A POSTERIORI ERROR ESTIMATION FOR STOCHASTIC GALERKIN APPROXIMATIONS

Alex Bespalov

School of Mathematics, University of Birmingham,

Edgbaston, Birmingham, United Kingdom

Joint work with **Catherine Powell** and **David Silvester**,

University of Manchester, United Kingdom

British Computational PDEs Colloquium: New Trends

January 23 – 24, 2014

ICMS, Edinburgh

What is this talk about...

* PDE problems with random data
  (or, more generally, parametric operator equations)
* Stochastic Galerkin finite element method
* A posteriori error estimation and h-p adaptivity
* A proof-of-concept implementation

Introduction: modelling uncertainty

Generic PDE-type problem:

$$\begin{aligned} Lu &= f \qquad \text{in} \quad D \subset \mathbf{R}^d, \\ Bu &= g \qquad \text{on} \quad \partial D. \end{aligned}$$

If the data (i.e., the domain D, coefficients in L and B, the source term f and boundary conditions) are known precisely, then we can compute numerical approximation $\tilde{u} \approx u$ to high accuracy (e.g., using FEM, BEM, ...).

Introduction: modelling uncertainty

Generic PDE-type problem:

$$\begin{aligned} Lu &= f \qquad \text{in } \; D \subset \mathbf{R}^d, \\ Bu &= g \qquad \text{on } \; \partial D. \end{aligned}$$

If the data (i.e., the domain D, coefficients in L and B, the source term f and boundary conditions) are known precisely, then we can compute numerical approximation $\tilde{u} \approx u$ to high accuracy (e.g., using FEM, BEM, ...).

In practice: we do not have access to all input data (*systematic* uncertainty).

Examples:

- permeability coefficients of a porous medium;
- modulus of elasticity of a stressed body;
- width of a channel in flow problems.

Introduction: modelling uncertainty

Generic PDE-type problem:

$$\begin{aligned} Lu &= f \qquad \text{in } \ D \subset \mathbf{R}^d, \\ Bu &= g \qquad \text{on } \ \partial D. \end{aligned}$$

If the data (i.e., the domain D, coefficients in L and B, the source term f and boundary conditions) are known precisely, then we can compute numerical approximation $\tilde{u} \approx u$ to high accuracy (e.g., using FEM, BEM, ...).

In practice: we do not have access to all input data (*systematic* uncertainty).

Mathematical model: consider data as random parameters/fields.

Introduction: modelling uncertainty

Generic PDE-type problem:

$$\begin{aligned} Lu &= f \qquad \text{in} \quad D \subset \mathbf{R}^d, \\ Bu &= g \qquad \text{on} \quad \partial D. \end{aligned}$$

If the data (i.e., the domain D, coefficients in L and B, the source term f and boundary conditions) are known precisely, then we can compute numerical approximation $\tilde{u} \approx u$ to high accuracy (e.g., using FEM, BEM, ...).

Randomness in the input data $\implies$ Randomness in the output (solution).

Questions to answer:

- What is the expected solution?
- What is the uncertainty in the solution?
- What is the probability of an (un)desirable event?

Elliptic PDEs with random data

Generic operator equation

$$A(\boldsymbol{\xi})\, u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}),$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$ is a multivariate random variable.

Elliptic PDEs with random data

Generic operator equation

$$A(\boldsymbol{\xi})\, u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}),$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$ is a multivariate random variable.

Assumptions and notation:

- $(\Omega, \mathcal{F}, \mathbb{P})$ – complete probability space;
- $\xi_m : \Omega \to \Gamma_m \subset \mathbb{R}$ are bounded real-valued random variables;

Elliptic PDEs with random data

Generic operator equation

$$A(\boldsymbol{\xi})\, u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}),$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$ is a multivariate random variable.

Assumptions and notation:

- $(\Omega, \mathcal{F}, \mathbb{P})$ – complete probability space;
- $\xi_m : \Omega \to \Gamma_m \subset \mathbb{R}$ are bounded real-valued random variables;
- $\boldsymbol{\xi}$ generates a distribution π satisfying $\pi(G) = \mathbb{P}(\boldsymbol{\xi}^{-1}(G))\ \forall\, G \in \mathcal{B}(\Gamma)$; π is also a probability measure (by definition);
- the observations $\mathbf{y}$ of $\boldsymbol{\xi}$ belong to $\Gamma := \prod_{m=1}^{\infty} \Gamma_m$ and $(\Gamma, \mathcal{B}(\Gamma), \pi)$ is a probability space;

Elliptic PDEs with random data

Generic operator equation

$$A(\boldsymbol{\xi})\, u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}),$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$ is a multivariate random variable.

Assumptions and notation:

- $(\Omega, \mathcal{F}, \mathbb{P})$ – complete probability space;
- $\xi_m : \Omega \to \Gamma_m \subset \mathbb{R}$ are bounded real-valued random variables;
- $\boldsymbol{\xi}$ generates a distribution π satisfying $\pi(G) = \mathbb{P}(\boldsymbol{\xi}^{-1}(G))\ \forall\, G \in \mathcal{B}(\Gamma)$;
  π is also a probability measure (by definition);
- the observations $\mathbf{y}$ of $\boldsymbol{\xi}$ belong to $\Gamma := \prod_{m=1}^{\infty} \Gamma_m$ and
  $(\Gamma, \mathcal{B}(\Gamma), \pi)$ is a probability space;
- $\xi_1, \xi_2, \ldots$ are independent $\implies \pi(\mathbf{y}) = \prod_{m=1}^{\infty} \pi_m(y_m)$,
  where π_m is the probability measure associated with ξ_m $(m = 1, 2, \ldots)$.

Elliptic PDEs with random data

Generic operator equation $\iff$ *Parametric operator equation*

$$A(\boldsymbol{\xi})\,u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}) \quad \Longleftrightarrow \quad A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}) \quad \forall\,\mathbf{y} \in \Gamma = \prod_{m=1}^{\infty} \Gamma_m.$$

Assumptions and notation:

- $(\Omega, \mathcal{F}, \mathbb{P})$ – complete probability space;
- $\xi_m : \Omega \to \Gamma_m \subset \mathbb{R}$ are bounded real-valued random variables;
- $\boldsymbol{\xi}$ generates a distribution π satisfying $\pi(G) = \mathbb{P}(\boldsymbol{\xi}^{-1}(G))\ \forall\, G \in \mathcal{B}(\Gamma)$; π is also a probability measure (by definition);
- the observations $\mathbf{y}$ of $\boldsymbol{\xi}$ belong to $\Gamma := \prod_{m=1}^{\infty} \Gamma_m$ and $(\Gamma, \mathcal{B}(\Gamma), \pi)$ is a probability space;
- $\xi_1, \xi_2, \ldots$ are independent $\implies \pi(\mathbf{y}) = \prod_{m=1}^{\infty} \pi_m(y_m)$, where π_m is the probability measure associated with ξ_m $(m = 1, 2, \ldots)$.

Parametric operator equation

$$A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}) \quad \forall\,\mathbf{y} \in \Gamma = \prod_{m=1}^{\infty} \Gamma_m,$$

where $A : \Gamma \to \mathcal{L}(H, H')$, $f : \Gamma \to H'$, H is a separable Hilbert space.

Parametric operator equation

$$A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}) \quad \forall\,\mathbf{y} \in \Gamma = \prod_{m=1}^{\infty} \Gamma_m,$$

where $A : \Gamma \to \mathcal{L}(H, H')$, $f : \Gamma \to H'$, H is a separable Hilbert space.

Solution space: $V := L^2_\pi(\Gamma; H)$ with the norm $\|\cdot\|_V := \left(\int_\Gamma \|\cdot\|_H^2 \,\mathrm{d}\pi(\mathbf{y})\right)^{1/2}$.

Weak formulation: find $u \in V$ such that

$$\underbrace{\int_\Gamma \langle A(\mathbf{y})u(\mathbf{y}), v(\mathbf{y})\rangle \,\mathrm{d}\pi(\mathbf{y})}_{B(u,v)} = \underbrace{\int_\Gamma \langle f(\mathbf{y}), v(\mathbf{y})\rangle \,\mathrm{d}\pi(\mathbf{y})}_{F(v)} \qquad \forall\, v \in V. \qquad (\star)$$

Motivating example

Find $u(\mathbf{x}, \mathbf{y})$ satisfying

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

where $D \subset \mathbb{R}^d$ $(d = 2, 3)$, $\Gamma := \prod_{m=1}^{\infty} [-1, 1]$ and

$$a(\mathbf{x}, \mathbf{y}) = a_0(\mathbf{x}) + \sum_{m=1}^{\infty} a_m(\mathbf{x})\, y_m, \qquad \mathbf{x} \in D, \;\; \mathbf{y} \in \Gamma.$$

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

Parametric operator equation: decomposition and well posedness

$$A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}), \quad \text{where} \quad A(\mathbf{y}) = A_0 + \sum_{m=1}^{\infty} y_m A_m \quad \forall\, \mathbf{y} \in \Gamma.$$

Parametric operator equation: decomposition and well posedness

$$A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}), \quad \text{where} \quad A(\mathbf{y}) = A_0 + \sum_{m=1}^{\infty} y_m A_m \quad \forall\, \mathbf{y} \in \Gamma.$$

Weak formulation: find $u \in V := L^2_\pi(\Gamma; H)$ such that

$$B(u, v) = F(v) \qquad \forall\, v \in V, \tag{$\star$}$$

where

$$B(u,v) = \int_\Gamma \langle A(\mathbf{y})u, v\rangle \,\mathrm{d}\pi(\mathbf{y}) = \int_\Gamma \langle A_0 u, v\rangle \,\mathrm{d}\pi(\mathbf{y}) + \sum_{m=1}^{\infty} \int_\Gamma \langle A_m u, v\rangle y_m \mathrm{d}\pi(\mathbf{y})$$

$$=: B_0(u,v) + \sum_{m=1}^{\infty} B_m(u,v) \quad \forall\, u, v \in V,$$

$$F(v) = \int_\Gamma \langle f(\mathbf{y}), v(\mathbf{y})\rangle \,\mathrm{d}\pi(\mathbf{y}).$$

Parametric operator equation: decomposition and well posedness

$$A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}), \quad \text{where} \quad A(\mathbf{y}) = A_0 + \sum_{m=1}^{\infty} y_m A_m \quad \forall\, \mathbf{y} \in \Gamma.$$

Weak formulation: find $u \in V := L^2_\pi(\Gamma; H)$ such that

$$B_0(u,v) + \sum_{m=1}^{\infty} B_m(u,v) = F(v) \qquad \forall\, v \in V. \tag{$\star$}$$

Assumptions ensuring the well posedness of $(\star)$*:*

- $A_0 \in \mathcal{L}(H, H')$ is symmetric and positive definite;
- $A_m \in \mathcal{L}(H, H')$ for $(m \in \mathbb{N})$ are symmetric and $\exists\, \tau \in [0,1)$ s.t. $\forall\, \mathbf{y} \in \Gamma$

$$\left| \left\langle \sum_{m=1}^{\infty} y_m A_m v, v \right\rangle \right| \leq \tau \left\langle A_0 v, v \right\rangle \quad \forall\, v \in H$$

(cf. [Schwab, Gittelson '11]).

Parametric operator equation: norm equivalence

$$A(\mathbf{y})\,u(\mathbf{y}) = f(\mathbf{y}), \quad \text{where} \quad A(\mathbf{y}) = A_0 + \sum_{m=1}^{\infty} y_m A_m \quad \forall\, \mathbf{y} \in \Gamma.$$

Weak formulation: find $u \in V := L^2_\pi(\Gamma; H)$ such that

$$\underbrace{B_0(u,v) + \sum_{m=1}^{\infty} B_m(u,v)}_{B(u,v)} = F(v) \qquad \forall\, v \in V. \qquad (\star)$$

Norm equivalence

Note that $(B(u,u))^{1/2} = \|u\|_B$ is equivalent to the B_0 norm: there exist positive constants $\lambda < 1 < \Lambda$ such that

$$\lambda\, B(v,v) \le B_0(v,v) \le \Lambda\, B(v,v) \qquad \forall\, v \in V. \qquad (B_0)$$

Discrete formulation

Solution space: $V := L^2_\pi(\Gamma; H) \sim H \otimes L^2_\pi(\Gamma)$ (isometric isomorphism).

Finite-dimensional subspace: $X \otimes \mathcal{P}_{\mathfrak{P}} =: V_{X\mathfrak{P}} \subset V$

Discrete formulation

Solution space: $V := L^2_\pi(\Gamma; H) \sim H \otimes L^2_\pi(\Gamma)$ (isometric isomorphism).

Finite-dimensional subspace: $X \otimes \mathcal{P}_{\mathfrak{P}} =: V_{X\mathfrak{P}} \subset V$, where

- X is a finite-dimensional subspace of H
  (corresponding, e.g., to spatial finite element discretization on D);
- $\mathcal{P}_{\mathfrak{P}} \subset L^2_\pi(\Gamma)$ is a polynomial space on Γ, more precisely,
  the set of tensor product polynomials associated with a finite subset $\mathfrak{P}$ of the index set $\mathfrak{J}$ of finitely supported sequences

$$\mathfrak{J} := \big\{\nu = (\nu_1, \nu_2, \ldots) \in \mathbb{N}_0^{\mathbb{N}};\ \#\operatorname{supp}\nu < \infty\big\},$$

  where $\operatorname{supp}\nu := \{m \in \mathbb{N};\ \nu_m \neq 0\}$ for any $\nu \in \mathbb{N}_0^{\mathbb{N}}$.

Discrete formulation

Solution space: $V := L^2_\pi(\Gamma; H) \sim H \otimes L^2_\pi(\Gamma)$ (isometric isomorphism).

Finite-dimensional subspace: $X \otimes \mathcal{P}_{\mathfrak{P}} =: V_{X\mathfrak{P}} \subset V$, where

- X is a finite-dimensional subspace of H
  (corresponding, e.g., to spatial finite element discretization on D);
- $\mathcal{P}_{\mathfrak{P}} \subset L^2_\pi(\Gamma)$ is a polynomial space on Γ, more precisely,
  the set of tensor product polynomials associated with a finite subset $\mathfrak{P}$ of the index set $\mathfrak{J}$ of finitely supported sequences

$$\mathfrak{J} := \big\{\nu = (\nu_1, \nu_2, \ldots) \in \mathbb{N}_0^{\mathbb{N}};\ \#\operatorname{supp}\nu < \infty\big\},$$

  where $\operatorname{supp}\nu := \{m \in \mathbb{N};\ \nu_m \neq 0\}$ for any $\nu \in \mathbb{N}_0^{\mathbb{N}}$.

Galerkin projection: find $u_{X\mathfrak{P}} \in V_{X\mathfrak{P}} = X \otimes \mathcal{P}_{\mathfrak{P}}$ such that

$$B(u_{X\mathfrak{P}}, v) = F(v) \quad \forall\, v \in V_{X\mathfrak{P}}.$$

Stochastic Galerkin FEM

- Approximation is sought in the form

$$u_{X\mathfrak{P}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{N_{\mathfrak{P}}} \sum_{j=1}^{N_X} \alpha_{ij}\, \varphi_j(\mathbf{x})\, \psi_i(\mathbf{y}), \quad \mathbf{x} \in D \subset \mathbf{R}^d, \quad \mathbf{y} \in \Gamma^M \subset \mathbf{R}^M,$$

  where $\{\varphi_j(\mathbf{x})\}$ – standard FE basis, $\{\psi_i(\mathbf{y})\}$ – orthogonal polynomials.

- Coefficients α_{ij} are found from Galerkin system

$$B(u_{X\mathfrak{P}},\ \varphi_j\,\psi_i) = F(\varphi_j\,\psi_i), \quad j = 1,\ldots,N_X, \quad i = 1,\ldots,N_{\mathfrak{P}},$$

  which is one linear system of $N_{\mathfrak{P}} \cdot N_X$ equations.

Stochastic Galerkin FEM

- Approximation is sought in the form

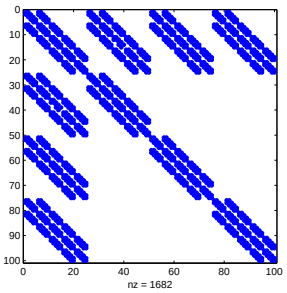
$$u_{X\mathfrak{P}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{N_{\mathfrak{P}}} \sum_{j=1}^{N_X} \alpha_{ij}\, \varphi_j(\mathbf{x})\, \psi_i(\mathbf{y}), \quad \mathbf{x} \in D \subset \mathbf{R}^d, \quad \mathbf{y} \in \Gamma^M \subset \mathbf{R}^M,$$

  where $\{\varphi_j(\mathbf{x})\}$ – standard FE basis, $\{\psi_i(\mathbf{y})\}$ – orthogonal polynomials.

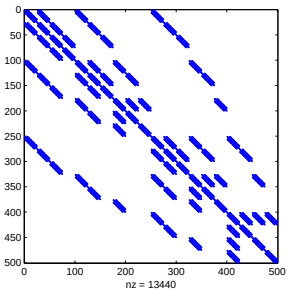
- Coefficients α_{ij} are found from Galerkin system

$$B(u_{X\mathfrak{P}},\ \varphi_j\,\psi_i) = F(\varphi_j\,\psi_i), \quad j = 1,\ldots,N_X, \quad i = 1,\ldots,N_{\mathfrak{P}},$$

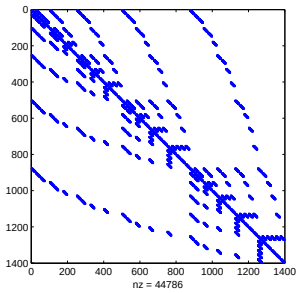
  which is one linear system of $N_{\mathfrak{P}} \cdot N_X$ equations.



$M = 3,\ k = 1$



$M = 3,\ k = 3$



$M = 5,\ k = 3$

Stochastic Galerkin FEM

- Approximation is sought in the form

$$u_{X\mathfrak{P}}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{N_{\mathfrak{P}}} \sum_{j=1}^{N_X} \alpha_{ij}\, \varphi_j(\mathbf{x})\, \psi_i(\mathbf{y}), \quad \mathbf{x} \in D \subset \mathbf{R}^d, \quad \mathbf{y} \in \Gamma^M \subset \mathbf{R}^M,$$

  where $\{\varphi_j(\mathbf{x})\}$ – standard FE basis, $\{\psi_i(\mathbf{y})\}$ – orthogonal polynomials.

- Coefficients α_{ij} are found from Galerkin system

$$B(u_{X\mathfrak{P}},\ \varphi_j\,\psi_i) = F(\varphi_j\,\psi_i), \quad j = 1,\ldots,N_X, \quad i = 1,\ldots,N_{\mathfrak{P}},$$

  which is one linear system of $N_{\mathfrak{P}} \cdot N_X$ equations.

- For polynomials of total degree k in M variables, $N_k = \frac{(M+k)!}{M!\,k!}$;

M = 5			M = 10		
k=2	k=4	k=6	k=2	k=4	k=6
21	126	462	66	1001	8008

Stochastic Galerkin FEM

Conclusion:

If stochastic Galerkin methods are to be competitive with traditional deterministic methodologies based on sampling (e.g., Monte-Carlo) one needs:

- special algorithms to construct approximations (e.g., in adaptive way);
- fast and robust linear algebra techniques to solve large linear systems.

Stochastic Galerkin FEM

Conclusion:

If stochastic Galerkin methods are to be competitive with traditional deterministic methodologies based on sampling (e.g., Monte-Carlo) one needs:

- special algorithms to construct approximations (e.g., in adaptive way);
- fast and robust linear algebra techniques to solve large linear systems.

Prerequisite to adaptivity:

A posteriori error estimation (the rest of this talk).

Error estimation I

$$u \in V: \qquad B(u,v) = F(v) \quad \forall\, v \in V. \tag{$\star$}$$

$$u_{X\mathfrak{P}} \in V_{X\mathfrak{P}} = X \otimes \mathcal{P}_{\mathfrak{P}}: \qquad B(u_{X\mathfrak{P}},v) = F(v) \quad \forall\, v \in V_{X\mathfrak{P}}.$$

Thus, the error $e := u - u_{X\mathfrak{P}} \in V$ satisfies

$$B(e,v) = \underbrace{F(v) - B(u_{X\mathfrak{P}},v)}_{=\ 0\ \forall\, v \in V_{X\mathfrak{P}}} \qquad \forall v \in V.$$

Error estimation I

$$u \in V: \qquad B(u,v) = F(v) \quad \forall\, v \in V. \tag{$\star$}$$

$$u_{X\mathfrak{P}} \in V_{X\mathfrak{P}} = X \otimes \mathcal{P}_{\mathfrak{P}}: \qquad B(u_{X\mathfrak{P}}, v) = F(v) \quad \forall\, v \in V_{X\mathfrak{P}}.$$

Thus, the error $e := u - u_{X\mathfrak{P}} \in V$ satisfies

$$B(e,v) = \underbrace{F(v) - B(u_{X\mathfrak{P}}, v)}_{=\ 0\ \forall\, v \in V_{X\mathfrak{P}}} \qquad \forall v \in V.$$

This motivates the estimator $e^* \in V^*_{X\mathfrak{P}}$ satisfying

$$B(e^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \qquad \forall\, v \in V^*_{X\mathfrak{P}},$$

where $V^*_{X\mathfrak{P}} \supset V_{X\mathfrak{P}}$ is an enriched finite-dimensional subspace of V, that is

$$V^*_{X\mathfrak{P}} = X^* \otimes \mathcal{P}_{\mathfrak{P}^*} \quad \text{with} \quad X^* \supset X, \quad \mathcal{P}_{\mathfrak{P}^*} \supset \mathcal{P}_{\mathfrak{P}}.$$

Error estimation I

$$u \in V: \qquad B(u,v) = F(v) \quad \forall\, v \in V. \tag{$\star$}$$

$$u_{X\mathfrak{P}} \in V_{X\mathfrak{P}} = X \otimes \mathcal{P}_{\mathfrak{P}}: \qquad B(u_{X\mathfrak{P}},v) = F(v) \quad \forall\, v \in V_{X\mathfrak{P}}.$$

Thus, the error $e := u - u_{X\mathfrak{P}} \in V$ satisfies

$$B(e,v) = \underbrace{F(v) - B(u_{X\mathfrak{P}},v)}_{=\ 0\ \forall\, v \in V_{X\mathfrak{P}}} \qquad \forall v \in V.$$

This motivates the estimator $e^* \in V^*_{X\mathfrak{P}}$ satisfying

$$B(e^*,v) = F(v) - B(u_{X\mathfrak{P}},v) \qquad \forall\, v \in V^*_{X\mathfrak{P}},$$

where $V^*_{X\mathfrak{P}} \supset V_{X\mathfrak{P}}$ is an enriched finite-dimensional subspace of V, that is

$$V^*_{X\mathfrak{P}} = X^* \otimes \mathcal{P}_{\mathfrak{P}^*} \ \text{ with } \ X^* \supset X, \quad \mathcal{P}_{\mathfrak{P}^*} \supset \mathcal{P}_{\mathfrak{P}}.$$

Is the estimator e^* practical? No.

Error estimation II

The first error estimator $e^* \in V^*_{X\mathfrak{P}}$ satisfies

$$B(e^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \qquad \forall\, v \in V^*_{X\mathfrak{P}}.$$

Recall that

$$B(u, v) = B_0(u, v) + \sum_{m=1}^{\infty} B_m(u, v).$$

Error estimation II

The first error estimator $e^* \in V^*_{X\mathfrak{P}}$ satisfies

$$B(e^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \qquad \forall\, v \in V^*_{X\mathfrak{P}}.$$

Recall that

$$B(u, v) = B_0(u, v) + \sum_{m=1}^{\infty} B_m(u, v).$$

This leads to the second error estimator $e_0^* \in V^*_{X\mathfrak{P}}$ satisfying

$$B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \qquad \forall\, v \in V^*_{X\mathfrak{P}}.$$

Note: the associate system matrix is block diagonal $\Longrightarrow$ efficient linear algebra!

Error estimation II

The first error estimator $e^* \in V^*_{X\mathfrak{P}}$ satisfies

$$B(e^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \qquad \forall\, v \in V^*_{X\mathfrak{P}}.$$

Recall that

$$B(u, v) = B_0(u, v) + \sum_{m=1}^{\infty} B_m(u, v).$$

This leads to the second error estimator $e_0^* \in V^*_{X\mathfrak{P}}$ satisfying

$$B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \qquad \forall\, v \in V^*_{X\mathfrak{P}}.$$

Note: the associate system matrix is block diagonal $\Longrightarrow$ efficient linear algebra!

The relation between e^* and e_0^* follows from (B_0):

$$\sqrt{\lambda}\, \|e_0^*\|_{B_0} \le \|e^*\|_B \le \sqrt{\Lambda}\, \|e_0^*\|_{B_0}. \qquad (B_0^*)$$

Error estimation III

The second estimator $e_0^* \in V_{X\mathfrak{P}}^*$: $B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{X\mathfrak{P}}^*$.

Exploit the structure of the enriched finite-dimensional space $V_{X\mathfrak{P}}^* = X^* \otimes \mathcal{P}_{\mathfrak{P}^*}$:

$$X^* = X \oplus Y, \quad X \cap Y = \{0\}; \quad \mathcal{P}_{\mathfrak{P}^*} = \mathcal{P}_{\mathfrak{P}} \oplus \mathcal{P}_{\mathfrak{Q}}, \quad \mathcal{P}_{\mathfrak{P}} \cap \mathcal{P}_{\mathfrak{Q}} = \{0\};$$

$$V_{X\mathfrak{P}}^* = (X \oplus Y) \otimes (\mathcal{P}_{\mathfrak{P}} \oplus \mathcal{P}_{\mathfrak{Q}}) = V_{X\mathfrak{P}} \oplus \underbrace{V_{X\mathfrak{Q}}}_{X \otimes \mathcal{P}_{\mathfrak{Q}}} \oplus \underbrace{V_{Y\mathfrak{P}}}_{Y \otimes \mathcal{P}_{\mathfrak{P}}} \oplus \underbrace{V_{Y\mathfrak{Q}}}_{Y \otimes \mathcal{P}_{\mathfrak{Q}}}.$$

Error estimation III

The second estimator $e_0^* \in V_{X\mathfrak{P}}^*$: $B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{X\mathfrak{P}}^*$.

Exploit the structure of the enriched finite-dimensional space $V_{X\mathfrak{P}}^* = X^* \otimes \mathcal{P}_{\mathfrak{P}^*}$:

$$X^* = X \oplus Y, \quad X \cap Y = \{0\}; \quad \mathcal{P}_{\mathfrak{P}^*} = \mathcal{P}_{\mathfrak{P}} \oplus \mathcal{P}_{\mathfrak{Q}}, \quad \mathcal{P}_{\mathfrak{P}} \cap \mathcal{P}_{\mathfrak{Q}} = \{0\};$$

$$V_{X\mathfrak{P}}^* = (X \oplus Y) \otimes (\mathcal{P}_{\mathfrak{P}} \oplus \mathcal{P}_{\mathfrak{Q}}) = V_{X\mathfrak{P}} \oplus \underbrace{V_{X\mathfrak{Q}}}_{X \otimes \mathcal{P}_{\mathfrak{Q}}} \oplus \underbrace{V_{Y\mathfrak{P}}}_{Y \otimes \mathcal{P}_{\mathfrak{P}}} \oplus \underbrace{V_{Y\mathfrak{Q}}}_{Y \otimes \mathcal{P}_{\mathfrak{Q}}}.$$

This leads to the third estimate

$$\eta := \left(2\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0}^2 + \|\bar{e}_{Y\mathfrak{P}}\|_{B_0}^2 + 2\,\|\bar{e}_{Y\mathfrak{Q}}\|_{B_0}^2\right)^{1/2},$$

where the contributing estimators are defined by

$$\begin{aligned}
\bar{e}_{X\mathfrak{Q}} \in V_{X\mathfrak{Q}}: \quad & B_0(\bar{e}_{X\mathfrak{Q}}, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{X\mathfrak{Q}}, \\
\bar{e}_{Y\mathfrak{P}} \in V_{Y\mathfrak{P}}: \quad & B_0(\bar{e}_{Y\mathfrak{P}}, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{Y\mathfrak{P}}, \\
\bar{e}_{Y\mathfrak{Q}} \in V_{Y\mathfrak{Q}}: \quad & B_0(\bar{e}_{Y\mathfrak{Q}}, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{Y\mathfrak{Q}}.
\end{aligned}$$

Error estimation III

The second estimator $e_0^* \in V_{X\mathfrak{P}}^*$: $B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{X\mathfrak{P}}^*$.

Exploit the structure of the enriched finite-dimensional space $V_{X\mathfrak{P}}^* = X^* \otimes \mathcal{P}_{\mathfrak{P}^*}$:

$$X^* = X \oplus Y, \quad X \cap Y = \{0\}; \quad \mathcal{P}_{\mathfrak{P}^*} = \mathcal{P}_{\mathfrak{P}} \oplus \mathcal{P}_{\mathfrak{Q}}, \quad \mathcal{P}_{\mathfrak{P}} \cap \mathcal{P}_{\mathfrak{Q}} = \{0\};$$

$$V_{X\mathfrak{P}}^* = (X \oplus Y) \otimes (\mathcal{P}_{\mathfrak{P}} \oplus \mathcal{P}_{\mathfrak{Q}}) = V_{X\mathfrak{P}} \oplus \underbrace{V_{X\mathfrak{Q}}}_{X \otimes \mathcal{P}_{\mathfrak{Q}}} \oplus \underbrace{V_{Y\mathfrak{P}}}_{Y \otimes \mathcal{P}_{\mathfrak{P}}} \oplus \underbrace{V_{Y\mathfrak{Q}}}_{Y \otimes \mathcal{P}_{\mathfrak{Q}}}.$$

This leads to the third estimate

$$\eta := \left(2\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0}^2 + \|\bar{e}_{Y\mathfrak{P}}\|_{B_0}^2 + 2\,\|\bar{e}_{Y\mathfrak{Q}}\|_{B_0}^2\right)^{1/2}.$$

The relation between the estimator e_0^* and the componentwise estimate η:

$$\frac{1}{\sqrt{5}}\,\eta \le \|e_0^*\|_{B_0} \le \frac{1}{\sqrt{1-\gamma^2}}\,\eta,$$

where $\gamma \in [0, 1)$ is the constant in the strengthened Cauchy–Schwarz inequality for the subspaces $X, Y \subset H$ w.r.t. inner product $\langle A_0 \cdot, \cdot \rangle$.

Error reduction indicators

Inspired by hp-adaptive algorithm in [A. Schmidt, K. G. Siebert '00].

Enhanced Galerkin approximations: $u_{X^*\mathfrak{P}} \in X^* \otimes \mathcal{P}_{\mathfrak{P}}$ and $u_{X\mathfrak{P}^*} \in X \otimes \mathcal{P}_{\mathfrak{P}^*}$.

Note: Pythagoras' theorem and Galerkin orthogonality imply (for $u_{X^*\mathfrak{P}}$)

$$\|e\|_B^2 = \|e_{X^*\mathfrak{P}}\|_B^2 + \|u_{X^*\mathfrak{P}} - u_{X\mathfrak{P}}\|_B^2, \quad \text{where } e_{X^*\mathfrak{P}} := u - u_{X^*\mathfrak{P}}.$$

Error reduction indicators

Inspired by hp-adaptive algorithm in [A. Schmidt, K. G. Siebert '00].

Enhanced Galerkin approximations: $u_{X^*\mathfrak{P}} \in X^* \otimes \mathcal{P}_{\mathfrak{P}}$ and $u_{X\mathfrak{P}^*} \in X \otimes \mathcal{P}_{\mathfrak{P}^*}$.

Note: Pythagoras' theorem and Galerkin orthogonality imply (for $u_{X^*\mathfrak{P}}$)

$$\|e\|_B^2 = \|e_{X^*\mathfrak{P}}\|_B^2 + \|u_{X^*\mathfrak{P}} - u_{X\mathfrak{P}}\|_B^2, \quad \text{where } e_{X^*\mathfrak{P}} := u - u_{X^*\mathfrak{P}}.$$

Recall

$$\eta := \left(2\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0}^2 + \|\bar{e}_{Y\mathfrak{P}}\|_{B_0}^2 + 2\,\|\bar{e}_{Y\mathfrak{Q}}\|_{B_0}^2\right)^{1/2},$$

with $\bar{e}_{X\mathfrak{Q}} \in V_{X\mathfrak{Q}}$, $\bar{e}_{Y\mathfrak{P}} \in V_{Y\mathfrak{P}}$, $\bar{e}_{Y\mathfrak{Q}} \in V_{Y\mathfrak{Q}}$.

Estimates of the error reduction

$$\sqrt{\lambda}\,\|\bar{e}_{Y\mathfrak{P}}\|_{B_0} \le \|u_{X^*\mathfrak{P}} - u_{X\mathfrak{P}}\|_B \le \frac{\sqrt{\Lambda}}{\sqrt{1-\gamma^2}}\,\|\bar{e}_{Y\mathfrak{P}}\|_{B_0},$$

$$\sqrt{\lambda}\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0} \le \|u_{X\mathfrak{P}^*} - u_{X\mathfrak{P}}\|_B \le \sqrt{\Lambda}\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0}.$$

Here, λ, Λ are the constants in (B_0) and $\gamma \in [0,1)$ is the constant in the (deterministic) strengthened Cauchy–Schwarz inequality.

Error reduction indicators

Inspired by hp-adaptive algorithm in [A. Schmidt, K. G. Siebert '00].

Enhanced Galerkin approximations: $u_{X^*\mathfrak{P}} \in X^* \otimes \mathcal{P}_{\mathfrak{P}}$ and $u_{X\mathfrak{P}^*} \in X \otimes \mathcal{P}_{\mathfrak{P}^*}$.

Note: Pythagoras' theorem and Galerkin orthogonality imply (for $u_{X^*\mathfrak{P}}$)

$$\|e\|_B^2 = \|e_{X^*\mathfrak{P}}\|_B^2 + \|u_{X^*\mathfrak{P}} - u_{X\mathfrak{P}}\|_B^2, \quad \text{where } e_{X^*\mathfrak{P}} := u - u_{X^*\mathfrak{P}}.$$

Recall

$$\eta := \left(2\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0}^2 + \|\bar{e}_{Y\mathfrak{P}}\|_{B_0}^2 + 2\,\|\bar{e}_{Y\mathfrak{Q}}\|_{B_0}^2\right)^{1/2},$$

with $\bar{e}_{X\mathfrak{Q}} \in V_{X\mathfrak{Q}}$, $\bar{e}_{Y\mathfrak{P}} \in V_{Y\mathfrak{P}}$, $\bar{e}_{Y\mathfrak{Q}} \in V_{Y\mathfrak{Q}}$.

Estimates of the error reduction

$$\sqrt{\lambda}\,\|\bar{e}_{Y\mathfrak{P}}\|_{B_0} \le \|u_{X^*\mathfrak{P}} - u_{X\mathfrak{P}}\|_B \le \frac{\sqrt{\Lambda}}{\sqrt{1-\gamma^2}}\,\|\bar{e}_{Y\mathfrak{P}}\|_{B_0},$$

$$\sqrt{\lambda}\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0} \le \|u_{X\mathfrak{P}^*} - u_{X\mathfrak{P}}\|_B \le \sqrt{\Lambda}\,\|\bar{e}_{X\mathfrak{Q}}\|_{B_0}.$$

This opens the door to $h\otimes p$ adaptivity ...

Numerical experiments: model problem

We solve the steady-state diffusion problem

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

with $\mathbf{x}=(x_1,x_2)\in D:=[-1,1]^2$, with data $f(\mathbf{x})=\frac{1}{8}(2-x_1^2-x_2^2)$ and

$$a(\mathbf{x},\mathbf{y}) = \underbrace{1}_{a_0=\mathbb{E}[a]} + \sigma\sqrt{3}\sum_{m=1}^{\infty}\sqrt{\lambda_m}\,\varphi_m(\mathbf{x})\,y_m.$$

Here $\mathbf{y}\in\Gamma=\prod_{m=1}^{\infty}[-1,1]$ and $\{(\lambda_m,\varphi_m)\}_{m=1}^{\infty}$ are the eigenpairs of the integral operator associated with the covariance function

$$C[a](\mathbf{x},\mathbf{x}') = \sigma^2\exp\left(-\tfrac{1}{2}\|\mathbf{x}-\mathbf{x}'\|_{\ell_1}\right), \quad \mathbf{x},\,\mathbf{x}'\in[-1,1]^2.$$

The weak formulation $(\star)$ of this problem is well posed as long as σ is small enough: in our experiments $\sigma^\star=0.2$.

Numerical experiments: approximation spaces

Solution space: $V := L^2_\pi(\Gamma; H^1_0(D)) \sim H^1_0(D) \otimes L^2_\pi(\Gamma)$.

Finite-dimensional subspace of V: $V_{X\mathfrak{P}} = X_h \otimes \mathcal{P}_{M,p}$, where

- $X_h \subset H^1_0(D)$ is the Q_1 finite element space on a uniform partition $\square_h$ of D;
- $\mathcal{P}_{M,p}$ is the space of polynomials of total degree $\leq p$ in y_m, $m = 1, \ldots, M$.

Numerical experiments: approximation spaces

Solution space: $V := L^2_\pi(\Gamma; H^1_0(D)) \sim H^1_0(D) \otimes L^2_\pi(\Gamma)$.

Finite-dimensional subspace of V: $V_{X\mathfrak{P}} = X_h \otimes \mathcal{P}_{M,p}$, where

- $X_h \subset H^1_0(D)$ is the Q_1 finite element space on a uniform partition $\square_h$ of D;
- $\mathcal{P}_{M,p}$ is the space of polynomials of total degree $\leq p$ in y_m, $m = 1, \ldots, M$.

Enriched finite-dimensional subspace:

$$V^*_{X\mathfrak{P}} = X_{h/2} \otimes \mathcal{P}_{M,p+1} = \big(X_h \oplus Y_h\big) \otimes \big(\mathcal{P}_{M,p} \oplus \mathcal{Q}_{M,p}\big);$$

Numerical experiments: approximation spaces

Solution space: $V := L^2_\pi(\Gamma; H^1_0(D)) \sim H^1_0(D) \otimes L^2_\pi(\Gamma)$.

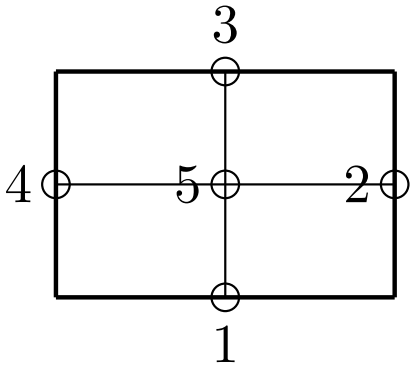
Finite-dimensional subspace of V: $V_{X\mathfrak{P}} = X_h \otimes \mathcal{P}_{M,p}$, where

- $X_h \subset H^1_0(D)$ is the Q_1 finite element space on a uniform partition $\square_h$ of D;
- $\mathcal{P}_{M,p}$ is the space of polynomials of total degree $\leq p$ in y_m, $m = 1, \ldots, M$.

Enriched finite-dimensional subspace:

$$V^*_{X\mathfrak{P}} = X_{h/2} \otimes \mathcal{P}_{M,p+1} = \big(X_h \oplus Y_h\big) \otimes \big(\mathcal{P}_{M,p} \oplus \mathcal{Q}_{M,p}\big);$$

- Y_h spans the set of bilinear bubble functions corresponding to the edge midpoints and element centroids of the original mesh $\square_h$ (defined locally)



Numerical experiments: approximation spaces

Solution space: $V := L^2_\pi(\Gamma; H^1_0(D)) \sim H^1_0(D) \otimes L^2_\pi(\Gamma)$.

Finite-dimensional subspace of V: $V_{X\mathfrak{P}} = X_h \otimes \mathcal{P}_{M,p}$, where

- $X_h \subset H^1_0(D)$ is the Q_1 finite element space on a uniform partition $\square_h$ of D;
- $\mathcal{P}_{M,p}$ is the space of polynomials of total degree $\leq p$ in y_m, $m = 1, \ldots, M$.

Enriched finite-dimensional subspace:

$$V^*_{X\mathfrak{P}} = X_{h/2} \otimes \mathcal{P}_{M,p+1} = \big(X_h \oplus Y_h\big) \otimes \big(\mathcal{P}_{M,p} \oplus \mathcal{Q}_{M,p}\big);$$

- Y_h spans the set of bilinear bubble functions corresponding to the edge midpoints and element centroids of the original mesh $\square_h$ (defined locally);
- $\mathcal{Q}_{M,p}$ spans M-variate polynomials of total degree equal to $p+1$.

Numerical experiments: error estimation

The smart choice of the detail space Y_h gives rise to local problems defined on all elements $K \in \square_h$.

For example, we compute $\bar{e}_{Y\mathfrak{P}}|_K \in Y_h\big|_K \otimes \mathcal{P}_{M,p}$ satisfying

$$\begin{aligned} B_{0,K}(\bar{e}_{Y\mathfrak{P}}|_K, v) \quad &= \quad F_K(v) \\ &\quad + \int_\Gamma \int_K \nabla \cdot \big(a(\mathbf{x},\mathbf{y})\, \nabla u_{X\mathfrak{P}}(\mathbf{x},\mathbf{y})\big)\, v(\mathbf{x},\mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\pi(\mathbf{y}) \\ &\quad - \frac{1}{2} \int_\Gamma \int_{\partial K \setminus \partial D} a(s,\mathbf{y}) \left[\!\!\left[\frac{\partial u_{X\mathfrak{P}}}{\partial n} \right]\!\!\right] v(s,\mathbf{y})\, \mathrm{d}s\, \mathrm{d}\pi(\mathbf{y}), \end{aligned}$$

for all $v \in Y_h|_K \otimes \mathcal{P}_{M,p}$.

The coefficient matrix associated with this local problem is the Kronecker product of a 5×5 (stiffness) matrix and an identity matrix of dimension $\dim(\mathcal{P}_{M,p}) = \frac{(p+M)!}{p!\,M!}$.

Numerical experiments: effectivity I

We compute a reference solution u_{ref} for the case $M = 3$, $\sigma^\star = 0.2$, using a 1024×1024 grid with $p^\star = 7$ (in this case $\dim(V_{X\mathfrak{P}}) = 125'583'480$).

fixed polynomial degree: $p = 2$

effectivity index : $\Theta_\eta := \eta(u_{X\mathfrak{P}}) / \|u_{\mathrm{ref}} - u_{X\mathfrak{P}}\|_B$

Grid	$\dim(V_{X\mathfrak{P}})$	$\eta(u_{X\mathfrak{P}})$	$\|u_{\mathrm{ref}} - u_{X\mathfrak{P}}\|_B$	Θ_η
8×8	490	1.8411×10^{-2}	1.8882×10^{-2}	0.98
16×16	2250	8.7125×10^{-3}	9.4502×10^{-3}	0.92
32×32	9610	4.3394×10^{-3}	4.7752×10^{-3}	0.91
64×64	39690	2.3500×10^{-3}	2.4889×10^{-3}	0.94
128×128	161290	1.5192×10^{-3}	1.4297×10^{-3}	1.06

Numerical experiments: effectivity II

We compute a reference solution u_{ref} for the case $M = 3$, $\sigma^\star = 0.2$, using a 1024×1024 grid with $p^\star = 7$ (in this case $\dim(V_{X\mathfrak{P}}) = 125'583'480$).

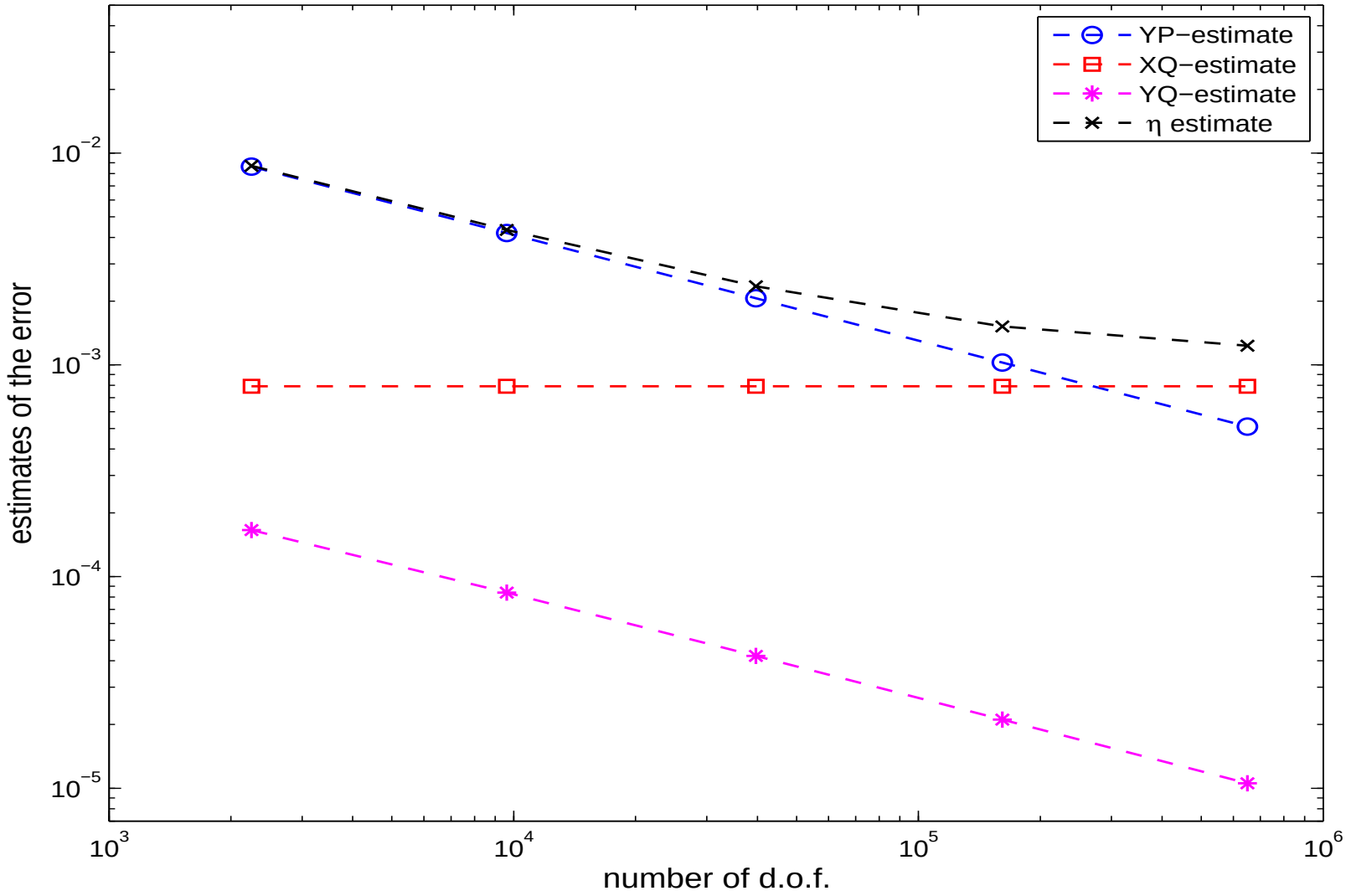
fixed spatial grid: 64×64

effectivity index : $\Theta_\eta := \eta(u_{X\mathfrak{P}}) / \|u_{\mathrm{ref}} - u_{X\mathfrak{P}}\|_B$

p	$\dim(V_{X\mathfrak{P}})$	$\eta(u_{X\mathfrak{P}})$	$\|u_{\mathrm{ref}} - u_{X\mathfrak{P}}\|_B$	Θ_η
1	15876	6.2228×10^{-3}	4.8980×10^{-3}	1.27
2	39690	2.3500×10^{-3}	2.4889×10^{-3}	0.94
3	79380	2.0770×10^{-3}	2.3539×10^{-3}	0.88
4	138915	2.0650×10^{-3}	2.3482×10^{-3}	0.88
5	222264	2.0645×10^{-3}	2.3479×10^{-3}	0.88

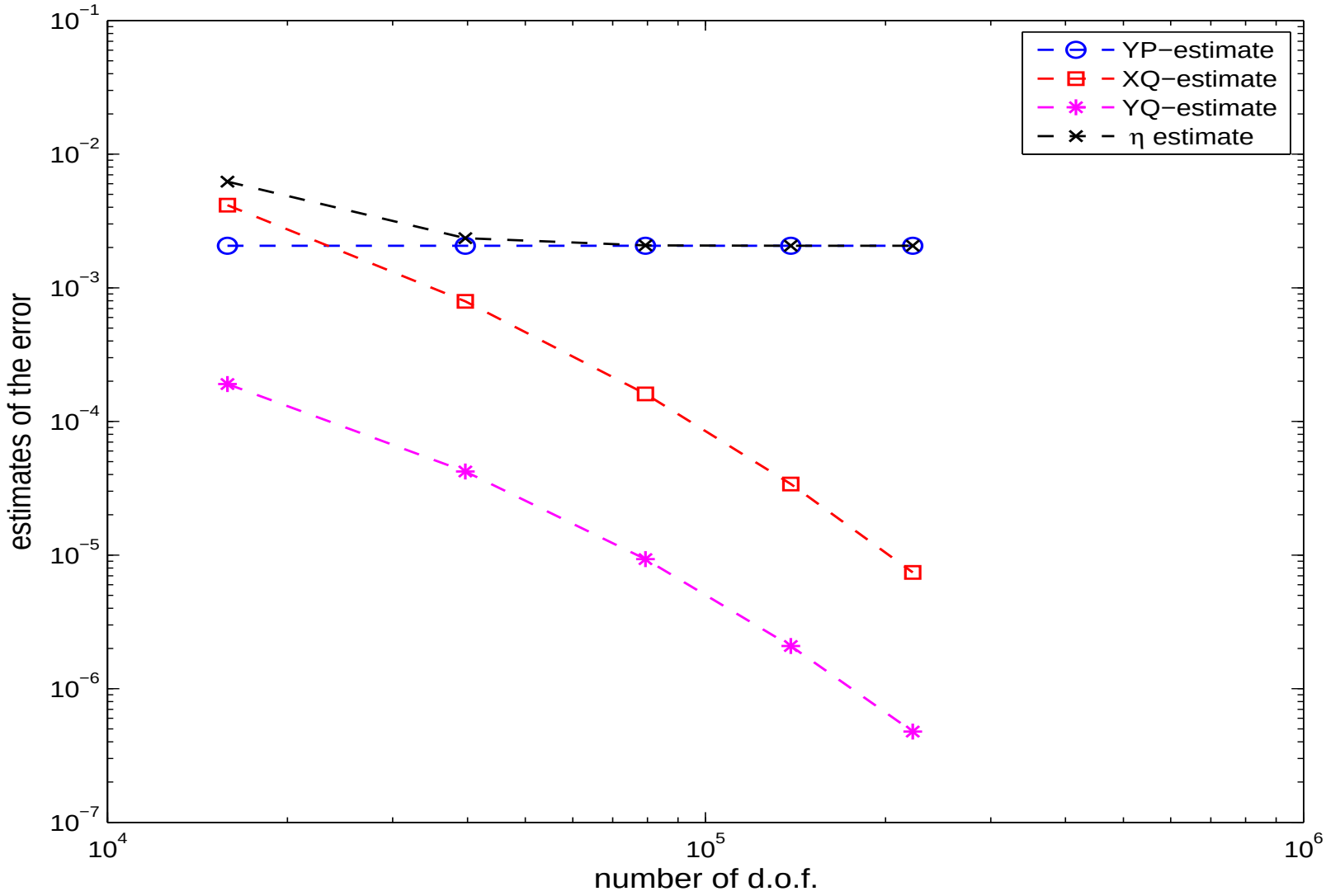
Numerical experiments: error reduction I

$M = 3, \quad \sigma^\star = 0.2, \quad$ fixed polynomial degree $p = 2$



Numerical experiments: error reduction II

$M = 3, \quad \sigma^\star = 0.2, \quad$ fixed spatial grid 64×64



Conclusions

What have we achieved?

⋆ Computable estimates of the error reduction: these exploit the tensor product structure of the approximation spaces.

⋆ Efficient linear algebra: CPU time for the error estimation is commensurate with the CPU time taken to solve the original problem.

Conclusions

What have we achieved?

⋆ Computable estimates of the error reduction: these exploit the tensor product structure of the approximation spaces.

⋆ Efficient linear algebra: CPU time for the error estimation is commensurate with the CPU time taken to solve the original problem.

A. B., C. Powell & D. Silvester

Energy norm a posteriori error estimation for parametric operator equations

SIAM J. Sci. Comp. (to appear)

Conclusions

What have we achieved?

⋆ Computable estimates of the error reduction: these exploit the tensor product structure of the approximation spaces.

⋆ Efficient linear algebra: CPU time for the error estimation is commensurate with the CPU time taken to solve the original problem.

What's next?

- Including local refinement in space
- Multilevel approximation spaces
- Designing a practical adaptive strategy ...

Conclusions

What have we achieved?

⋆ Computable estimates of the error reduction: these exploit the tensor product structure of the approximation spaces.

⋆ Efficient linear algebra: CPU time for the error estimation is commensurate with the CPU time taken to solve the original problem.

What's next?

- Including local refinement in space
- Multilevel approximation spaces
- Designing a practical adaptive strategy ...

Thank you for your attention!