Adaptive multilevel stochastic Galerkin FEM for parametric PDE problems

Alex Bespalov
University of Birmingham

Joint work with: **Dirk Praetorius** (TU Wien)
Michele Ruggeri (University of Strathclyde)

Workshop
"Interplay of discretization and algebraic solvers: a posteriori error estimates and adaptivity"

INRIA Paris

8 – 10 June 2022

Engineering and Physical Sciences Research Council

UNIVERSITY OF BIRMINGHAM

What is this talk about...

* Elliptic PDE problems with random inputs
  (or, more generally, high-dimensional parametric elliptic PDEs)
* Stochastic Galerkin finite element method (SGFEM)
* Multilevel adaptivity for stochastic Galerkin FEM
* Theoretical convergence and rate optimality analysis
* Implementation and numerical results

Parametric model problem (1/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$
\begin{aligned}
-\nabla_x \cdot (a(x,\mathbf{y})\nabla_x u(x,\mathbf{y})) &= f(x) \qquad && x \in D,\ \mathbf{y} \in \Gamma, \\
u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma
\end{aligned}
$$

- Domains
  - $D \subset \mathbb{R}^2 \rightsquigarrow$ physical domain
  - $\Gamma := [-1,1]^{\mathbb{N}} \rightsquigarrow$ parameter domain

Parametric model problem (1/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$\begin{aligned} -\nabla_x \cdot (a(x,\mathbf{y})\nabla_x u(x,\mathbf{y})) &= f(x) \qquad && x \in D,\ \mathbf{y} \in \Gamma, \\ u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma \end{aligned}$$

- Domains
  - $D \subset \mathbb{R}^2 \rightsquigarrow$ physical domain
  - $\Gamma := [-1,1]^{\mathbb{N}} \rightsquigarrow$ parameter domain

Remark: parameters $y_1, y_2, \ldots$ can be seen as images (observations) of independent real-valued random variables with cumulative distribution functions $\pi_1(y_1), \pi_2(y_2), \ldots$. Then, the joint cumulative distribution function is defined as

$$\pi(\mathbf{y}) := \prod_{m=1}^{\infty} \pi_m(y_m), \quad \text{and} \quad \int_{-1}^{1} d\pi_m(y_m) = \int_{\Gamma} d\pi(\mathbf{y}) = 1.$$

Parametric model problem (1/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$\begin{aligned} -\nabla_x \cdot (a(x,\mathbf{y})\nabla_x u(x,\mathbf{y})) &= f(x) \qquad && x \in D,\ \mathbf{y} \in \Gamma, \\ u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma \end{aligned}$$

- Affine-parametric diffusion coefficient
  - $\Gamma := [-1,1]^{\mathbb{N}} \rightsquigarrow$ parameter domain
  - $a(x,\mathbf{y}) = a_0(x) + \sum_{m \in \mathbb{N}} y_m a_m(x)$ for $x \in D$, $\mathbf{y} = (y_m)_{m \in \mathbb{N}} \in \Gamma$

Parametric model problem (1/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$-\nabla_x \cdot (a(x, \mathbf{y}) \nabla_x u(x, \mathbf{y})) = f(x) \qquad x \in D,\ \mathbf{y} \in \Gamma,$$
$$u(x, \mathbf{y}) = 0 \qquad x \in \partial D,\ \mathbf{y} \in \Gamma$$

- Affine-parametric diffusion coefficient
  - $\Gamma := [-1, 1]^{\mathbb{N}} \rightsquigarrow$ parameter domain
  - $a(x, \mathbf{y}) = a_0(x) + \sum_{m \in \mathbb{N}} y_m a_m(x)$ for $x \in D$, $\mathbf{y} = (y_m)_{m \in \mathbb{N}} \in \Gamma$
  - $0 < a_0^{\min} \le a_0(x) \le a_0^{\max} < \infty$ for almost all $x \in D$
  - $\tau := \dfrac{1}{a_0^{\min}} \left\| \sum_{m \in \mathbb{N}} |a_m| \right\|_{L^\infty(D)} < 1$ & $\sum_{m \in \mathbb{N}} \|a_m\|_{L^\infty(D)} < \infty$

Remark: $a_0(x)$ typically represents the mean field, i.e., $a_0(x) = \int_\Gamma a(x, \mathbf{y}) \mathrm{d}\pi(\mathbf{y})$.

Parametric model problem (2/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$\begin{aligned} -\nabla_x \cdot (a(x,\mathbf{y})\nabla_x u(x,\mathbf{y})) &= f(x) \qquad && x \in D,\ \mathbf{y} \in \Gamma, \\ u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma \end{aligned}$$

- $\mathbb{X} := H^1_0(D)$, $\mathbb{P} := L^2_\pi(\Gamma)$, $\pi(\mathbf{y}) = \prod_{m\in\mathbb{N}} \pi_m(y_m)$; $\quad \mathbb{V} := L^2_\pi(\Gamma;\mathbb{X}) \cong \mathbb{X} \otimes \mathbb{P}$

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$\begin{aligned} -\nabla_x \cdot (a(x,\mathbf{y}) \nabla_x u(x,\mathbf{y})) &= f(x) \qquad && x \in D,\ \mathbf{y} \in \Gamma, \\ u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma \end{aligned}$$

- $\mathbb{X} := H_0^1(D)$, $\mathbb{P} := L^2_\pi(\Gamma)$, $\pi(\mathbf{y}) = \prod_{m\in\mathbb{N}} \pi_m(y_m)$; $\quad \mathbb{V} := L^2_\pi(\Gamma;\mathbb{X}) \cong \mathbb{X} \otimes \mathbb{P}$
- Bilinear forms on $\mathbb{V}$
  - $B_0(u,v) := \displaystyle\int_\Gamma \int_D a_0(x) \nabla u(x,\mathbf{y}) \cdot \nabla v(x,\mathbf{y})\, \mathrm{d}x\, \mathrm{d}\pi(\mathbf{y})$
  - $B(u,v) := B_0(u,v) + \displaystyle\sum_{m\in\mathbb{N}} \int_\Gamma y_m \int_D a_m(x) \nabla u(x,\mathbf{y}) \cdot \nabla v(x,\mathbf{y})\, \mathrm{d}x\, \mathrm{d}\pi(\mathbf{y})$

Parametric model problem (2/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$
\begin{aligned}
-\nabla_x \cdot (a(x,\mathbf{y})\nabla_x u(x,\mathbf{y})) &= f(x) && x \in D,\ \mathbf{y} \in \Gamma,\\
u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma
\end{aligned}
$$

- $\mathbb{X} := H_0^1(D)$, $\mathbb{P} := L_\pi^2(\Gamma)$, $\pi(\mathbf{y}) = \prod_{m\in\mathbb{N}} \pi_m(y_m)$; $\mathbb{V} := L_\pi^2(\Gamma;\mathbb{X}) \cong \mathbb{X} \otimes \mathbb{P}$
- Bilinear forms on $\mathbb{V}$
  - $B_0(u,v) := \displaystyle\int_\Gamma \int_D a_0(x)\nabla u(x,\mathbf{y}) \cdot \nabla v(x,\mathbf{y})\,\mathrm{d}x\,\mathrm{d}\pi(\mathbf{y})$
  - $B(u,v) := B_0(u,v) + \displaystyle\sum_{m\in\mathbb{N}} \int_\Gamma y_m \int_D a_m(x)\nabla u(x,\mathbf{y}) \cdot \nabla v(x,\mathbf{y})\,\mathrm{d}x\,\mathrm{d}\pi(\mathbf{y})$

Weak formulation: given $f \in L^2(D)$, find $u \in \mathbb{V}$ such that

$$
B(u,v) = F(v) := \int_\Gamma \int_D f(x) v(x,\mathbf{y})\,\mathrm{d}x\,\mathrm{d}\pi(\mathbf{y}) \quad \text{for all } v \in \mathbb{V} \qquad (\star)
$$

Parametric model problem (2/2)

Problem formulation: find $u : D \times \Gamma \to \mathbb{R}$ satisfying

$$\begin{aligned} -\nabla_x \cdot (a(x,\mathbf{y})\nabla_x u(x,\mathbf{y})) &= f(x) \qquad && x \in D,\ \mathbf{y} \in \Gamma, \\ u(x,\mathbf{y}) &= 0 && x \in \partial D,\ \mathbf{y} \in \Gamma \end{aligned}$$

- $\mathbb{X} := H^1_0(D)$, $\mathbb{P} := L^2_\pi(\Gamma)$, $\pi(\mathbf{y}) = \prod_{m\in\mathbb{N}} \pi_m(y_m)$; $\quad \mathbb{V} := L^2_\pi(\Gamma; \mathbb{X}) \cong \mathbb{X} \otimes \mathbb{P}$
- Bilinear forms on $\mathbb{V}$
  - $B_0(u,v) := \int_\Gamma \int_D a_0(x) \nabla u(x,\mathbf{y}) \cdot \nabla v(x,\mathbf{y})\, \mathrm{d}x\, \mathrm{d}\pi(\mathbf{y})$
  - $B(u,v) := B_0(u,v) + \sum_{m\in\mathbb{N}} \int_\Gamma y_m \int_D a_m(x) \nabla u(x,\mathbf{y}) \cdot \nabla v(x,\mathbf{y})\, \mathrm{d}x\, \mathrm{d}\pi(\mathbf{y})$

Weak formulation: given $f \in L^2(D)$, find $u \in \mathbb{V}$ such that

$$B(u,v) = F(v) := \int_\Gamma \int_D f(x) v(x,\mathbf{y})\, \mathrm{d}x\, \mathrm{d}\pi(\mathbf{y}) \quad \text{for all } v \in \mathbb{V} \qquad (\star)$$

[Schwab, Gittelson '11]: the assumptions on $a(x,\mathbf{y})$ ensure the wellposedness of $(\star)$.

Stochastic Galerkin FEM (1/2)

- Finite dimensional subspace

$$\mathbb{V}_\bullet \subset \mathbb{V} \cong \mathbb{X} \otimes \mathbb{P}$$

Galerkin projection:

$$\text{find} \quad u_\bullet \in \mathbb{V}_\bullet \text{ such that } \quad B(u_\bullet, v_\bullet) = F(v_\bullet) \quad \text{for all } v \in \mathbb{V}_\bullet$$

Stochastic Galerkin FEM (1/2)

- Finite dimensional subspace

$$\mathbb{V}_\bullet \subset \mathbb{V} \cong \mathbb{X} \otimes \mathbb{P}$$

Galerkin projection:

$$\text{find} \quad u_\bullet \in \mathbb{V}_\bullet \text{ such that } \quad B(u_\bullet, v_\bullet) = F(v_\bullet) \quad \text{for all } v \in \mathbb{V}_\bullet$$

- Galerkin orthogonality

$$B(u - u_\bullet, v_\bullet) = 0 \quad \text{for all } v_\bullet \in \mathbb{V}_\bullet$$

- Best approximation property

$$|||\, u - u_\bullet \,||| = \min_{v_\bullet \in \mathbb{V}_\bullet} |||\, u - v_\bullet \,|||, \quad \text{where} \quad |||\, \cdot \,||| := B(\cdot, \cdot)^{1/2}$$

Stochastic Galerkin FEM (1/2)

- Finite dimensional subspace

$$\mathbb{V}_\bullet \subset \mathbb{V} \cong \mathbb{X} \otimes \mathbb{P}$$

Galerkin projection:

find $u_\bullet \in \mathbb{V}_\bullet$ such that $B(u_\bullet, v_\bullet) = F(v_\bullet)$ for all $v \in \mathbb{V}_\bullet$

- Galerkin orthogonality

$$B(u - u_\bullet, v_\bullet) = 0 \quad \text{for all } v_\bullet \in \mathbb{V}_\bullet$$

- Best approximation property

$$|||\, u - u_\bullet \,||| = \min_{v_\bullet \in \mathbb{V}_\bullet} |||\, u - v_\bullet \,|||, \quad \text{where} \quad |||\cdot||| := B(\cdot,\cdot)^{1/2}$$

- Main question: how to choose $\mathbb{V}_\bullet$?

Stochastic Galerkin FEM (2/2)

- $\{P_\nu : \nu \in \mathfrak{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathfrak{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$

Stochastic Galerkin FEM (2/2)

- $\{P_\nu : \nu \in \mathfrak{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathfrak{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$
- gPC expansion: $\mathbb{V} \ni u(x, \mathbf{y}) = \sum_{\nu \in \mathfrak{I}} u_\nu(x) P_\nu(\mathbf{y})$ with unique $u_\nu \in \mathbb{X} = H^1_0(D)$

Stochastic Galerkin FEM (2/2)

- $\{P_\nu : \nu \in \mathcal{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathcal{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$
- gPC expansion: $\mathbb{V} \ni u(x, \mathbf{y}) = \sum_{\nu \in \mathcal{I}} u_\nu(x) P_\nu(\mathbf{y})$ with unique $u_\nu \in \mathbb{X} = H^1_0(D)$
- Discretisation in the parameter domain
  - finite index set $\mathcal{P}_\bullet \subset \mathcal{I} \implies \mathbb{P}_\bullet = \mathrm{span}\{P_\nu : \nu \in \mathcal{P}_\bullet\} \subset \mathbb{P} = L^2_\pi(\Gamma)$

Stochastic Galerkin FEM (2/2)

- $\{P_\nu : \nu \in \mathfrak{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathfrak{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$
- gPC expansion: $\mathbb{V} \ni u(x, \mathbf{y}) = \sum_{\nu \in \mathfrak{I}} u_\nu(x) P_\nu(\mathbf{y})$ with unique $u_\nu \in \mathbb{X} = H^1_0(D)$
- Discretisation in the parameter domain
  - finite index set $\mathfrak{P}_\bullet \subset \mathfrak{I} \implies \mathbb{P}_\bullet = \mathrm{span}\{P_\nu : \nu \in \mathfrak{P}_\bullet\} \subset \mathbb{P} = L^2_\pi(\Gamma)$
  - semidiscrete approximation via truncation of gPC expansion

$$u(x, \mathbf{y}) \approx \sum_{\nu \in \mathfrak{P}_\bullet} u_\nu(x) P_\nu(\mathbf{y}) \in \mathbb{X} \otimes \mathbb{P}_\bullet \quad \text{with coefficients } u_\nu \in \mathbb{X} = H^1_0(D)$$

Stochastic Galerkin FEM (2/2)

- $\{P_\nu : \nu \in \mathcal{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathcal{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$
- gPC expansion: $\mathbb{V} \ni u(x, \mathbf{y}) = \sum_{\nu \in \mathcal{I}} u_\nu(x) P_\nu(\mathbf{y})$ with unique $u_\nu \in \mathbb{X} = H^1_0(D)$
- Discretisation in the parameter domain
  - finite index set $\mathcal{P}_\bullet \subset \mathcal{I} \implies \mathbb{P}_\bullet = \mathrm{span}\{P_\nu : \nu \in \mathcal{P}_\bullet\} \subset \mathbb{P} = L^2_\pi(\Gamma)$
  - semidiscrete approximation via truncation of gPC expansion

$$u(x, \mathbf{y}) \approx \sum_{\nu \in \mathcal{P}_\bullet} u_\nu(x) P_\nu(\mathbf{y}) \in \mathbb{X} \otimes \mathbb{P}_\bullet \quad \text{with coefficients } u_\nu \in \mathbb{X} = H^1_0(D)$$

- Discretisation in the physical domain (multilevel SGFEM)

- $\{P_\nu : \nu \in \mathcal{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathcal{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$
- gPC expansion: $\mathbb{V} \ni u(x, \mathbf{y}) = \sum_{\nu \in \mathcal{I}} u_\nu(x) P_\nu(\mathbf{y})$ with unique $u_\nu \in \mathbb{X} = H^1_0(D)$
- Discretisation in the parameter domain
  - finite index set $\mathcal{P}_\bullet \subset \mathcal{I} \implies \mathbb{P}_\bullet = \mathrm{span}\{P_\nu : \nu \in \mathcal{P}_\bullet\} \subset \mathbb{P} = L^2_\pi(\Gamma)$
  - semidiscrete approximation via truncation of gPC expansion

$$u(x, \mathbf{y}) \approx \sum_{\nu \in \mathcal{P}_\bullet} u_\nu(x) P_\nu(\mathbf{y}) \in \mathbb{X} \otimes \mathbb{P}_\bullet \quad \text{with coefficients } u_\nu \in \mathbb{X} = H^1_0(D)$$

- Discretisation in the physical domain (multilevel SGFEM)
  - a hierarchy of meshes $\rightsquigarrow \{\mathcal{T}_{\bullet\nu}\}_{\nu \in \mathcal{P}_\bullet}$
  - $u_\bullet(x, \mathbf{y}) = \sum_{\nu \in \mathcal{P}_\bullet} u_{\bullet\nu}(x) P_\nu(\mathbf{y})$ with $u_{\bullet\nu} \in \mathbb{X}_{\bullet\nu} = \mathcal{S}^1_0(\mathcal{T}_{\bullet\nu})$ for all $\nu \in \mathcal{P}_\bullet$

Stochastic Galerkin FEM (2/2)

- $\{P_\nu : \nu \in \mathcal{I}\}$ is a countable orthonormal polynomial basis of $\mathbb{P} = L^2_\pi(\Gamma)$
- $\mathcal{I} := \{\nu \in \mathbb{N}_0^{\mathbb{N}} : \#\mathrm{supp}(\nu) < \infty\}$ where $\mathrm{supp}(\nu) = \{m \in \mathbb{N} : \nu_m \neq 0\}$
- gPC expansion: $\mathbb{V} \ni u(x, \mathbf{y}) = \sum_{\nu\in\mathcal{I}} u_\nu(x) P_\nu(\mathbf{y})$ with unique $u_\nu \in \mathbb{X} = H^1_0(D)$
- Discretisation in the parameter domain
  - finite index set $\mathcal{P}_\bullet \subset \mathcal{I} \implies \mathbb{P}_\bullet = \mathrm{span}\{P_\nu : \nu \in \mathcal{P}_\bullet\} \subset \mathbb{P} = L^2_\pi(\Gamma)$
  - semidiscrete approximation via truncation of gPC expansion

$$u(x, \mathbf{y}) \approx \sum_{\nu\in\mathcal{P}_\bullet} u_\nu(x) P_\nu(\mathbf{y}) \in \mathbb{X} \otimes \mathbb{P}_\bullet \quad \text{with coefficients } u_\nu \in \mathbb{X} = H^1_0(D)$$

- Discretisation in the physical domain (multilevel SGFEM)
  - a hierarchy of meshes $\rightsquigarrow \{\mathcal{T}_{\bullet\nu}\}_{\nu\in\mathcal{P}_\bullet}$
  - $u_\bullet(x, \mathbf{y}) = \sum_{\nu\in\mathcal{P}_\bullet} u_{\bullet\nu}(x) P_\nu(\mathbf{y})$ with $u_{\bullet\nu} \in \mathbb{X}_{\bullet\nu} = \mathcal{S}^1_0(\mathcal{T}_{\bullet\nu})$ for all $\nu \in \mathcal{P}_\bullet$
- $\mathbb{V}_\bullet := \bigoplus_{\nu\in\mathcal{P}_\bullet} \left[\mathbb{X}_{\bullet\nu} \otimes \mathrm{span}\{P_\nu\}\right]$, $\dim \mathbb{V}_\bullet = \sum_{\nu\in\mathcal{P}_\bullet} \dim \mathbb{X}_{\bullet\nu}$

Multilevel methods

- Multilevel methods in UQ
  - Multilevel Monte Carlo
    [Giles '08], [Cliffe, Giles, Scheichl, Teckentrup '11], [Giles '15]
  - Multilevel quasi-Monte Carlo
    [Kuo, Schwab, Sloan '15]
  - Multilevel stochastic collocation
    [Teckentrup, Jantsch, Webster, Gunzburger '15]
    [Lang, Scheichl, Silvester '20]

Multilevel methods

- Multilevel methods in UQ
  - Multilevel Monte Carlo
    [Giles '08], [Cliffe, Giles, Scheichl, Teckentrup '11], [Giles '15]
  - Multilevel quasi-Monte Carlo
    [Kuo, Schwab, Sloan '15]
  - Multilevel stochastic collocation
    [Teckentrup, Jantsch, Webster, Gunzburger '15]
    [Lang, Scheichl, Silvester '20]
  - Multilevel stochastic Galerkin FEM
    - theoretical benchmarks for convergence analysis
      [Cohen, DeVore, Schwab '10, '11], [Gittelson '13]
    - practical realisations
      [Eigel, Gittelson, Schwab, Zander '14], [Crowder, Powell, B. '19]

Multilevel methods

- Multilevel methods in UQ
  - Multilevel Monte Carlo
    [Giles '08], [Cliffe, Giles, Scheichl, Teckentrup '11], [Giles '15]
  - Multilevel quasi-Monte Carlo
    [Kuo, Schwab, Sloan '15]
  - Multilevel stochastic collocation
    [Teckentrup, Jantsch, Webster, Gunzburger '15]
    [Lang, Scheichl, Silvester '20]
  - Multilevel stochastic Galerkin FEM
    - theoretical benchmarks for convergence analysis
      [Cohen, DeVore, Schwab '10, '11], [Gittelson '13]
    - practical realisations
      [Eigel, Gittelson, Schwab, Zander '14], [Crowder, Powell, B. '19]
    - optimal convergence rates for practical realisations $\rightsquigarrow$ open problem...

Hierarchical a posteriori error estimates(1/3)

- Why hierarchical error estimates?
  - provide effective error estimation in the energy norm (eff. indices ≈ 1)
  - provide reliable and efficient estimates of the error reduction;
    error reduction indicators $\rightsquigarrow$ key to adaptivity
  - facilitate convergence analysis of adaptive algorithms

Hierarchical a posteriori error estimates(1/3)

- Why hierarchical error estimates?
  - provide effective error estimation in the energy norm (eff. indices ≈ 1)
  - provide reliable and efficient estimates of the error reduction;
    error reduction indicators $\rightsquigarrow$ key to adaptivity
  - facilitate convergence analysis of adaptive algorithms

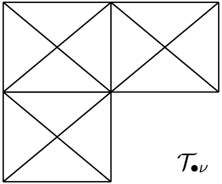
- Main ideas
  - use an enhanced approximation $\widehat{u}_\bullet \in \widehat{V}_\bullet \supset V_\bullet$ to estimate

$$\left|\!\left|\!\left| u - u_\bullet \right|\!\right|\!\right| \approx \underbrace{\left|\!\left|\!\left| \widehat{u}_\bullet - u_\bullet \right|\!\right|\!\right|}_{\text{computable}}$$

  - the enhanced approximation is based on ‘uniform refinement’
  - avoiding the computation of $\widehat{u}_\bullet$ via a decomposition of $\widehat{V}_\bullet$
    in the spirit of [Bank, Weiser ’85]

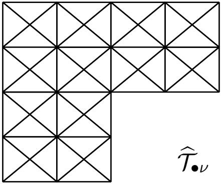
Hierarchical a posteriori error estimates (2/3)

- Enhancement of approximations in physical domain
  - initial mesh $\mathcal{T}_0$
  - add new vertices to $\mathcal{T}_{\bullet\nu}$ $\rightsquigarrow$ mesh refinement
  - mesh refinement by newest vertex bisection (NVB)
  - $\widehat{\mathcal{T}}_{\bullet\nu}$ $\rightsquigarrow$ uniform refinement of $\mathcal{T}_{\bullet\nu}$



Hierarchical a posteriori error estimates (2/3)

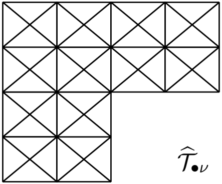
- Enhancement of approximations in physical domain
  - initial mesh $\mathcal{T}_0$
  - add new vertices to $\mathcal{T}_{\bullet\nu}$ $\rightsquigarrow$ mesh refinement
  - mesh refinement by newest vertex bisection (NVB)
  - $\widehat{\mathcal{T}}_{\bullet\nu}$ $\rightsquigarrow$ uniform refinement of $\mathcal{T}_{\bullet\nu}$



$\widehat{\mathcal{T}}_{\bullet\nu}$

Hierarchical a posteriori error estimates (2/3)

- Enhancement of approximations in physical domain
  - initial mesh $\mathcal{T}_0$
  - add new vertices to $\mathcal{T}_{\bullet\nu}$ $\rightsquigarrow$ mesh refinement
  - mesh refinement by newest vertex bisection (NVB)
  - $\widehat{\mathcal{T}}_{\bullet\nu}$ $\rightsquigarrow$ uniform refinement of $\mathcal{T}_{\bullet\nu}$



$\widehat{\mathcal{T}}_{\bullet\nu}$

- Enhancement of approximations in the parameter domain
  - add new indices to $\mathcal{P}_\bullet$
  - finite set $\mathcal{Q}_\bullet \subset \mathcal{I} \setminus \mathcal{P}_\bullet$ $\rightsquigarrow$ detail index set ('boundary' of $\mathcal{P}_\bullet$)
  - $\widehat{\mathcal{P}}_\bullet = \mathcal{P}_\bullet \cup \mathcal{Q}_\bullet$ $\rightsquigarrow$ uniform enrichment of $\mathcal{P}_\bullet$
  - $\widehat{\mathbb{P}}_\bullet = \mathrm{span}\{P_\nu : \nu \in \widehat{\mathcal{P}}_\bullet\} \supset \mathbb{P}_\bullet$

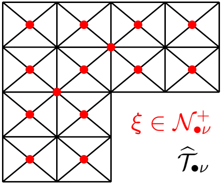
Example

- $\mathcal{P}_\bullet = \{(0,0,\dots);(1,0,\dots);(0,1,0,\dots)\}$

  $\Longrightarrow \mathcal{Q}_\bullet = \{(2,0,\dots);(1,1,0,\dots);(0,2,0,\dots);(0,0,1,0,\dots)\}$

Hierarchical a posteriori error estimates (2/3)

- Enhancement of approximations in physical domain
  - initial mesh $\mathcal{T}_0$
  - add new vertices to $\mathcal{T}_{\bullet\nu}$ $\rightsquigarrow$ mesh refinement
  - mesh refinement by newest vertex bisection (NVB)
  - $\widehat{\mathcal{T}}_{\bullet\nu}$ $\rightsquigarrow$ uniform refinement of $\mathcal{T}_{\bullet\nu}$



- Enhancement of approximations in the parameter domain
  - add new indices to $\mathcal{P}_\bullet$
  - finite set $\mathcal{Q}_\bullet \subset \mathcal{I} \setminus \mathcal{P}_\bullet$ $\rightsquigarrow$ detail index set ('boundary' of $\mathcal{P}_\bullet$)
  - $\widehat{\mathcal{P}}_\bullet = \mathcal{P}_\bullet \cup \mathcal{Q}_\bullet$ $\rightsquigarrow$ uniform enrichment of $\mathcal{P}_\bullet$
  - $\widehat{\mathbb{P}}_\bullet = \mathrm{span}\{P_\nu : \nu \in \widehat{\mathcal{P}}_\bullet\} \supset \mathbb{P}_\bullet$

- Local error indicators
  - local spatial indicators $\eta_\bullet(\nu, \xi)$ for each $\nu \in \mathcal{P}_\bullet$ and for each $\xi \in \mathcal{N}_{\bullet\nu}^{+}$
  - individual parametric indicators $\eta_\bullet(\nu)$ for each $\nu \in \mathcal{Q}_\bullet$

Hierarchical a posteriori error estimates (3/3)

- Total error estimate (for multilevel SGFEM)

$$\eta_\bullet^2 = (\text{estim. } \mathbb{X}\text{-error})^2 + (\text{estim. } \mathbb{P}\text{-error})^2 = \sum_{\nu \in \mathcal{P}_\bullet} \sum_{\xi \in \mathcal{N}_{\bullet,\nu}^+} \eta_\bullet^2(\nu, \xi) + \sum_{\nu \in \mathcal{Q}_\bullet} \eta_\bullet^2(\nu)$$

Hierarchical a posteriori error estimates (3/3)

- Total error estimate (for multilevel SGFEM)

$$\eta_\bullet^2 = (\text{estim. } \mathbb{X}\text{-error})^2 + (\text{estim. } \mathbb{P}\text{-error})^2 = \sum_{\nu \in \mathfrak{P}_\bullet} \sum_{\xi \in \mathcal{N}_{\bullet,\nu}^+} \eta_\bullet^2(\nu, \xi) + \sum_{\nu \in \mathfrak{Q}_\bullet} \eta_\bullet^2(\nu)$$

Theorem 1 [B., Praetorius, Ruggeri; SIAM/ASA JUQ '21]

- $\eta_\bullet \simeq |||\, \widehat{u}_\bullet - u_\bullet \,|||$
- $|||\, u - u_\bullet \,||| \gtrsim \eta_\bullet$ (efficiency)
- saturation assumption $\Longrightarrow$ $|||\, u - u_\bullet \,||| \lesssim \eta_\bullet$ (reliability)

Adaptive SGFEM algorithm

INPUT: initial mesh $\mathcal{T}_0$, initial index set $\mathcal{P}_0 = \{(0, 0, \dots)\}$, tolerance tol

FOR $\ell = 0, 1, 2, 3, \dots$ DO:

- SOLVE: compute $u_\ell \in \mathbb{V}_\ell$ for index set $\mathcal{P}_\ell$ and meshes $\mathcal{T}_{\ell,\nu}$ $(\nu \in \mathcal{P}_\ell)$
- ESTIMATE: compute *local* error indicators and the *total* error estimate
  - spatial & parametric indicators
    $\{\eta_\ell(\nu, \xi);\ \xi \in \mathcal{N}^+_{\ell,\nu},\ \nu \in \mathcal{P}_\ell\}$ & $\{\eta_\ell(\nu);\ \nu \in \mathcal{Q}_\ell\}$
  - energy error estimate η_ℓ
  - IF $\eta_\ell <$ tol THEN STOP
- MARK: mark certain vertices $\mathcal{M}_{\ell,\nu} \subseteq \mathcal{N}^+_{\ell,\nu}$ $(\nu \in \mathcal{P}_\ell)$ and indices $\mathcal{R}_\ell \subseteq \mathcal{Q}_\ell$
- REFINE: enhance approximation space
  - mesh refinement (NVB) $\rightsquigarrow \mathcal{T}_{\ell+1,\nu} = \textbf{refine}(\mathcal{T}_{\ell,\nu}, \mathcal{M}_{\ell,\nu})\ \ \forall\, \nu \in \mathcal{P}_\ell$
  - parametric enrichment $\rightsquigarrow \mathcal{P}_{\ell+1} = \mathcal{P}_\ell \cup \mathcal{R}_\ell,\ \ \mathcal{T}_{\ell,\nu} = \mathcal{T}_0\ \ \forall\, \nu \in \mathcal{R}_\ell$

OUTPUT: stochastic Galerkin approximations $\{u_\ell\}$ and error estimates $\{\eta_\ell\}$

Marking strategy

- `MARK`: mark certain vertices $\mathcal{M}_{\ell\nu} \subseteq \mathcal{N}_{\ell\nu}^{+}$ ($\nu \in \mathcal{P}_\ell$) and indices $\mathcal{R}_\ell \subseteq \mathcal{Q}_\ell$

Marking strategy

- `MARK`: mark certain vertices $\mathcal{M}_{\ell\nu} \subseteq \mathcal{N}^+_{\ell\nu}$ ($\nu \in \mathcal{P}_\ell$) and indices $\mathcal{R}_\ell \subseteq \mathcal{Q}_\ell$

- Recall the error estimate: $\eta_\ell^2 = \sum_{\nu \in \mathcal{P}_\ell} \sum_{\xi \in \mathcal{N}^+_{\ell,\nu}} \eta_\ell^2(\nu, \xi) + \sum_{\nu \in \mathcal{Q}_\ell} \eta_\ell^2(\nu)$

- Separate marking/enrichment of spatial and parametric components
  - two independent marking thresholds $\theta_{\mathbb{X}} \in (0, 1]$ and $\theta_{\mathbb{P}} \in (0, 1]$

Marking strategy

- `MARK`: mark certain vertices $\mathcal{M}_{\ell\nu} \subseteq \mathcal{N}_{\ell\nu}^{+}$ $(\nu \in \mathcal{P}_\ell)$ and indices $\mathcal{R}_\ell \subseteq \mathcal{Q}_\ell$

- Recall the error estimate: $\eta_\ell^2 = \sum_{\nu \in \mathcal{P}_\ell} \sum_{\xi \in \mathcal{N}_{\ell,\nu}^{+}} \eta_\ell^2(\nu, \xi) + \sum_{\nu \in \mathcal{Q}_\ell} \eta_\ell^2(\nu)$

- Separate marking/enrichment of spatial and parametric components
  - two independent marking thresholds $\theta_{\mathbb{X}} \in (0, 1]$ and $\theta_{\mathbb{P}} \in (0, 1]$

- Combined marking/enrichment of spatial and parametric components
  - use Dörfler marking for

$$\{\eta_\ell(\nu, \xi);\ \nu \in \mathcal{P}_\ell,\ \xi \in \mathcal{N}_{\ell,\nu}^{+}\} \cup \{\eta_\ell(\nu);\ \nu \in \mathcal{Q}_\ell\}$$

    with $\theta \in (0, 1]$ yields $\mathcal{M}_{\ell,\nu} \subseteq \mathcal{N}_{\ell,\nu}^{+}$ $(\nu \in \mathcal{P}_\ell)$ and $\mathcal{R}_\ell \subseteq \mathcal{Q}_\ell$ satisfying

$$\theta\, \eta_\ell^2 \le \sum_{\nu \in \mathcal{P}_\ell} \sum_{\xi \in \mathcal{M}_{\ell,\nu}} \eta_\ell^2(\nu, \xi) + \sum_{\nu \in \mathcal{R}_\ell} \eta_\ell^2(\nu)$$

  - too costly for the single-level SGFEM

Convergence of adaptive algorithm (1/2)

Single-level SGFEM: [B., Praetorius, Rocchi, Ruggeri; SINUM '19]

Multilevel SGFEM: [B., Praetorius, Ruggeri; IMANUM (appeared online)]

Theorem 2 (plain convergence)

For any $\theta_{\mathbb{X}}$, $\theta_{\mathbb{P}} \in (0, 1]$ (resp., $\theta \in (0, 1]$) $\Longrightarrow \lim_{\ell \to \infty} \eta_\ell = 0$.

Convergence of adaptive algorithm (1/2)

Single-level SGFEM: [B., Praetorius, Rocchi, Ruggeri; SINUM '19]

Multilevel SGFEM: [B., Praetorius, Ruggeri; IMANUM (appeared online)]

Theorem 2 (plain convergence)

For any $\theta_{\mathbb{X}}$, $\theta_{\mathbb{P}} \in (0, 1]$ (resp., $\theta \in (0, 1]$) $\Longrightarrow \lim_{\ell \to \infty} \eta_\ell = 0$.

Remarks

- Theorem 2 provides a theoretical guarantee that the adaptive algorithm terminates after a finite number of iterations
- No assumptions on the Galerkin approximations generated by the algorithm
- The result extends to more general marking criteria and to $D \subset \mathbb{R}^3$

Proof of Theorem 2

Energy error estimate: $\eta_\ell^2 = \sum_{\nu \in \mathcal{P}_\ell} \sum_{\xi \in \mathcal{N}^+_{\ell,\nu}} \eta_\ell^2(\nu, \xi) + \sum_{\nu \in \mathcal{Q}_\ell} \eta_\ell^2(\nu)$

Proof of Theorem 2

Energy error estimate: $\eta_\ell^2 = \sum_{\nu \in \mathcal{P}_\ell} \sum_{\xi \in \mathcal{N}_{\ell,\nu}^+} \eta_\ell^2(\nu, \xi) + \sum_{\nu \in \mathcal{Q}_\ell} \eta_\ell^2(\nu)$

Proposition 1 (Extension of the results in [Morin, Siebert, Veeser '08])

The multilevel spatial refinement that follows Dörfler marking strategy *along a subsequence* $(\ell_k)_{k \in \mathcal{N}_0}$ guarantees convergence of spatial error estimates *along this subsequence*, i.e., $\sum_{\nu \in \mathcal{P}_{\ell_k}} \sum_{\xi \in \mathcal{N}_{\ell_k}^+} \eta_{\ell_k}(\nu, \xi)^2 \to 0$ as $k \to \infty$.

Proof of Theorem 2

Energy error estimate: $\eta_\ell^2 = \sum_{\nu \in \mathcal{P}_\ell} \sum_{\xi \in \mathcal{N}^+_{\ell,\nu}} \eta_\ell^2(\nu, \xi) + \sum_{\nu \in \mathcal{Q}_\ell} \eta_\ell^2(\nu)$

Proposition 1 (Extension of the results in [Morin, Siebert, Veeser '08])

The multilevel spatial refinement that follows Dörfler marking strategy *along a subsequence* $(\ell_k)_{k \in \mathcal{N}_0}$ guarantees convergence of spatial error estimates *along this subsequence*, i.e., $\sum_{\nu \in \mathcal{P}_{\ell_k}} \sum_{\xi \in \mathcal{N}^+_{\ell_k}} \eta_{\ell_k}(\nu, \xi)^2 \to 0$ as $k \to \infty$.

Proposition 2

The parametric enrichment that follows the Dörfler marking strategy *along a subsequence* $(\mathcal{P}_{\ell_k})_{k \in \mathcal{N}_0} \subset (\mathcal{P}_\ell)_{\ell \in \mathcal{N}_0}$ guarantees convergence of the *whole sequence* of parametric error estimates, i.e., $\sum_{\nu \in \mathcal{Q}_\ell} \eta_\ell(\nu)^2 \to 0$ as $\ell \to \infty$.

Convergence of adaptive algorithm (2/2)

Single-level SGFEM: [B., Praetorius, Rocchi, Ruggeri; SINUM '19]

Multilevel SGFEM: [B., Praetorius, Ruggeri; IMANUM (appeared online)]

Theorem 3 (linear convergence)

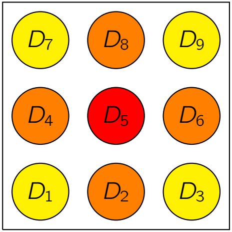
Saturation assumption $\Longrightarrow$ linear converg. for any $\theta_{\mathbb{X}}$, $\theta_{\mathbb{P}} \in (0,1]$ (resp., $\theta \in (0,1]$):

$$\exists\, q_{\mathrm{lin}} \in (0,1) \text{ s.t. } |||\, u - u_{\ell+1} \,||| \le q_{\mathrm{lin}} \,|||\, u - u_\ell \,||| \quad \forall\, \ell \in \mathbb{N}_0$$

Numerical results: cookie problem

[B., Praetorius, Ruggeri; IMANUM (appeared online)]

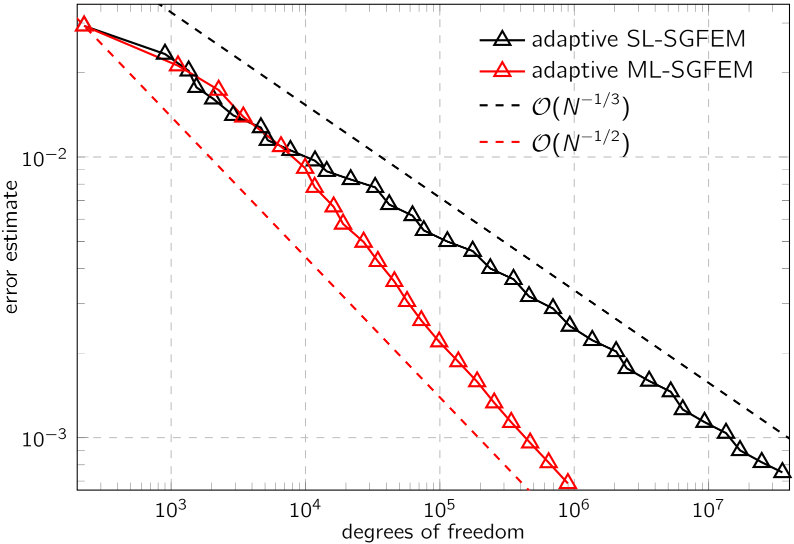
- $-\nabla \cdot (a \nabla u) = f$ in $D \times \Gamma$, $\quad u = 0$ on $\partial D \times \Gamma$
- $a(x, \mathbf{y}) = a_0(x) + \sum_{m \in \mathbb{N}} y_m a_m(x)$
- $D = (0,1)^2 \rightsquigarrow$ square domain
  - nine circular inclusions $D_m \subset D$ $(m = 1, \ldots, 9)$
- Expansion coefficients $\{a_m\}_{m \in \mathbb{N}_0}$
  - $a_0 \equiv 1$
  - $a_m = 0.5\, \chi_{D_m}$ for $m = 1, 3, 7, 9$
  - $a_m = 0.7\, \chi_{D_m}$ for $m = 2, 4, 6, 8$
  - $a_m = 0.9\, \chi_{D_m}$ for $m = 5$
  - $a_m \equiv 0$ for $m > 9$

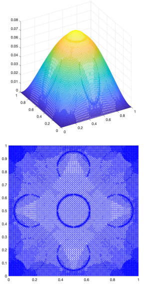


D

- $f \equiv 1$
- $\mathrm{d}\pi_m(y_m) = \frac{1}{2}\, \mathrm{d}y_m \rightsquigarrow$ uniform probability measure on $[-1, 1]$

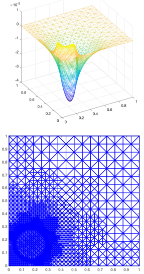
Numerical results: single-level vs. multilevel adaptive SGFEM





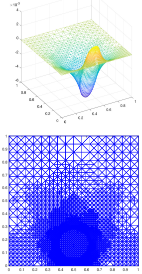
$\nu = \mathbf{0}$

$\#\mathcal{T}_{\ell,\nu} = 84050$



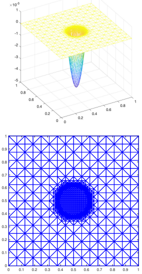
$\nu = (1\ 0)$

$\#\mathcal{T}_{\ell,\nu} = 10994$



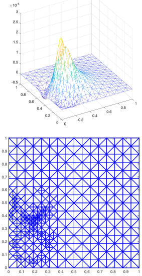
$\nu = (0\ 1)$

$\#\mathcal{T}_{\ell,\nu} = 16420$



$\nu = (0\ 0\ 0\ 0\ 1)$

$\#\mathcal{T}_{\ell,\nu} = 9528$



$\nu = (1\ 0\ 0\ 1)$

$\#\mathcal{T}_{\ell,\nu} = 839$

Optimal convergence of adaptive multilevel SGFEM

[B., Praetorius, Ruggeri; IMANUM (appeared online)]

- Concept of 'multilevel structure' $\rightsquigarrow \mathbb{P}_\bullet = [\mathcal{P}_\bullet, (\mathcal{T}_{\bullet\nu})_{\nu\in\mathcal{P}_\bullet}]$, $\#\mathbb{P}_\bullet \simeq \dim \mathbb{V}_\bullet$

Optimal convergence of adaptive multilevel SGFEM

[B., Praetorius, Ruggeri; IMANUM (appeared online)]

- Concept of ‘multilevel structure’ $\rightsquigarrow \mathbb{P}_\bullet = [\mathcal{P}_\bullet, (\mathcal{T}_{\bullet\nu})_{\nu\in\mathcal{P}_\bullet}]$, $\#\mathbb{P}_\bullet \simeq \dim \mathbb{V}_\bullet$
- Concept of ‘multilevel refinement’ $\rightsquigarrow \mathbb{P}_\circ = \mathbb{REFINE}(\mathbb{P}_\bullet, \mathbb{M}_\bullet)$

Optimal convergence of adaptive multilevel SGFEM

[B., Praetorius, Ruggeri; IMANUM (appeared online)]

- Concept of 'multilevel structure' $\rightsquigarrow$ $\mathbb{P}_\bullet = [\mathcal{P}_\bullet, (\mathcal{T}_{\bullet\nu})_{\nu\in\mathcal{P}_\bullet}]$, $\#\mathbb{P}_\bullet \simeq \dim \mathbb{V}_\bullet$
- Concept of 'multilevel refinement' $\rightsquigarrow$ $\mathbb{P}_\circ = \mathbb{REFINE}(\mathbb{P}_\bullet, \mathbb{M}_\bullet)$
- Concept of optimality $\rightsquigarrow$ approximation class $\mathbb{A}_s$ $(s > 0)$

$$w \in \mathbb{A}_s \quad \Longleftrightarrow \quad \exists \left\{\mathbb{P}^\star_\ell\right\}_{\ell\in\mathbb{N}_0} \text{ such that } |||\, w - w^\star_\ell \,||| = \mathcal{O}\big((\dim \mathbb{V}^\star_\ell)^{-s}\big)$$

Optimal convergence of adaptive multilevel SGFEM

[B., Praetorius, Ruggeri; IMANUM (appeared online)]

- Concept of 'multilevel structure' $\rightsquigarrow$ $\mathbb{P}_\bullet = [\mathcal{P}_\bullet, (\mathcal{T}_{\bullet\nu})_{\nu\in\mathcal{P}_\bullet}]$, $\#\mathbb{P}_\bullet \simeq \dim \mathbb{V}_\bullet$
- Concept of 'multilevel refinement' $\rightsquigarrow$ $\mathbb{P}_\circ = \mathbb{REFINE}(\mathbb{P}_\bullet, \mathbb{M}_\bullet)$
- Concept of optimality $\rightsquigarrow$ approximation class $\mathbb{A}_s$ $(s > 0)$

$$w \in \mathbb{A}_s \quad \Longleftrightarrow \quad \exists \left\{\mathbb{P}_\ell^\star\right\}_{\ell\in\mathbb{N}_0} \text{ such that } |||\, w - w_\ell^\star \,||| = \mathcal{O}\big((\dim \mathbb{V}_\ell^\star)^{-s}\big)$$

Theorem 4 (rate optimality)

Strong saturation assumption $\Longrightarrow$ optimal convergence for sufficiently small θ, i.e., if $s > 0$ and $u \in \mathbb{A}_s$, then

$$\sup_{\ell\in\mathbb{N}_0} (\#\mathbb{P}_\ell - \#\mathbb{P}_0 + 1)^s |||\, u - u_\ell \,||| \le C\, \|u\|_{\mathbb{A}_s},$$

where $\|u\|_{\mathbb{A}_s} := \sup\limits_{N\in\mathbb{N}_0} (N+1)^s \min\limits_{\substack{\mathbb{P}_{\rm opt}\in\mathbb{REFINE}(\mathbb{P}_0)\\ \#\mathbb{P}_{\rm opt}-\#\mathbb{P}_0\le N}} \min\limits_{v_{\rm opt}\in\mathbb{V}_{\rm opt}} |||\, u - v_{\rm opt} \,|||$

Saturation assumption vs. strong saturation assumption

- Saturation assumption
  - SGFEM solution $u_\bullet \in \mathbb{V}_\bullet$
  - Enhanced (uniformly refined) SGFEM solution $\widehat{u}_\bullet \in \widehat{\mathbb{V}}_\bullet$
  - There exist a constant $q_{\rm sat} \in (0,1)$ s.t. $|||\, u - \widehat{u}_\bullet \,||| \le q_{\rm sat} |||\, u - u_\bullet \,|||$
  - $\widehat{\mathbb{V}}_\bullet = \widehat{\mathbb{X}}_\bullet \otimes \widehat{\mathbb{P}}_\bullet \Longrightarrow q_{\rm sat} = q_{{\rm sat},\mathbb{X}} \cdot q_{{\rm sat},\mathbb{P}} \in (0,1)$ if $q_{{\rm sat},\mathbb{X}} < 1$ or $q_{{\rm sat},\mathbb{P}} < 1$

Saturation assumption vs. strong saturation assumption

- Saturation assumption
  - SGFEM solution $u_\bullet \in \mathbb{V}_\bullet$
  - Enhanced (uniformly refined) SGFEM solution $\widehat{u}_\bullet \in \widehat{\mathbb{V}}_\bullet$
  - There exist a constant $q_{\rm sat} \in (0,1)$ s.t. $|||\, u - \widehat{u}_\bullet \,||| \le q_{\rm sat} |||\, u - u_\bullet \,|||$
  - $\widehat{\mathbb{V}}_\bullet = \widehat{\mathbb{X}}_\bullet \otimes \widehat{\mathbb{P}}_\bullet \Longrightarrow q_{\rm sat} = q_{{\rm sat},\mathbb{X}} \cdot q_{{\rm sat},\mathbb{P}} \in (0,1)$ if $q_{{\rm sat},\mathbb{X}} < 1$ or $q_{{\rm sat},\mathbb{P}} < 1$
- Strong saturation assumption
  - $\mathbb{P}_\bullet \in \mathbb{REFINE}(\mathbb{P}_0) \rightsquigarrow$ any multilevel structure obtained from $\mathbb{P}_0$
  - $\mathbb{P}_\star \in \mathbb{REFINE}(\mathbb{P}_\bullet) \rightsquigarrow$ a refined multilevel structure obtained from $\mathbb{P}_\bullet$
  - $\mathbb{P}_\circ := \mathbb{REFINE}(\mathbb{P}_\bullet, \mathbb{M}_\bullet) \rightsquigarrow$ a multilevel structure obtained from $\mathbb{P}_\bullet$ by one step of multilevel refinement towards $\mathbb{P}_\star$
  - There exist constants $0 < \kappa_{\rm sat} \le q_{\rm sat} < 1$ such that

$$|||\, u - u_\star \,||| \le \kappa_{\rm sat} \,|||\, u - u_\bullet \,||| \quad \Longrightarrow \quad |||\, u - u_\circ \,||| \le q_{\rm sat} \,|||\, u - u_\bullet \,|||$$

Multilevel SGFEM: implementation challenges

- linear algebra

$$\mathbf{A} = (\mathbf{A}_{\nu\mu})_{\nu,\mu\in\mathcal{P}_\bullet}, \quad \mathbf{A}_{\nu\mu} = \sum_{m=0}^{\infty} [G_m]_{\nu\mu} K_m^{\nu\mu} = \sum_{m=0}^{M} [G_m]_{\nu\mu} K_m^{\nu\mu}$$

with

$$[G_m]_{\nu\mu} = \begin{cases} \delta_{\nu\mu} & \text{if } m = 0, \\ \displaystyle\int_\Gamma y_m P_\mu(\mathbf{y}) P_\nu(\mathbf{y})\, d\pi(\mathbf{y}) & \text{if } m \in \mathbb{N}, \end{cases}$$

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla\varphi_{\bullet\mu,z_j}(x) \cdot \nabla\varphi_{\bullet\nu,z_i}(x)\, dx$$

Multilevel SGFEM: implementation challenges

- computation of non-square stiffness matrices

Multilevel SGFEM: implementation challenges

- computation of non-square stiffness matrices

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla \varphi_{\bullet\mu,z_j}(x) \cdot \nabla \varphi_{\bullet\nu,z_i}(x)\, dx$$

$$= \sum_{T_\nu \in \mathcal{T}_{\bullet\nu}} \sum_{\substack{T_\mu \in \mathcal{T}_{\bullet\mu} \\ |T_\mu \cap T_\nu| \neq 0}} \left(\nabla \varphi_{\bullet\mu,z_j}|_{T_\mu} \cdot \nabla \varphi_{\bullet\nu,z_i}|_{T_\nu} \right) \int_{T_\mu \cap T_\nu} a_m(x)\, dx$$

Multilevel SGFEM: implementation challenges

- computation of non-square stiffness matrices

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla \varphi_{\bullet\mu,z_j}(x) \cdot \nabla \varphi_{\bullet\nu,z_i}(x)\, dx$$

$$= \sum_{T_\nu \in \mathcal{T}_{\bullet\nu}} \sum_{\substack{T_\mu \in \mathcal{T}_{\bullet\mu} \\ |T_\mu \cap T_\nu| \neq 0}} \left(\nabla \varphi_{\bullet\mu,z_j}|_{T_\mu} \cdot \nabla \varphi_{\bullet\nu,z_i}|_{T_\nu} \right) \int_{T_\mu \cap T_\nu} a_m(x)\, dx$$

- key observation: $\mathcal{T}_{\bullet,\nu}$ and $\mathcal{T}_{\bullet,\mu}$ are NVB refinements of the same $\mathcal{T}_0$;
  $\rightsquigarrow$ exploit the binary tree structure of the NVB refinement

Multilevel SGFEM: implementation challenges

- computation of non-square stiffness matrices

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla\varphi_{\bullet\mu,z_j}(x) \cdot \nabla\varphi_{\bullet\nu,z_i}(x)\, dx$$

$$= \sum_{T_\nu \in \mathcal{T}_{\bullet\nu}} \sum_{\substack{T_\mu \in \mathcal{T}_{\bullet\mu} \\ |T_\mu \cap T_\nu| \neq 0}} \left(\nabla\varphi_{\bullet\mu,z_j}|_{T_\mu} \cdot \nabla\varphi_{\bullet\nu,z_i}|_{T_\nu} \right) \int_{T_\mu \cap T_\nu} a_m(x)\, dx$$

- key observation: $\mathcal{T}_{\bullet,\nu}$ and $\mathcal{T}_{\bullet,\mu}$ are NVB refinements of the same $\mathcal{T}_0$;
$\rightsquigarrow$ exploit the binary tree structure of the NVB refinement

- $T_\mu \cap T_\nu$ is either T_μ, or T_ν, or a set of measure zero.

Multilevel SGFEM: implementation challenges

- computation of non-square stiffness matrices

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla \varphi_{\bullet\mu, z_j}(x) \cdot \nabla \varphi_{\bullet\nu, z_i}(x)\, dx$$

$$= \sum_{T_\nu \in \mathcal{T}_{\bullet\nu}} \sum_{\substack{T_\mu \in \mathcal{T}_{\bullet\mu} \\ |T_\mu \cap T_\nu| \neq 0}} \left(\nabla \varphi_{\bullet\mu, z_j}|_{T_\mu} \cdot \nabla \varphi_{\bullet\nu, z_i}|_{T_\nu} \right) \int_{T_\mu \cap T_\nu} a_m(x)\, dx$$

- key observation: $\mathcal{T}_{\bullet,\nu}$ and $\mathcal{T}_{\bullet,\mu}$ are NVB refinements of the same $\mathcal{T}_0$;
  $\rightsquigarrow$ exploit the binary tree structure of the NVB refinement

- $T_\mu \cap T_\nu$ is either T_μ, or T_ν, or a set of measure zero.

- naive implementation $\rightsquigarrow$ $\mathcal{O}((\#\mathcal{T}_{\bullet\nu})(\#\mathcal{T}_{\bullet\mu}))$ complexity, BUT...

Multilevel SGFEM: implementation challenges

- computation of non-square stiffness matrices

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla \varphi_{\bullet\mu, z_j}(x) \cdot \nabla \varphi_{\bullet\nu, z_i}(x)\, dx$$

$$= \sum_{T_\nu \in \mathcal{T}_{\bullet\nu}} \sum_{\substack{T_\mu \in \mathcal{T}_{\bullet\mu} \\ |T_\mu \cap T_\nu| \neq 0}} \left(\nabla \varphi_{\bullet\mu, z_j}|_{T_\mu} \cdot \nabla \varphi_{\bullet\nu, z_i}|_{T_\nu} \right) \int_{T_\mu \cap T_\nu} a_m(x)\, dx$$

- key observation: $\mathcal{T}_{\bullet,\nu}$ and $\mathcal{T}_{\bullet,\mu}$ are NVB refinements of the same $\mathcal{T}_0$;
  $\rightsquigarrow$ exploit the binary tree structure of the NVB refinement

- $T_\mu \cap T_\nu$ is either T_μ, or T_ν, or a set of measure zero.

- naive implementation $\rightsquigarrow$ $\mathcal{O}((\#\mathcal{T}_{\bullet\nu})(\#\mathcal{T}_{\bullet\mu}))$ complexity, BUT...

- the binary search tree algorithm $\rightsquigarrow$ $\mathcal{O}(\#\mathcal{T}_{\bullet\nu} + \#\mathcal{T}_{\bullet\mu})$ complexity (tbc...)

Multilevel SGFEM: implementation challenges

- linear algebra

$$\mathbf{A} = (\mathbf{A}_{\nu\mu})_{\nu,\mu\in\mathcal{P}_\bullet}, \quad \mathbf{A}_{\nu\mu} = \sum_{m=0}^{\infty} [G_m]_{\nu\mu} K_m^{\nu\mu} = \sum_{m=0}^{M} [G_m]_{\nu\mu} K_m^{\nu\mu}$$

with

$$[G_m]_{\nu\mu} = \begin{cases} \delta_{\nu\mu} & \text{if } m = 0, \\ \int_\Gamma y_m P_\mu(\mathbf{y}) P_\nu(\mathbf{y})\, d\pi(\mathbf{y}) & \text{if } m \in \mathbb{N}, \end{cases}$$

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla\varphi_{\bullet\mu,z_j}(x) \cdot \nabla\varphi_{\bullet\nu,z_i}(x)\, dx$$

- solver: preconditioned MINRES or CG

Multilevel SGFEM: implementation challenges

- linear algebra

$$\mathbf{A} = (\mathbf{A}_{\nu\mu})_{\nu,\mu\in\mathfrak{P}_\bullet}, \quad \mathbf{A}_{\nu\mu} = \sum_{m=0}^{\infty} [G_m]_{\nu\mu} K_m^{\nu\mu} = \sum_{m=0}^{M} [G_m]_{\nu\mu} K_m^{\nu\mu}$$

with

$$[G_m]_{\nu\mu} = \begin{cases} \delta_{\nu\mu} & \text{if } m = 0, \\ \int_\Gamma y_m P_\mu(\mathbf{y}) P_\nu(\mathbf{y})\, d\pi(\mathbf{y}) & \text{if } m \in \mathbb{N}, \end{cases}$$

$$[K_m^{\nu\mu}]_{ij} = \int_D a_m(x) \nabla\varphi_{\bullet\mu,z_j}(x) \cdot \nabla\varphi_{\bullet\nu,z_i}(x)\, dx$$

- solver: preconditioned MINRES or CG
- preconditioner: mean-based preconditioner with diagonal blocks $K_0^{\nu\nu}$

Concluding remarks

What have we achieved?

- Design of adaptive algorithms for single-level and multilevel SGFEM
- Theoretical analysis
  - convergence of the adaptive algorithm
  - linear convergence (under the saturation assumption)
  - rate-optimality of adaptive multilevel SGFEM (under the strong saturation assumption)
- Implementation: Stochastic T-IFISS
  `http://web.mat.bham.ac.uk/A.Bespalov/software/`

Concluding remarks

What have we achieved?

- Design of adaptive algorithms for single-level and multilevel SGFEM
- Theoretical analysis
  - convergence of the adaptive algorithm
  - linear convergence (under the saturation assumption)
  - rate-optimality of adaptive multilevel SGFEM (under the strong saturation assumption)
- Implementation: Stochastic T-IFISS
  `http://web.mat.bham.ac.uk/A.Bespalov/software/`

References

- AB, D. Praetorius, L. Rocchi, M. Ruggeri, *Convergence of adaptive stochastic Galerkin FEM*, SIAM J. Numer. Anal., Vol. 57 (2019), pp. 2359–2382.
- AB, D. Praetorius, M. Ruggeri, *Two-level a posteriori error estimation for adaptive multilevel stochastic Galerkin FEM*, SIAM/ASA J. Uncertain. Quantif., Vol.9 (2021), pp. 1184–1216.
- AB, D. Praetorius, M. Ruggeri, *Convergence and rate optimality of adaptive multilevel stochastic Galerkin FEM*, IMA J. Numer. Anal. (appeared online).