Dimension reduction of the input parameter space of vector-valued functions

Olivier Zahm
INRIA - AIRSEA project team

olivier.zahm@inria.fr
https://team.inria.fr/airsea/

Nantes, April 11th, 2018.

Collaborators: Youssef Marzouk, Clémentine Prieur and Paul Constantine.

$$x \mapsto f(x_1, \ldots, x_d) \ \in V$$

Computational issues:

- expensive-to-evaluate map
- high inputs space dimension $d \gg 1$

Building a surrogate for f requires to exploit some **underlying structure** of f

- small Kolmogorov n-width
- low-rank structure
- sparse structure
- low-effective dimension
- ...

$$f \approx h \circ A \quad \text{where} \quad \begin{cases} \mathbb{R}^d \xrightarrow{\qquad f \qquad} V \\ \mathbb{R}^d \xrightarrow{A} \mathbb{R}^r \xrightarrow{h} V \end{cases}$$

Low-effective dimension via linear maps A

Ridge functions = functions which are constant along a subspace

$$x \mapsto h(A\,x) \qquad \text{where} \begin{cases} A \in \mathbb{R}^{r \times d} \\ h : \mathbb{R}^r \to V \end{cases}$$

or equivalently

$$x \mapsto g(P_r x) \qquad \text{where} \begin{cases} P_r \in \mathbb{R}^{d \times d} \text{ rank-}r \text{ projector} \\ g : \mathbb{R}^d \to V \end{cases}$$

Low-effective dimension via linear maps A

Ridge functions = functions which are constant along a subspace

$$x \mapsto h(Ax) \qquad \text{where} \begin{cases} A \in \mathbb{R}^{r\times d} \\ h : \mathbb{R}^r \to V \end{cases}$$

or equivalently

$$x \mapsto g(P_r x) \qquad \text{where} \begin{cases} P_r \in \mathbb{R}^{d\times d} \text{ rank-}r \text{ projector} \\ g : \mathbb{R}^d \to V \end{cases}$$

The function f has a **low-effective dimension** if it is *close* to a ridge function $g \circ P_r$ with $r = \mathrm{rank}(P_r) \ll d$.

$$\boxed{\mathbb{R}^d = \underbrace{\mathrm{Im}(P_r)}_{\text{“}f\text{ varies ”}} \oplus \underbrace{\mathrm{Ker}(P_r)}_{\text{“}f\text{ is (almost) constant”}}}$$

Today's task

Consider the $L^2_\mu(\mathbb{R}^d; V)$ norm $\|\cdot\|$ defined by

$$\|v\|^2 = \int \|v(x)\|_V^2 \; \mu(\mathrm{d}x), \qquad \text{where } \begin{cases} \mu = \mathcal{N}(m, \Sigma) \\ \|\cdot\|_V = (\cdot,\cdot)_V^{1/2} \end{cases}$$

Controlled approximation problem

Given $\varepsilon \geq 0$, find g and a projector P_r such that

$$\|f - g \circ P_r\| \leq \varepsilon \tag{1}$$

with $r = \mathrm{rank}(P_r)$ much smaller than d.

Note that (1) is equivalent to

$$\mathbb{E}\big(\|f(X) - g(P_r X)\|_V^2\big) \leq \varepsilon^2$$

where $X \sim \mu$.

This talk in a nutshell

Methodology

1. derive an upper bound for the error

$$\|f - g \circ P_r\| \leq \mathcal{R}(g, P_r)$$

2. fix r and solve

$$\min_{g, P_r} \mathcal{R}(g, P_r)$$

3. increase r until

$$\min_{g, P_r} \mathcal{R}(g, P_r) \ \leq \varepsilon$$

Road map:

1. Lipschitz-based upper bound
2. Poincaré-based upper bound
3. Examples
4. Conclusion

Lipschitz-based upper bound

Lipschitz-based upper bound

Find a projector P_r so that the ridge function

$$\widetilde{f} : x \mapsto f(P_r x + (I_d - P_r)m)$$

is good approximation of f.

Lipschitz-based upper bound

Find a projector P_r so that the ridge function

$$\widetilde{f} : x \mapsto f(P_r x + (I_d - P_r)m)$$

is good approximation of f.

Proposition

Assume f is L-Lipschitz, meaning

$$\|f(x) - f(y)\|_V \leq L\|x - y\|_2 \qquad \forall x, y \in \mathbb{R}^d.$$

Then for any projector P_r we have

$$\|f - \widetilde{f}\| \leq L \sqrt{\text{trace}\big((I_d - P_r)\Sigma(I_d - P_r)^T\big)}$$

Lipschitz-based upper bound

Find a projector P_r so that the ridge function

$$\widetilde{f} : x \mapsto f(P_r x + (I_d - P_r)m)$$

is good approximation of f.

Proposition

Assume f is L-Lipschitz, meaning

$$\|f(x) - f(y)\|_V \leq L\|x - y\|_2 \qquad \forall x, y \in \mathbb{R}^d.$$

Then for any projector P_r we have

$$\|f - \widetilde{f}\| \leq L\sqrt{\mathrm{trace}\big((I_d - P_r)\Sigma(I_d - P_r)^T\big)}$$

Proof. By letting $X \sim \mathcal{N}(m, \Sigma)$, can write

$$\begin{aligned}\|f - \widetilde{f}\|^2 &= \mathbb{E}\big(\|f(X) - \widetilde{f}(X)\|_V^2\big) \\ &= \mathbb{E}\big(\|f(X) - f(P_r X + (I_d - P_r)m)\|_V^2\big)\end{aligned}$$

Lipschitz-based upper bound

Find a projector P_r so that the ridge function

$$\widetilde{f} : x \mapsto f(P_r x + (I_d - P_r)m)$$

is good approximation of f.

Proposition

Assume f is L-Lipschitz, meaning

$$\|f(x) - f(y)\|_V \leq L\|x - y\|_2 \qquad \forall x, y \in \mathbb{R}^d.$$

Then for any projector P_r we have

$$\|f - \widetilde{f}\| \leq L \sqrt{\mathrm{trace}\big((I_d - P_r)\Sigma(I_d - P_r)^T\big)}$$

Proof. By letting $X \sim \mathcal{N}(m, \Sigma)$, can write

$$\begin{aligned}\|f - \widetilde{f}\|^2 &= \mathbb{E}\big(\|f(X) - \widetilde{f}(X)\|_V^2\big) \\ &= \mathbb{E}\big(\|f(X) - f(P_r X + (I_d - P_r)m)\|_V^2\big) \\ &\leq \mathbb{E}\big(L^2 \, \|X - (P_r X + (I_d - P_r)m)\|_2^2\big)\end{aligned}$$

Lipschitz-based upper bound

Find a projector P_r so that the ridge function

$$\widetilde{f} : x \mapsto f(P_r x + (I_d - P_r)m)$$

is good approximation of f.

Proposition

Assume f is L-Lipschitz, meaning

$$\|f(x) - f(y)\|_V \leq L\|x - y\|_2 \qquad \forall x, y \in \mathbb{R}^d.$$

Then for any projector P_r we have

$$\|f - \widetilde{f}\| \leq L \sqrt{\text{trace}\big((I_d - P_r)\Sigma(I_d - P_r)^T\big)}$$

Proof. By letting $X \sim \mathcal{N}(m, \Sigma)$, can write

$$\begin{aligned}
\|f - \widetilde{f}\|^2 &= \mathbb{E}\big(\|f(X) - \widetilde{f}(X)\|_V^2\big) \\
&= \mathbb{E}\big(\|f(X) - f(P_r X + (I_d - P_r)m)\|_V^2\big) \\
&\leq \mathbb{E}\big(L^2 \|X - (P_r X + (I_d - P_r)m)\|_2^2\big) \\
&= L^2 \, \mathbb{E}\big(\|(X - m) - P_r(X - m)\|_2^2\big) \\
&= L^2 \, \text{trace}\big((I_d - P_r)\Sigma(I_d - P_r)^T\big)
\end{aligned}$$

Minimizing the bound

$$\min_{P_r} L \sqrt{\operatorname{trace}\big((I_d - P_r)\Sigma(I_d - P_r)^T\big)} = L \sqrt{\sum_{i=r+1}^{d} \sigma_i^2}$$

- $\sigma_1^2 \geq \ldots \geq \sigma_d^2$ are the eigenvalues of Σ.
- The solution is the orthogonal projector on the leading eigenspace of Σ $\Rightarrow$ *cf.* truncated Karhunen-Loève decomposition of X
- A fast decay in σ_i^2 ensures $L\sqrt{\sum_{i=r+1}^{d} \sigma_i^2} \leq \varepsilon$ for $r = r(\varepsilon) \ll d$.

But

- the construction of P_r is independent of f...
- in practice L is not available: no certification of the error...

Poincaré-based upper bound

Optimal function g

A well known property of the conditional expectation

For any given (fixed) P_r, we have

$$\|f - \underbrace{\mathbb{E}_\mu(f|\sigma(P_r))}_{=g^* \circ P_r}\| = \min_g \|f - g \circ P_r\|$$

Optimal function g

A well known property of the conditional expectation

For any given (fixed) P_r, we have

$$\|f - \underbrace{\mathbb{E}_\mu(f|\sigma(P_r))}_{=g^* \circ P_r}\| = \min_g \|f - g \circ P_r\|$$

The conditional expectation $\mathbb{E}_\mu(f|\sigma(P_r)) : \mathbb{R}^d \to V$ can be expressed as follow:

- **Statistical viewpoint:** by letting $X \sim \mu$, we can write

$$\mathbb{E}_\mu(f|\sigma(P_r)) : x \mapsto \mathbb{E}(f(X)|P_r X = P_r x)$$

Optimal function g

A well known property of the conditional expectation

For any given (fixed) P_r, we have

$$\|f - \underbrace{\mathbb{E}_\mu(f|\sigma(P_r))}_{=g^* \circ P_r}\| = \min_g \|f - g \circ P_r\|$$

The conditional expectation $\mathbb{E}_\mu(f|\sigma(P_r)) : \mathbb{R}^d \to V$ can be expressed as follow:

- **Statistical viewpoint:** by letting $X \sim \mu$, we can write

$$\mathbb{E}_\mu(f|\sigma(P_r)) : x \mapsto \mathbb{E}(f(X)|P_r X = P_r x)$$

- **A simple expression:** recall $\mu = \mathcal{N}(m, \Sigma)$. If P_r is Σ^{-1}-orthogonal then

$$\mathbb{E}_\mu(f|\sigma(P_r)) : x \mapsto \int f(P_r x + (I_d - P_r)y)\ \mu(\mathrm{d}y)$$

Poincaré-type inequalities

- The normal distribution $\mu = \mathcal{N}(m, \Sigma)$ satisfies the **Poincaré inequality**

$$\int \big(h - \mathbb{E}_\mu(h)\big)^2 \mathrm{d}\mu \leq \int \|\nabla h\|_\Sigma^2 \, \mathrm{d}\mu,$$

for any smooth function $h : \mathbb{R}^d \to \mathbb{R}$, where $\|x\|_\Sigma^2 = x^T \Sigma x$.

Poincaré-type inequalities

- The normal distribution $\mu = \mathcal{N}(m, \Sigma)$ satisfies the **Poincaré inequality**

$$\int \big(h - \mathbb{E}_\mu(h)\big)^2 \mathrm{d}\mu \leq \int \|\nabla h\|_\Sigma^2 \, \mathrm{d}\mu,$$

for any smooth function $h : \mathbb{R}^d \to \mathbb{R}$, where $\|x\|_\Sigma^2 = x^T \Sigma x$.

- It also satisfies the **"subspace" Poincaré inequality**

$$\int \big(h - \mathbb{E}_\mu(h|\sigma(P_r))\big)^2 \mathrm{d}\mu \leq \int \|(I_d - P_r^T)\nabla h\|_\Sigma^2 \, \mathrm{d}\mu$$

for any smooth function $h : \mathbb{R}^d \to \mathbb{R}$ **and for any projector** P_r.

Poincaré-type inequalities

- The normal distribution $\mu = \mathcal{N}(m, \Sigma)$ satisfies the **Poincaré inequality**

$$\int \left(h - \mathbb{E}_\mu(h)\right)^2 \mathrm{d}\mu \leq \int \|\nabla h\|_\Sigma^2 \, \mathrm{d}\mu,$$

  for any smooth function $h : \mathbb{R}^d \to \mathbb{R}$, where $\|x\|_\Sigma^2 = x^T \Sigma x$.
- It also satisfies the **"subspace" Poincaré inequality**

$$\int \left(h - \mathbb{E}_\mu(h|\sigma(P_r))\right)^2 \mathrm{d}\mu \leq \int \|(I_d - P_r^T)\nabla h\|_\Sigma^2 \, \mathrm{d}\mu$$

  for any smooth function $h : \mathbb{R}^d \to \mathbb{R}$ **and for any projector** P_r.

Proposition

Given a **smooth vector-valued** function $f : \mathbb{R}^d \to V$ we have

$$\|f - \mathbb{E}_\mu(f|\sigma(P_r))\| \leq \sqrt{\mathrm{trace}\left(H(I_d - P_r)\Sigma(I_d - P_r)^T\right)}$$

for any projector P_r. The matrix $H \in \mathbb{R}^{d \times d}$ is defined by

$$\boxed{H = \int (\nabla f)^* (\nabla f) \mathrm{d}\mu} \quad \text{where} \quad \begin{cases} \nabla f(x) : \mathbb{R}^d \to V \text{ Jacobian of } f \text{ at } x \\ \nabla f(x)^* \text{ is the adjoint of } \nabla f(x) \end{cases}$$

What is the matrix H ?

$$H = \int (\nabla f)^* (\nabla f) \mathrm{d}\mu \quad \in \mathbb{R}^{d \times d}$$

What is the matrix H ?

$$H = \int (\nabla f)^* (\nabla f) \mathrm{d}\mu \quad \in \mathbb{R}^{d \times d}$$

- Algebraic case: $V = \mathbb{R}^m$ with $\|\cdot\|_V$ such that $\|v\|_V^2 = v^T R_V v$ for some SPD matrix $R_V \in \mathbb{R}^{m \times m}$. Then

$$H = \int (\nabla f)^T R_V (\nabla f) \, \mathrm{d}\mu, \qquad \text{where} \quad \nabla f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_d} \end{pmatrix}$$

What is the matrix H ?

$$H = \int (\nabla f)^* (\nabla f) \mathrm{d}\mu \quad \in \mathbb{R}^{d \times d}$$

- Algebraic case: $V = \mathbb{R}^m$ with $\|\cdot\|_V$ such that $\|v\|_V^2 = v^T R_V v$ for some SPD matrix $R_V \in \mathbb{R}^{m\times m}$. Then

$$H = \int (\nabla f)^T R_V (\nabla f)\, \mathrm{d}\mu, \qquad \text{where} \quad \nabla f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_d} \end{pmatrix}$$

- Scalar-valued case: $V = \mathbb{R}$ with $\|\cdot\|_V = |\cdot|$, then

$$H = \int (\nabla f)(\nabla f)^T \, \mathrm{d}\mu, \qquad \text{where} \quad \nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_d} \end{pmatrix}$$

$\rightsquigarrow$ Active-Subspace method.

Minimizing the upper bound

$$\min_{P_r} \sqrt{\mathrm{trace}\big(H(I_d - P_r)\Sigma(I_d - P_r)^T\big)} = \sqrt{\sum_{i=r+1}^{d} \lambda_i}$$

- Let (v_i, λ_i) be the i-th generalized eigenpair of (H, Σ^{-1}):

$$Hv_i = \lambda_i \Sigma^{-1} v_i$$

- The solution is a Σ^{-1}-orthogonal projector onto $\mathrm{span}\{v_1, \ldots, v_r\}$.
- A fast decay in λ_i ensures $\sqrt{\sum_{i=r+1}^{d} \lambda_i} \leq \varepsilon$ for $r = r(\varepsilon) \ll d$.
- H provides a test that reveals the low-effective dimension.

Minimizing the upper bound

$$\min_{P_r} \sqrt{\text{trace}\big(H(I_d - P_r)\Sigma(I_d - P_r)^T\big)} = \sqrt{\sum_{i=r+1}^{d} \lambda_i}$$

- Let (v_i, λ_i) be the i-th generalized eigenpair of (H, Σ^{-1}):

$$H v_i = \lambda_i \Sigma^{-1} v_i$$

- The solution is a Σ^{-1}-orthogonal projector onto $\text{span}\{v_1, \ldots, v_r\}$.
- A fast decay in λ_i ensures $\sqrt{\sum_{i=r+1}^{d} \lambda_i} \leq \varepsilon$ for $r = r(\varepsilon) \ll d$.
- H provides a test that reveals the low-effective dimension.

" Poincaré-based bound $\leq$ Lipschitz-based bound "

$$\sqrt{\sum_{i=r+1}^{d} \lambda_i} \;\leq\; L\sqrt{\sum_{i=r+1}^{d} \sigma_i^2}$$

Examples

An analytical example

- Let $\mu = \mathcal{N}(0, I_d)$ be the standard Gaussian distribution and consider the **scalar-valued** function

$$f : x \mapsto \sum_{i=1}^{d} a_i \sin(\omega_i x_i)$$

- Restrict P_r to be an orthogonal projector onto the canonical coordinates

$$P_r = \sum_{i \in \Lambda_r} e_i e_i^T, \quad \text{where} \begin{cases} \Lambda_r \subset \{1, \ldots, d\} \\ \#\Lambda_r = r \end{cases}$$

An analytical example

- Let $\mu = \mathcal{N}(0, I_d)$ be the standard Gaussian distribution and consider the **scalar-valued** function

$$f : x \mapsto \sum_{i=1}^{d} a_i \sin(\omega_i x_i)$$

- Restrict P_r to be an orthogonal projector onto the canonical coordinates

$$P_r = \sum_{i \in \Lambda_r} e_i e_i^T, \quad \text{where} \begin{cases} \Lambda_r \subset \{1, \dots, d\} \\ \#\Lambda_r = r \end{cases}$$

- We can analytically compute the true error and its Poincaré-based bound:

$$\|f - \mathbb{E}_\mu(f|\sigma(P_r))\| = \sqrt{\frac{1}{2} \sum_{i \in \Lambda_r^c} a_i^2 (1 - \exp(-2\omega_i^2))}$$

$$\sqrt{\mathrm{trace}\big(H(I_d - P_r)\Sigma(I_d - P_r)^T\big)} = \sqrt{\frac{1}{2} \sum_{i \in \Lambda_r^c} a_i^2 \omega_i^2 (1 + \exp(-2\omega_i^2))}$$

An analytical example

$$\underbrace{\sqrt{\frac{1}{2}\sum_{i\in\Lambda_r^c} a_i^2(1-\exp(-2\omega_i^2))}}_{\left\{error\right\}} \leq \underbrace{\sqrt{\frac{1}{2}\sum_{i\in\Lambda_r^c} a_i^2\omega_i^2(1+\exp(-2\omega_i^2))}}_{\left\{bound\right\}}$$

- If $\omega_i = \omega$ for all $1 \leq i \leq d$, we have

$$\arg\min_{P_r}\left\{error\right\} = \arg\min_{P_r}\left\{bound\right\}$$

An analytical example

$$\underbrace{\sqrt{\frac{1}{2}\sum_{i\in\Lambda_r^c} a_i^2(1-\exp(-2\omega_i^2))}}_{\left\{error\right\}} \leq \underbrace{\sqrt{\frac{1}{2}\sum_{i\in\Lambda_r^c} a_i^2\omega_i^2(1+\exp(-2\omega_i^2))}}_{\left\{bound\right\}}$$

- If $\omega_i = \omega$ for all $1 \leq i \leq d$, we have

$$\arg\min_{P_r}\left\{error\right\} = \arg\min_{P_r}\left\{bound\right\}$$

but when $\omega \to \infty$,

$$\sqrt{\sum_{i\in\Lambda_r^c} a_i^2} \quad \xleftarrow[\omega\to\infty]{} \quad \min_{P_r}\left\{error\right\} \leq \min_{P_r}\left\{bound\right\} \quad \xrightarrow[\omega\to\infty]{} \quad \infty$$

An analytical example

$$\underbrace{\sqrt{\frac{1}{2}\sum_{i\in\Lambda_r^c} a_i^2(1-\exp(-2\omega_i^2))}}_{\left\{error\right\}} \leq \underbrace{\sqrt{\frac{1}{2}\sum_{i\in\Lambda_r^c} a_i^2\omega_i^2(1+\exp(-2\omega_i^2))}}_{\left\{bound\right\}}$$

- If $\omega_i = \omega$ for all $1 \leq i \leq d$, we have

$$\arg\min_{P_r}\left\{error\right\} = \arg\min_{P_r}\left\{bound\right\}$$

but when $\omega \to \infty$,

$$\sqrt{\sum_{i\in\Lambda_r^c} a_i^2} \quad \xleftarrow[\omega\to\infty]{} \quad \min_{P_r}\left\{error\right\} \leq \min_{P_r}\left\{bound\right\} \quad \xrightarrow[\omega\to\infty]{} \quad \infty$$

- If $\omega_i = a_i^{-2} \geq 1$:

$$\arg\max_{P_r}\left\{error\right\} = \arg\min_{P_r}\left\{bound\right\}$$

A numerical example

A numerical example

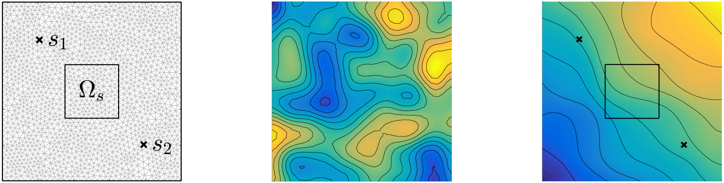
- Diffusion problem on $\Omega = [0,1]^2$: $\begin{cases} \nabla \cdot \kappa \nabla u &= 0 & \text{in } \Omega \\ u &= x + y & \text{on } \partial\Omega \end{cases}$
- Random diffusion field κ, log-normal distribution.

A numerical example

- Diffusion problem on $\Omega = [0,1]^2$: $\begin{cases} \nabla \cdot \kappa \nabla u = 0 & \text{in } \Omega \\ u = x + y & \text{on } \partial\Omega \end{cases}$
- Random diffusion field κ, log-normal distribution.
- After finite element discretization:

$$x = \log(\kappa) \in \mathbb{R}^{3252} \quad \text{and} \quad \mu = \mathcal{N}(0, \Sigma)$$

(a) mesh, 3252 elements (b) log. diffusion field (c) solution

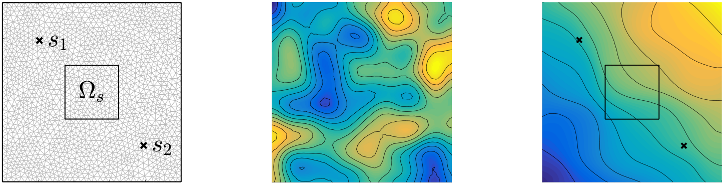


A numerical example

- Diffusion problem on $\Omega = [0,1]^2$: $\begin{cases} \nabla \cdot \kappa \nabla u &= 0 & \text{in } \Omega \\ u &= x + y & \text{on } \partial\Omega \end{cases}$
- Random diffusion field κ, log-normal distribution.
- After finite element discretization:

$$x = \log(\kappa) \in \mathbb{R}^{3252} \quad \text{and} \quad \mu = \mathcal{N}(0, \Sigma)$$

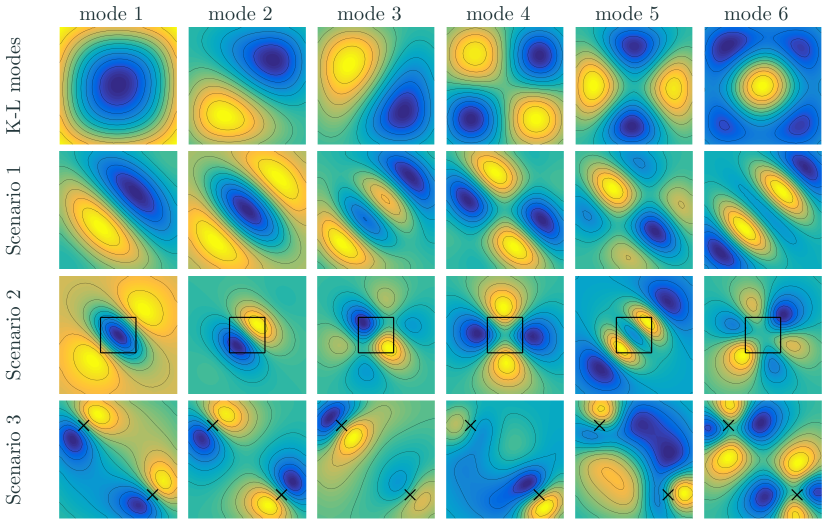
(a) mesh, 3252 elements (b) log. diffusion field (c) solution



Three scenarios

1. $f : x \mapsto u \quad \in V \subset H^1(\Omega)$
2. $f : x \mapsto u_{|\Omega_s} \quad \in V \subset H^1(\Omega_s)$
3. $f : x \mapsto (u_{|s_1}, u_{|s_2}) \quad \in V = \mathbb{R}^2$ (canonical norm)

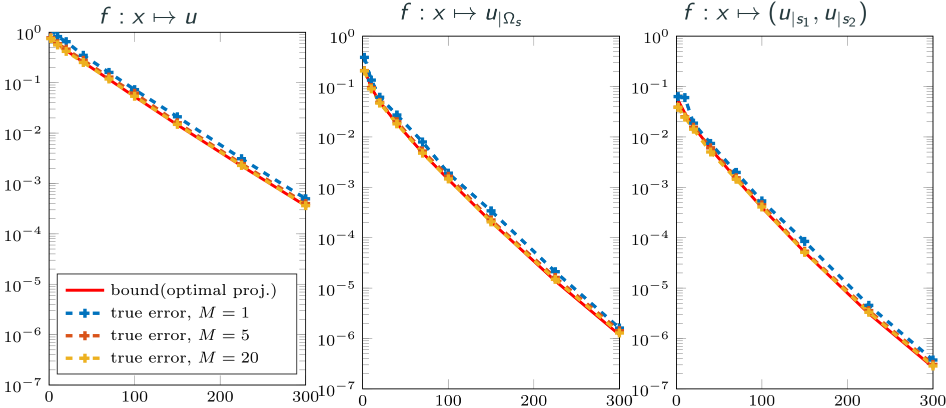
Modes $v_1, v_2, \dots$



$$\mathrm{Im}(P_r) = \mathrm{span}\{v_1, v_2, \dots, v_r\}$$

Approximation of the conditional expectation assuming H is known

$$\mathbb{E}_\mu(f|\sigma(P_r)) \;\approx\; \hat{F}_r : x \mapsto \frac{1}{M}\sum_{k=1}^{M} f(P_r x + (I_d - P_r)Y_i), \qquad Y_i \overset{iid}{\sim} \mu$$

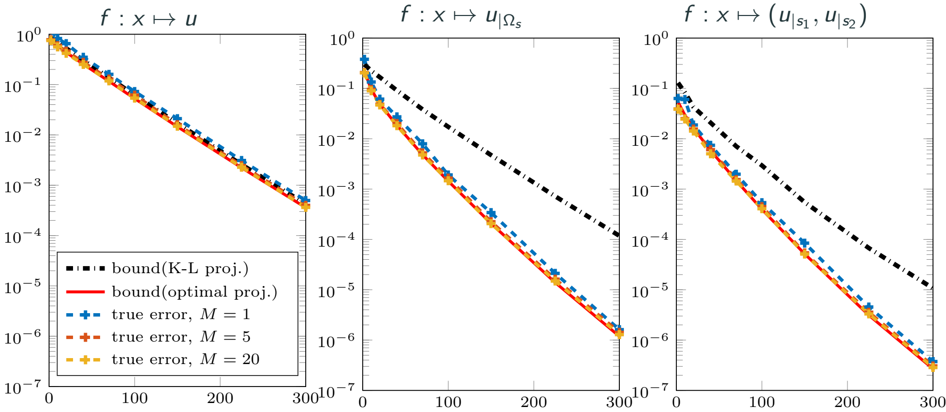


$$\|f - \hat{F}_r\| = function(r)$$

We can show that

$$\mathbb{E}\left(\|f - \hat{F}_r\|^2\right) \leq (1 + M^{-1})\ \mathrm{trace}(\Sigma(I_d - P_r^T)H(I_d - P_r))$$

$$\mathbb{E}_\mu(f|\sigma(P_r)) \approx \hat{F}_r : x \mapsto \frac{1}{M}\sum_{k=1}^{M} f(P_r x + (I_d - P_r)Y_i), \quad Y_i \overset{iid}{\sim} \mu$$



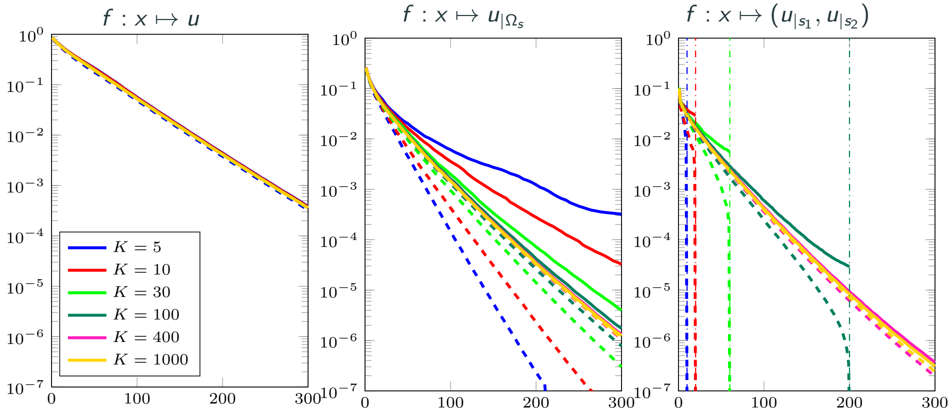
$$\|f - \hat{F}_r\| = function(r)$$

We can show that

$$\mathbb{E}\left(\|f - \hat{F}_r\|^2\right) \leq (1 + M^{-1})\ \mathrm{trace}(\Sigma(I_d - P_r^T)H(I_d - P_r))$$

Approximation of H to get the projector

$$H \approx \hat{H} = \frac{1}{K}\sum_{k=1}^{K}(\nabla f(X_i))^*(\nabla f(X_i)), \quad X_i \overset{iid}{\sim} \mu$$



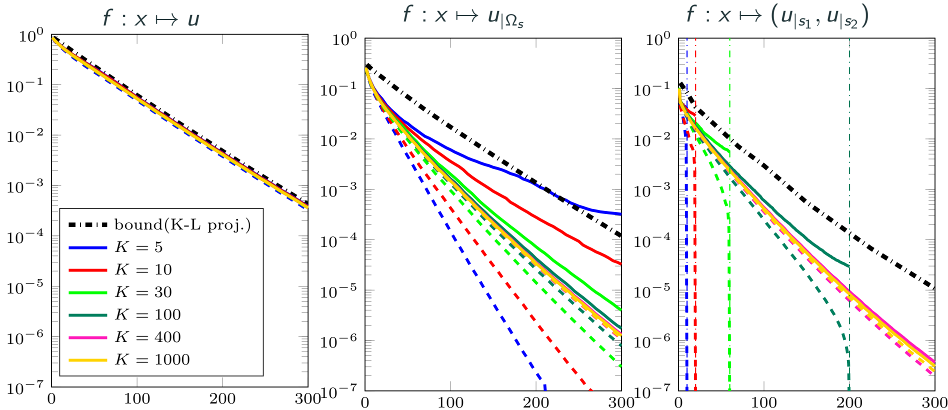
$$\sqrt{\mathrm{trace}(\Sigma(I_d - \hat{P}_r^T)\hat{H}(I_d - \hat{P}_r))} = function(r)$$ (dashed curves)

$$\sqrt{\mathrm{trace}(\Sigma(I_d - \hat{P}_r^T)H(I_d - \hat{P}_r))} = function(r)$$ (solid curves)

Notice that $\mathrm{rank}(\hat{H}) \leq K\ \mathrm{rank}(\nabla f(X))$

Approximation of H to get the projector

$$H \approx \hat{H} = \frac{1}{K}\sum_{k=1}^{K} (\nabla f(X_i))^*(\nabla f(X_i)), \quad X_i \overset{iid}{\sim} \mu$$



$$\sqrt{\mathrm{trace}(\Sigma(I_d - \hat{P}_r^T)\hat{H}(I_d - \hat{P}_r))} = function(r) \qquad \text{(dashed curves)}$$

$$\sqrt{\mathrm{trace}(\Sigma(I_d - \hat{P}_r^T)H(I_d - \hat{P}_r))} = function(r) \qquad \text{(solid curves)}$$

Notice that $\mathrm{rank}(\hat{H}) \leq K \ \mathrm{rank}(\nabla f(X))$

Conclusion

Summary

- Methodology: minimize an upper bound derived from Poincaré inequalities
- Provides a certified error indicator.
- Performs generally better than the truncated Karhunen-Loève.
- Fundamentally gradient-based...

For more information:

O. Zahm, P. Constantine, C. Prieur and Y. Marzouk
Gradient-based dimension reduction of multivariate vector-valued functions.
(2018) preprint: hal-01701425.

Thank you !