\varepsilon}\right)^{\alpha_5}$$
$$\times \frac{1}{\log_2\left(\frac{\sqrt{\overline{\kappa}}+1}{\sqrt{\overline{\kappa}}-1}\right)}\left[\alpha_6 \log_2\left(\frac{3C_{\mathrm{sg}}}{\varepsilon}\right) + \log_2(\sqrt{\overline{\kappa}}\tau_0) + \alpha_7 N + \alpha_8\right],$$

where the constants $\alpha_1, \ldots, \alpha_8$ are independent of L, N and ε.

Computational cost analyses

with hierarchical acceleration

- With hierarchical acceleration, $\tau_0^l \leq C_{\text{sg}} 2^{-2l} + 2^N e_{\text{cg}}$ for $l = 1, \ldots, L$, so that the minimum cost $\mathcal{C}_{\min}$ to achieve $\|e\|| \leq \varepsilon$ can be bounded by

$$\mathcal{C}_{\min} \leq \sum_{l=0}^{L_k} |\Delta \mathcal{H}_l(\Gamma)| \cdot J(\tau_0^l, \varepsilon, \overline{\kappa}, L_k, N)$$

whose estimate is given as follows:

Theorem [GWZ14], complexity with hierarchical acceleration

The minimum cost $\mathcal{C}_{\min}$ for building the standard piecewise linear SG approximation $\widetilde{u}_{N_h, M_L}(x, \boldsymbol{y})$ with the prescribed accuracy $\varepsilon > 0$ can be bounded by

$$\mathcal{C}_{\min} \leq \alpha_1 \left[\alpha_2 + \alpha_3 \frac{\log_2\left(\frac{2C_{\text{sg}}}{\varepsilon}\right)}{N} \right]^{\alpha_4 N} \left(\frac{2C_{\text{sg}}}{\varepsilon}\right)^{\alpha_5}$$
$$\times \frac{1}{\log_2\left(\frac{\sqrt{\overline{\kappa}}+1}{\sqrt{\overline{\kappa}}-1}\right)} \left[2N - \log_2(N) + \alpha_9 + \log_2(\sqrt{\overline{\kappa}})\right],$$

where the constants $\alpha_1, \ldots, \alpha_5$ and α_9 are independent of L, N and ε.

Numerical example

Case 1: linear elliptic problem with random inputs

We consider the 2D Poisson equation with random diffusivity and forcing term, i.e.,

$$\begin{cases} \nabla \cdot (a(x, \boldsymbol{y}) \nabla u(x, \boldsymbol{y})) = f(x, \boldsymbol{y}) & \text{in} \quad [0,1]^2 \times \Gamma, \\ u(x, \boldsymbol{y}) = 0 & \text{on} \quad \partial D \times \Gamma, \end{cases}$$

where a and f arethe nonlinear functions of the random vector $\boldsymbol{y}$ given by

$$a(x, \boldsymbol{y}) = 0.1 + \exp\Big[y_1 \cos(\pi x_1) + y_2 \sin(\pi x_2)\Big],$$

and

$$f(x, \boldsymbol{y}) = 10 + \exp\Big[y_3 \cos(\pi x_1) + y_4 \sin(\pi x_2)\Big],$$

where y_n for $n = 1, 2, 3, 4$ are i. i. d. random variables following the uniform distribution $U([-1, 1])$. The quantity of interest is the mean value of the solution over $D \times \Gamma$, i.e.

$$\text{QoI} = \mathbb{E}\left[\int_D u(x, \boldsymbol{y}) dx\right],$$

Numerical example

Case 1: linear elliptic problem with random inputs

The computational savings of the piecewise SG approach with hierarchical acceleration

Basis type	Error	# SG points	hSGSC cost	hSGSC+acceleration cost	hSGSC+acceleration saving
Linear	1.0e-2	377	13,841	7,497	45.8%
	1.0e-3	1,893	81,068	38,670	52.2%
	1.0e-4	7,777	376,287	167,832	55.3%
Quadratic	1.0e-3	701	29,874	11,877	60.2%
	1.0e-4	2,285	110,744	36,760	66.8%
	1.0e-5	6,149	329,294	100,420	69.5%
Cubic	1.0e-4	1,233	59,344	23,228	60.8%
	1.0e-5	3,233	172,845	57,777	66.5%
	1.0e-6	7,079	415,760	129,433	68.8%

Numerical example

Case 2: A stochastic elliptic problem using **glocal** sparse grid SC

Let $\mathbf{x} = (x_1, x_2)$ and consider the following linear elliptic SPDE:

$$\begin{cases} -\nabla \cdot (a(x_1, \boldsymbol{y}(\omega)) \nabla u(\mathbf{x}, \boldsymbol{y}) &= \cos(x_1)\sin(x_2) \quad [0,1]^2 \times \Gamma \\ u(\mathbf{x}, \boldsymbol{y}) &= 0 \quad \text{on } \partial D \times \Gamma \end{cases}$$

The diffusion coefficient is a 1d random field (varies only in x_1) and is $a(x_1, \boldsymbol{y}) = 0.5 + \exp\{\gamma(x_1, \boldsymbol{y})\}$, where γ is a truncated 1d random field with correlation length L and covariance

$$Cov[\gamma](x_1, \tilde{x}_1) = \exp\left(-\frac{(x_1 - \tilde{x}_1)^2}{L^2}\right), \quad \forall (x_1, \tilde{x}_1) \in [0,1]$$

$$\gamma(x_1, \boldsymbol{y}) = 1 + y_1(\omega)\left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} + \sum_{n=2}^{N} \beta_n \, \varphi_n(x_1) \, y_n(\omega)$$

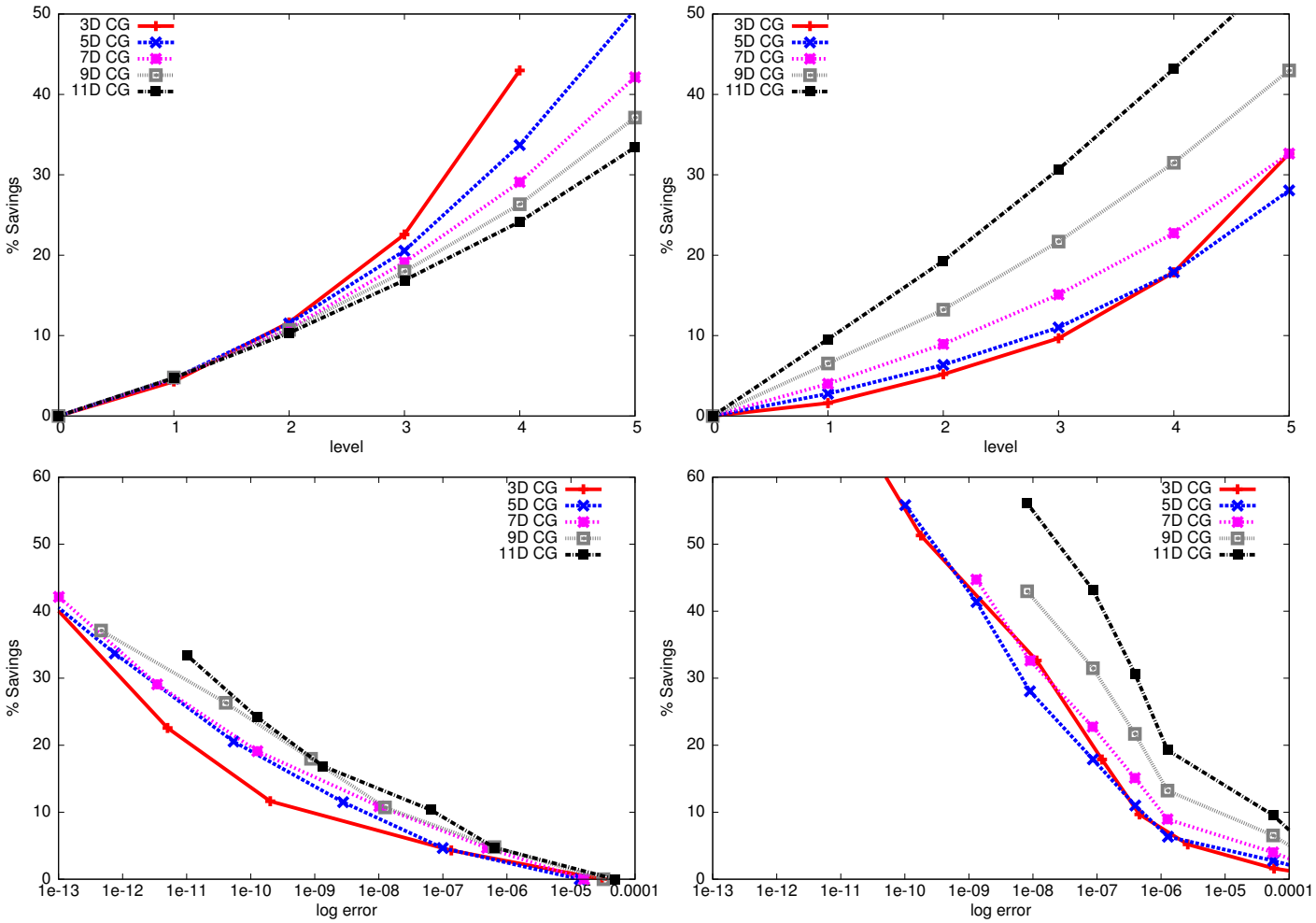
$$\beta_n := \left(\sqrt{\pi}L\right)^{1/2} e^{\frac{-\left(\lfloor \frac{n}{2} \rfloor \pi L\right)^2}{8}}, \quad \varphi_n(x_1) := \begin{cases} \sin\left(\lfloor \frac{n}{2} \rfloor \pi x_1\right), & \text{if } n \text{ even}, \\ \cos\left(\lfloor \frac{n}{2} \rfloor \pi x_1\right), & \text{if } n \text{ odd} \end{cases}$$

- $\mathbb{E}[y_n] = 0$ and $\mathbb{E}[y_n y_m] = \delta_{nm}$ for $n, m \in \mathbb{N}_+$ and iid in $U(-\sqrt{3}, \sqrt{3})$

Computational savings

QoI $= \mathbb{E}[u]$

Savings versus level and savings versus error for $L = 1/64$ (left) and $L = 1/2$ (right)



Conclusion and future work

- The bulk of the computational cost of high-dimensional extreme-scale computational simulations is associated with linear or nonlinear iterative solvers, and the convergence of such methods can be dramatically improved by using the hierarchical acceleration approach
- The new method can be easily extended to other non-intrusive methods, including global sparse-grid approximation, polynomial chaos expansion, etc..
- Extend the new approach to time-dependent problems
- Incorporate this approach into a multi-level framework using model hierarchies $\{u^s(x, \boldsymbol{y})\}_{s=1}^S$ of increasing complexity.

$$\|\widehat{u}^S - u^S\| \leq \|u^0_{M_L} - u^0\| + \sum_{s=1}^{S} \|(u^s_{M_{L-s}} - u^{s-1}_{M_{L-s}}) - (u^s - u^{s-1})\| \leq \epsilon$$

References

[GWZ14] M. Gunzburger, C. Webster, G. Zhang, *Stochastic finite element methods for partial differential equations with random input data*, **Acta Numerica**, Vol. 23, pp. 521–650, 2014

[ZWG14] G. Zhang, C. Webster, M. Gunzburger, *A hierarchical stochastic collocation method for adaptive acceleration of PDEs with random input data*, **ORNL Technical Report**, 2014

[BG04] H. Bungartz and M. Griebel, *Sparse grids*, **Acta Numerica**, Vol. 13, pp. 147–269, 2004.

[WW95] G. W Wasilkowski, H. Wożniakowski, *Explicit cost bounds of algorithms for multivariate tensor product problems*, **J. Complexity**, Vol. 11, pp. 1–56, 1995

[NTW08a] F. Nobile, R. Tempone, C. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, **SIAM Journal on Numerical Analysis**, Vol. 46(5), pp. 2309–2345, 2008

[NTW08b] F. Nobile, R. Tempone, C. Webster, *An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data*, **SIAM Journal on Numerical Analysis**, Vol. 46(5), pp. 2411–2442, 2008

[ZWGB14] G. Zhang, C. Webster, M. Gunzburger, J. Burkardt, *A Hyper-Spherical Adaptive Sparse-Grid Approach for High-Dimensional Discontinuity Detection*, submitted

[GWZ13] M. Gunzburger, C. Webster G. Zhang, *An Adaptive Wavelet Stochastic Collocation Method for Irregular Solutions of PDEs with Random Input Data*, **Springer Lecture Notes on CS&E**, Vol. 97, pp. 137–170, 2014