Manchester Numerical Analysis

A posteriori error estimation for stochastic Galerkin approximation

Alex Bespalov

Catherine Powell

David Silvester

University of Manchester

Supported by EPSRC grant EP/H021205/1

Outline of the talk ...

- PDEs with random data
- Stochastic Galerkin and h–p adaptivity:
  - estimates of the error reduction
- A proof-of-concept implementation:
  - efficient linear algebra
  - the IFISS MATLAB Toolbox

Outline of the talk ...

- PDEs with random data
- Stochastic Galerkin and h–p adaptivity:
  - estimates of the error reduction
- A proof-of-concept implementation:
  - efficient linear algebra
  - the IFISS MATLAB Toolbox

Alex Bespalov, Catherine Powell & David Silvester.
A posteriori error estimation for parametric operator equations with applications to PDEs with random data.
MIMS Eprint 2013.16

elliptic PDEs with random data

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We consider the following generic operator equation

$$A(\boldsymbol{\xi})u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}) \iff A(\mathbf{y})u(\mathbf{y}) = f(\mathbf{y}) \quad \forall \mathbf{y} \in \Gamma$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$ is a vector of bounded real-valued random variables $\xi_m : \Omega \to \Gamma_m \subset \mathbb{R}$ with $\Gamma = \prod_{m=1}^{\infty} \Gamma_m$.

elliptic PDEs with random data

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We consider the following generic operator equation

$$A(\boldsymbol{\xi})u(\boldsymbol{\xi}) = f(\boldsymbol{\xi}) \iff A(\mathbf{y})u(\mathbf{y}) = f(\mathbf{y}) \quad \forall \mathbf{y} \in \Gamma$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$ is a vector of bounded real-valued random variables $\xi_m : \Omega \to \Gamma_m \subset \mathbb{R}$ with $\Gamma = \prod_{m=1}^{\infty} \Gamma_m$.

The multivariate random variable $\boldsymbol{\xi}$ generates a distribution π satisfying $\pi(G) = \mathbb{P}(\boldsymbol{\xi}^{-1}(G))$ for $G \in \mathcal{B}(\Gamma)$, which is also a probability measure (by definition). Thus, the observations $\mathbf{y}$ of $\boldsymbol{\xi}$ belong to Γ and $(\Gamma, \mathcal{B}(\Gamma), \pi)$ is a probability space.

weak formulation | energy estimates

Define the solution space $V := L^2_\pi(\Gamma; H)$, where H is a separable Hilbert space with norm $\|\cdot\|_H$. The natural norm on V is $\|\cdot\|_V := \left(\int_\Gamma \|\cdot\|_H^2 \,\mathrm{d}\pi(\mathbf{y})\right)^{1/2}$.

The weak formulation is to find $u \in V$ such that

$$\begin{aligned} B(u,v) &= F(v) \qquad \forall\, v \in V, \qquad\qquad (\star) \\ \text{with} \quad B(u,v) &:= \int_\Gamma \langle A(\mathbf{y})u(\mathbf{y}), v(\mathbf{y})\rangle \,\mathrm{d}\pi(\mathbf{y}) \\ F(v) &:= \int_\Gamma \langle f(\mathbf{y}), v(\mathbf{y})\rangle \,\mathrm{d}\pi(\mathbf{y}). \end{aligned}$$

If $(\star)$ is well posed then the energy norm $\|u\|_B = (B(u,u))^{1/2}$ is equivalent to the (natural) norm $\|u\|_V$.

model problem | steady-state diffusion

Let $D \subset \mathbb{R}^d$ ($d = 2,\, 3$). Suppose that the diffusion coefficient $a = a(\mathbf{x}, \boldsymbol{\xi}) : D \times \Omega \to \mathbb{R}$ is a second-order correlated random field that is a function of a multivariate random variable $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots)$. One possible PDE problem is: find the function $u(\mathbf{x}, \mathbf{y})$ satisfying

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

where $\Gamma := \prod_{m=1}^{\infty} [-1, 1]$, with the series representation

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

The parameter-free term $a_0(\mathbf{x})$ typically represents $\int_\Gamma a(\mathbf{x}, \mathbf{y})\, \mathrm{d}\pi(\mathbf{y})$; the mean coefficient at a given point $\mathbf{x} \in D$.

model problem | spatial formulation

Define the (mean) solution space $H := H_0^1(D)$.

For all $\mathbf{y} \in \Gamma$ we define the operator $A(\mathbf{y}) \in \mathcal{L}(H_0^1(D), H^{-1}(D))$ by

$$\langle A(\mathbf{y})v, w\rangle := \int_D a(\mathbf{x}, \mathbf{y}) \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, \mathsf{d}\mathbf{x} \quad \forall\, v, w \in H_0^1(D).$$

Next, using the series expansion, we can write $A(\mathbf{y})$ as the sum

$$A(\mathbf{y}) = A_0 + \sum_{m=1}^{\infty} y_m A_m \quad \text{for all } \mathbf{y} \in \Gamma$$

with A_m satisfying

$$\begin{aligned} \langle A_m v, w\rangle &= \int_D a_m(\mathbf{x}) \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, \mathsf{d}\mathbf{x} \quad \forall\, v, w \in H_0^1(D) \\ \langle A_0 v, w\rangle &= \int_D a_0(\mathbf{x}) \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, \mathsf{d}\mathbf{x} \quad \forall\, v, w \in H_0^1(D). \end{aligned}$$

model problem | well posedness

$$\langle A_m v, w \rangle = \int_D a_m(\mathbf{x}) \, \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, \mathsf{d}\mathbf{x} \quad \forall\, v, w \in H_0^1(D)$$

$$\langle A_0 v, w \rangle = \int_D a_0(\mathbf{x}) \, \nabla v(\mathbf{x}) \cdot \nabla w(\mathbf{x}) \, \mathsf{d}\mathbf{x} \quad \forall\, v, w \in H_0^1(D)$$

The weak formulation of (M) is well posed if $a_0 \in L^\infty(D)$ is uniformly bounded away from zero, that is,

$$\exists\, \alpha_0^{\min},\, \alpha_0^{\max} > 0 \text{ such that } \alpha_0^{\min} \leq a_0(\mathbf{x}) \leq \alpha_0^{\max} \text{ a.e. in } D,$$

and that for $a_m(\mathbf{x}) \in L^\infty(D)$, $m \in \mathbb{N}$, there exists a constant τ so that

$$\tau := \frac{1}{\alpha_0^{\min}} \sum_{m=1}^{\infty} \|a_m\|_{L^\infty(D)} < 1.$$

model problem | energy estimates

Define $V := L^2_\pi(\Gamma; H^1_0(D))$ and let $\|u\|_H := \|\nabla u\|$. Using the series expansion, the weak formulation of (M) is to find $u \in V$ such that

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

with

$$\begin{aligned} B_0(u,v) &:= \int_\Gamma \langle A_0 u(\mathbf{y}), v(\mathbf{y})\rangle \,\mathrm{d}\pi(\mathbf{y}), \\ B_m(u,v) &:= \int_\Gamma \langle A_m u(\mathbf{y}), v(\mathbf{y})\rangle \, y_m \,\mathrm{d}\pi(\mathbf{y}) \quad \forall\, m \in \mathbb{N}. \end{aligned}$$

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

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

stochastic Galerkin | approximation spaces

Given the isometric isomorphism: $V = L^2_\pi(\Gamma, H) \sim H \otimes L^2_\pi(\Gamma)$. It is natural[a] to take a tensor product approximation

$$X \otimes \mathcal{P}_{\mathfrak{P}} =: V_{X\mathfrak{P}} \subset V, \qquad \text{with components:}$$

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

$$\mathfrak{I} := \left\{\nu = (\nu_1, \nu_2, \ldots) \in \mathbb{N}_0^{\mathbb{N}};\ \#\ \mathrm{supp}\ \nu < \infty\right\},$$

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

---

[a] see e.g., C. Schwab & C. Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Numer. **20** (2011).

stochastic Galerkin | error estimation I

$$B(u,v) = F(v) \quad \forall v \in V \qquad (\star)$$

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

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

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

stochastic Galerkin | error estimation I

$$B(u,v) = F(v) \quad \forall v \in V \qquad\qquad (\star)$$

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

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

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

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

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

$$\Longleftrightarrow \quad V^*_{X\mathfrak{P}} := V_{X\mathfrak{P}} \oplus \underbrace{V_{X\mathfrak{Q}}}_{X\otimes\mathcal{P}_{\mathfrak{Q}}} \oplus \underbrace{V_{Y\mathfrak{P}}}_{Y\otimes\mathcal{P}_{\mathfrak{P}}} \oplus \underbrace{V_{Y\mathfrak{Q}}}_{Y\otimes\mathcal{P}_{\mathfrak{Q}}}.$$

stochastic Galerkin | error estimation II

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

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

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

stochastic Galerkin | error estimation II

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

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

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

A second estimator e^*_0 is obtained by exploiting the series structure of $(\star)$: find $e^*_0 \in V^*_{X\mathfrak{P}}$ such that

$$B_0(e^*_0, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V^*_{X\mathfrak{P}}.$$

This leads to efficient linear algebra—the associated system matrix is block diagonal with each block representing a deterministic problem associated with the A_0 component of the operator $A(\mathbf{y})$.

Note that using the norm equivalence (B_0) gives

$$\sqrt{\lambda}\, \|e^*_0\|_{B_0} \le \|e^*\|_B \le \sqrt{\Lambda}\, \|e^*_0\|_{B_0}. \qquad (B^*_0)$$

stochastic Galerkin | error estimation III

The second error estimator $e_0^* \in V_{X\mathfrak{P}}^*$ satisfies

$$B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{X\mathfrak{P}}^*, \text{ with}$$

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

stochastic Galerkin | error estimation III

The second error estimator $e_0^* \in V_{X\mathfrak{P}}^*$ satisfies

$$B_0(e_0^*, v) = F(v) - B(u_{X\mathfrak{P}}, v) \quad \forall\, v \in V_{X\mathfrak{P}}^*, \text{ with}$$

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

A third energy estimate η is obtained by exploiting the structure of the space $V_{X\mathfrak{P}}^*$: This is given by

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

where the contributing estimators $\bar{e}_{X\mathfrak{Q}} \in V_{X\mathfrak{Q}}$, $\bar{e}_{Y\mathfrak{P}} \in V_{Y\mathfrak{P}}$ and $\bar{e}_{Y\mathfrak{Q}} \in V_{Y\mathfrak{Q}}$ satisfy

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

Since $\langle A_0\cdot,\cdot\rangle$ defines an inner product in H and $X \cap Y = \{0\}$, there exists a constant $\gamma \in [0,1)$ such that[a]

$$|\langle A_0 u_X, v_Y\rangle| \le \gamma \langle A_0 u_X, u_X\rangle^{1/2} \langle A_0 v_Y, v_Y\rangle^{1/2} \quad \forall\, u_X \in X,\ \forall\, v_Y \in Y.$$

Note that for any finite index set $\mathfrak{P} \subset \mathfrak{I}$, the finite-dimensional subspaces $V_{X\mathfrak{P}}$, $V_{Y\mathfrak{P}} \subset V$ are such that the following strengthened Cauchy–Schwarz inequality holds:

$$|B_0(u,v)| \le \gamma\, \|u\|_{B_0}\, \|v\|_{B_0} \qquad \forall\, u \in V_{X\mathfrak{P}},\ \forall\, v \in V_{Y\mathfrak{P}}.$$

This allows us to establish the following equivalence between the estimator $e_0^* \in V_{X\mathfrak{P}}^*$ and the componentwise estimate η :

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

---

[a] see, for example, V. Eijkhout & P. Vassilevski, *The role of the strengthened Cauchy–Buniakowskiĭ–Schwarz inequality in multilevel methods*, SIAM Review **33** (1991).

error reduction indicators

Recall

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

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

The construction of η also gives information regarding the error reduction that would result by enriching the approximation either spatially or parametrically:

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

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

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

This opens the door to h–p adaptivity ...

model problem | example

We solve the steady-state diffusion problem $(\star)$ 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 correlation function

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

The problem $(\star)$ is well posed as long as σ is small enough: our reference value is $\sigma^\star = 0.2$.

model problem | example

We solve the steady-state diffusion problem $(\star)$ 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

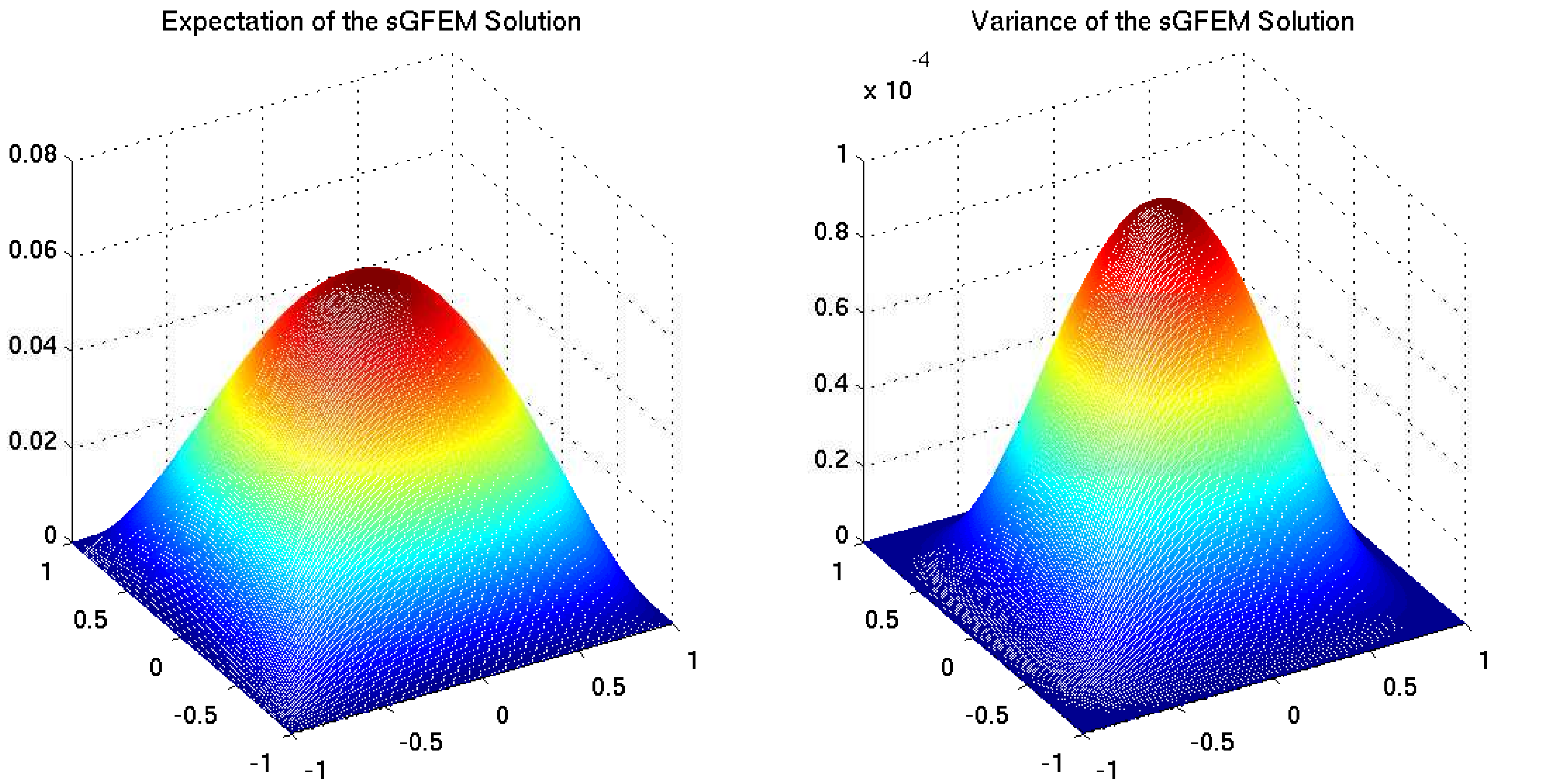
$$a(\mathbf{x}, \mathbf{y}) = \underbrace{1}_{a_0 = \mathbb{E}[a]} + \sigma \sqrt{3} \sum_{m=1}^{\infty} \sqrt{\lambda_m}\, \varphi_m(\mathbf{x})\, y_m.$$

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 correlation function

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

The problem $(\star)$ is well posed as long as σ is small enough: our reference value is $\sigma^\star = 0.2$. The Galerkin approximation is given by tensor product of the $\boldsymbol{Q}_1$ approximation space X_h defined on a uniform partition $\square_h$ of D with the subspace $\mathcal{P}_{M,p}$ of polynomials of total degree p in the first M parameters y_m, $m = 1, \ldots, M$.

model problem | sample solution

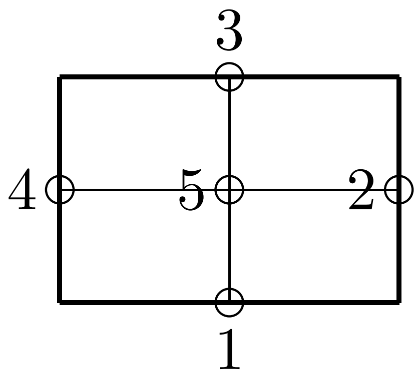


model problem | detail spaces

The detail spaces Y_h and $\mathcal{Q}_{M,p}$ are chosen so that

$$X_{h/2} = X_h \oplus Y_h \quad \text{and} \quad \mathcal{P}_{M,p+1} = \mathcal{P}_{M,p} \oplus \mathcal{Q}_{M,p},$$

respectively. Thus, 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, see below)



whereas $\mathcal{Q}_{M,p}$ spans M-variate polynomials of total degree equal to $p+1$. This gives a decomposition of the error estimate η with

$$V^*_{X\mathfrak{P}} := V_{X\mathfrak{P}} \oplus (X_h \otimes \mathcal{Q}_{M,p}) \oplus (Y_h \otimes \mathcal{P}_{M,p}) \oplus (Y_h \otimes \mathcal{Q}_{M,p}).$$

spatial error estimation | example

Compute the harmonic function

$$u(x_1, x_2) = \frac{2(1 + x_2)}{(3 + x_1)^2 + (1 + x_2)^2}$$

by solving $\nabla^2 u = 0$ on a uniform subdivision of $D = [-1, 1] \times [-1, 1]$ with boundary data g obtained by Q_1 interpolation of u on ∂D.

Global Effectivity Index : $X_\eta := \eta(u_h)/\|\nabla(u - u_h)\|$

Grid	$\|\nabla(u - u_h)\|$	$\eta(u_h)$	X_η
4×4	5.032×10^{-2}	4.954×10^{-2}	0.9845
8×8	2.516×10^{-2}	2.511×10^{-2}	0.9980
16×16	1.258×10^{-2}	1.257×10^{-2}	0.9992
32×32	6.291×10^{-3}	6.288×10^{-2}	0.9995

model problem | error estimation

$$V^*_{X\mathfrak{P}} := V_{X\mathfrak{P}} \oplus (X_h \otimes \mathcal{Q}_{M,p}) \oplus (Y_h \otimes \mathcal{P}_{M,p}) \oplus (Y_h \otimes \mathcal{Q}_{M,p}).$$

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

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

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 $|P| = \dim(\mathcal{P}_{M,p}) = \frac{(p+M)!}{p!\, M!}$.

example | effectivity I

We compute a reference solution u_{ref} for the case $M = 3$, $\sigma^\star = 0.2$, using a 512×512 grid with $p^\star = 7$.

fixed polynomial degree: $p = 2$

effectivity Index : $X_\eta := \eta(u_{X\mathfrak{P}})/\|u_{\text{ref}} - u_{X\mathfrak{P}}\|_B$

Grid	$\|u_{\text{ref}} - u_{X\mathfrak{P}}\|_B$	$\eta(u_{X\mathfrak{P}})$	X_η
8×8	1.841×10^{-2}	1.888×10^{-2}	0.98
16×16	8.712×10^{-3}	9.450×10^{-3}	0.92
32×32	4.339×10^{-3}	4.775×10^{-3}	0.91
64×64	2.350×10^{-3}	2.488×10^{-3}	0.94
128×128	1.519×10^{-3}	1.429×10^{-3}	1.06

example | effectivity II

We compute a reference solution u_{ref} for the case $M = 3$, $\sigma^\star = 0.2$, using a 512×512 grid with $p^\star = 7$.

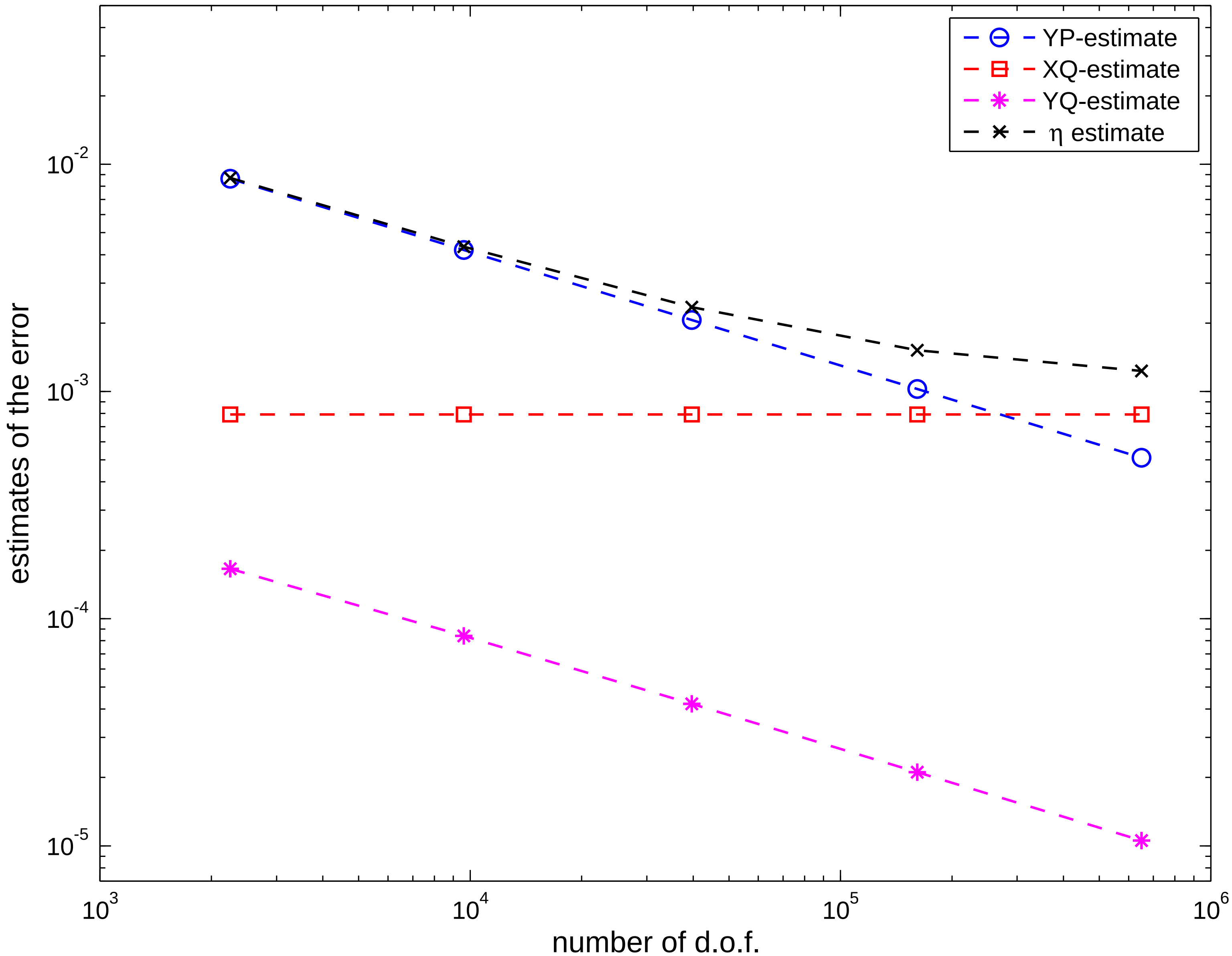
fixed spatial grid: 64×64

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

p	$\|u_{\mathrm{ref}} - u_{X\mathfrak{P}}\|_B$	$\eta(u_{X\mathfrak{P}})$	X_η
1	6.222×10^{-3}	4.898×10^{-3}	1.27
2	2.350×10^{-3}	2.488×10^{-3}	0.94
3	2.077×10^{-3}	2.353×10^{-3}	0.88
4	2.065×10^{-3}	2.348×10^{-3}	0.88
5	2.064×10^{-3}	2.347×10^{-3}	0.88

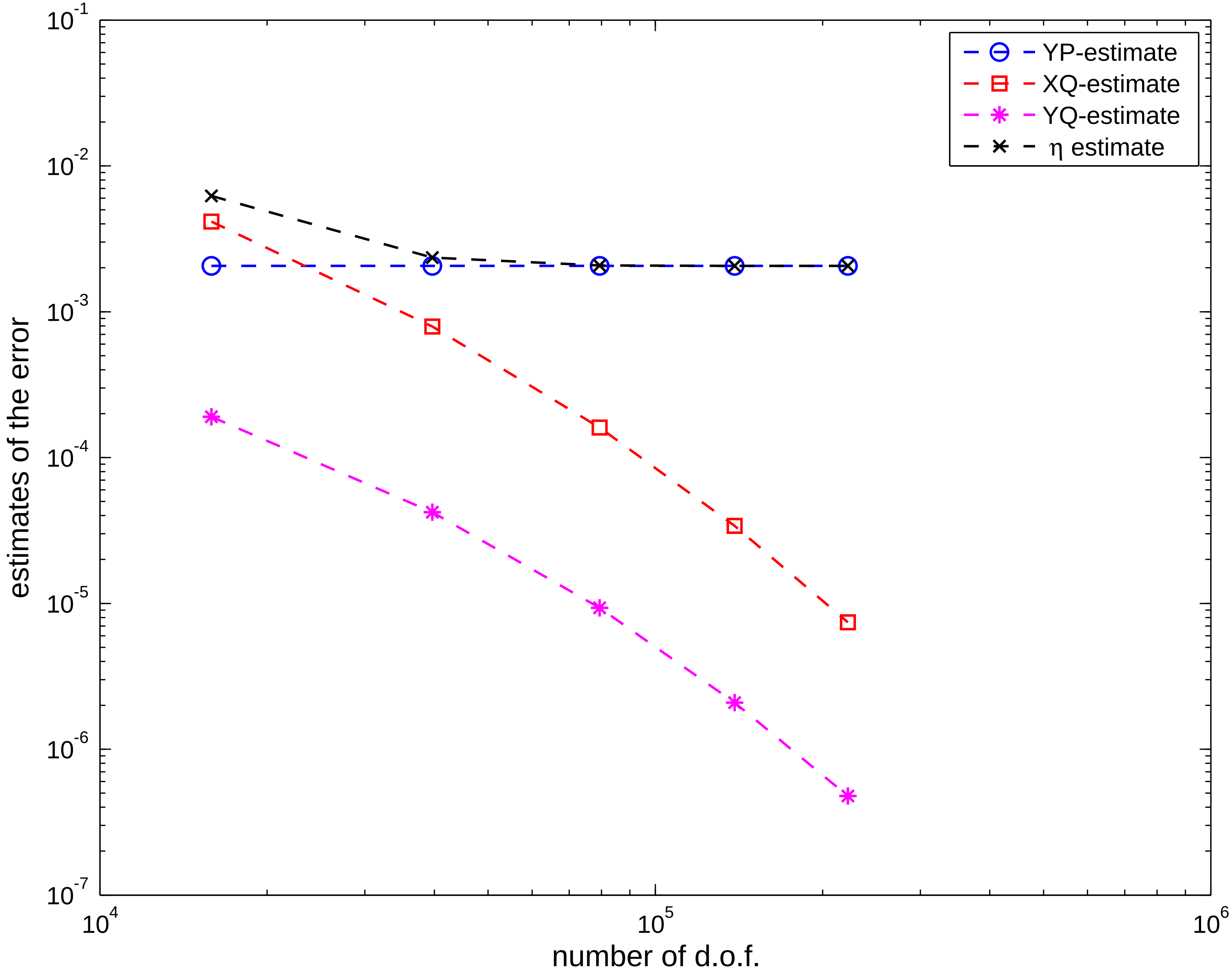
error reduction | decreasing h

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



error reduction | increasing p

$M = 3$, $\sigma^{\star} = 0.2$, fixed spatial grid: 64×64



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...

But, is the linear algebra done efficiently (or not)?

$$A(\mathbf{y}) = A_0 + \sigma \sum_{m=1}^{\infty} y_m A_m$$

$$X_h = \mathtt{span}\,\{\phi_j(\boldsymbol{x})\}_{j=1}^{n_x}\,, \quad \mathcal{P}_{M,p} = \mathtt{span}\,\{\psi_k(\mathbf{y})\}_{k=1}^{n_y}$$

The Galerkin approximation is given by tensor product of the $\boldsymbol{Q}_1$ approximation space X_h defined on a uniform partition $\square_h$ of D with the subspace $\mathcal{P}_{M,p}$ of polynomials of total degree p in the first M parameters y_m, $m = 1, \ldots, M$.

$$A(\mathbf{y}) = A_0 + \sigma \sum_{m=1}^{\infty} y_m A_m$$

$$X_h = \mathtt{span}\,\{\phi_j(\boldsymbol{x})\}_{j=1}^{n_x}, \quad \mathcal{P}_{M,p} = \mathtt{span}\,\{\psi_k(\mathbf{y})\}_{k=1}^{n_y}$$

The Galerkin approximation is given by tensor product of the $\boldsymbol{Q}_1$ approximation space X_h defined on a uniform partition $\square_h$ of D with the subspace $\mathcal{P}_{M,p}$ of polynomials of total degree p in the first M parameters y_m, $m = 1, \ldots, M$.

Discrete system

$$\mathcal{A}\mathbf{u} = \mathbf{f} \quad \Longleftrightarrow \quad \big(I \otimes K_0 + \sigma \textstyle\sum_{m=1}^{M} G_m \otimes K_m\big)\mathbf{u} = \mathbf{f} \qquad (S)$$

$$[G_m]_{rs} = \langle\, y_m\, \psi_r(\mathbf{y})\psi_s(\mathbf{y})\rangle$$

- The dimension of the system (S) is $n := n_x \times n_y$.
- An effective solver for (S) is preconditioned MINRES with a block-diagonal preconditioner : $\mathcal{P} = I \otimes K_0$.

Sample performance | decreasing h

$M = 5$ random variables

cubic polynomial approximation $p = 3$

EST_MINRES tolerance is 10^{-13}

	$1/h$	32	64	128	256
	n	60,984	236,600	931,896	$3,698,744$
$\frac{\sigma}{\mu} = 0.1$	# iterations	13	13	13	13
	total solve time	1.0s	4.8s	19.6s	$60.6s$
$\frac{\sigma}{\mu} = 0.2$	# iterations	18	18	18	18
	total solve time	1.2s	6.7s	24.5s	$84.9s$
$\frac{\sigma}{\mu} = 0.3$	# iterations	24	24	24	24
	total solve time	1.6s	9.2s	35.6s	$114.7s$

What have we achieved?

- Efficient linear algebra: convergence rate is independent of $n := n_x \times n_y$.
- Convergence deteriorates as σ is increased. The critical eigenvalue is computed on the fly using EST_MINRES.

What's next?

- ♡ including local refinement in space
- ♡♡ designing a practical adaptive strategy
- ♡♡♡ stopping criteria for the solver (EST_MINRES)