



Beyond Petrov-Galerkin projection by using “ multi-space ” priors

Cédric Herzet, Mamadou Diallo, Patrick Héas

► To cite this version:

Cédric Herzet, Mamadou Diallo, Patrick Héas. Beyond Petrov-Galerkin projection by using “ multi-space ” priors. European Conference on Numerical Mathematics and Advanced Applications (ENUMATH), Sep 2017, Vos, Norway. hal-02173637

HAL Id: hal-02173637

<https://inria.hal.science/hal-02173637>

Submitted on 4 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Beyond Petrov-Galerkin projection by using « multi-space » priors

Cédric Herzet
Inria, France

Joint work with M. Diallo & P. Héas

The target problem...

Find $\mathbf{h}^* \in \mathcal{H}$ such that $a(\mathbf{h}^*, \mathbf{h}) = b(\mathbf{h}) \quad \forall \mathbf{h} \in \mathcal{H}$

where $\mathcal{H}$ is a Hilbert space ($\langle \cdot, \cdot \rangle$ and $\|\cdot\|$)
 $a : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$ is a bilinear operator
 $b : \mathcal{H} \rightarrow \mathbb{R}$ is a linear operator

... and its Petrov-Galerkin approximation

Find $\hat{\boldsymbol{h}}_{\text{PG}} \in V_n$ such that $a(\hat{\boldsymbol{h}}_{\text{PG}}, \boldsymbol{h}) = b(\boldsymbol{h}) \quad \forall \boldsymbol{h} \in Z_n$

where $V_n \subset \mathcal{H}$, $Z_n \subset \mathcal{H}$ are n -dimensional subspaces

The precision of Petrov-Galerkin can be quantified by an « instance optimal property »

$$\left\| \boldsymbol{h}^* - \hat{\boldsymbol{h}}_{\text{PG}} \right\| \leq C(V_n, Z_n) \text{dist}(\boldsymbol{h}^*, V_n),$$

Standard methods constructing V_n often return a set of subspaces and their « widths »

Standard outputs of methods constructing V_n :

$$V_0 \subset V_1 \subset \dots \subset V_n, \quad \dim(V_k) = k$$

such that

$$\text{dist}(\mathbf{h}^*, V_k) \leq \hat{\epsilon}_k, \quad k = 0 \dots n.$$

E.g., « reduced basis » methods

The Petrov-Galerkin projection discards most of the available information

Standard outputs of methods constructing V_n :

$$\mathbf{X} \subset \mathbf{X} \subset \dots \subset V_n, \quad \dim(V_k) = k$$

such that

$$\text{dist}(\mathbf{h}^*, \mathbf{X}) \leq \hat{\epsilon}_k, \quad k = 0 \dots n.$$

E.g., « reduced basis » methods



Can we use this information to improve the projection process?

The Petrov-Galerkin projection can be reformulated as a variational problem

$$\hat{\boldsymbol{h}}_{\text{PG}} = \arg \min_{\boldsymbol{h} \in V_n} \sum_{j=1}^n (b_j - \langle \boldsymbol{a}_j, \boldsymbol{h} \rangle)^2$$

where $\text{span} \left(\{\boldsymbol{z}_j\}_{j=1}^n \right) = Z_n$
 $\boldsymbol{a}_j$ is the Riesz's representer of $a(\cdot, \boldsymbol{z}_j)$
 $b_j = b(\boldsymbol{z}_j)$

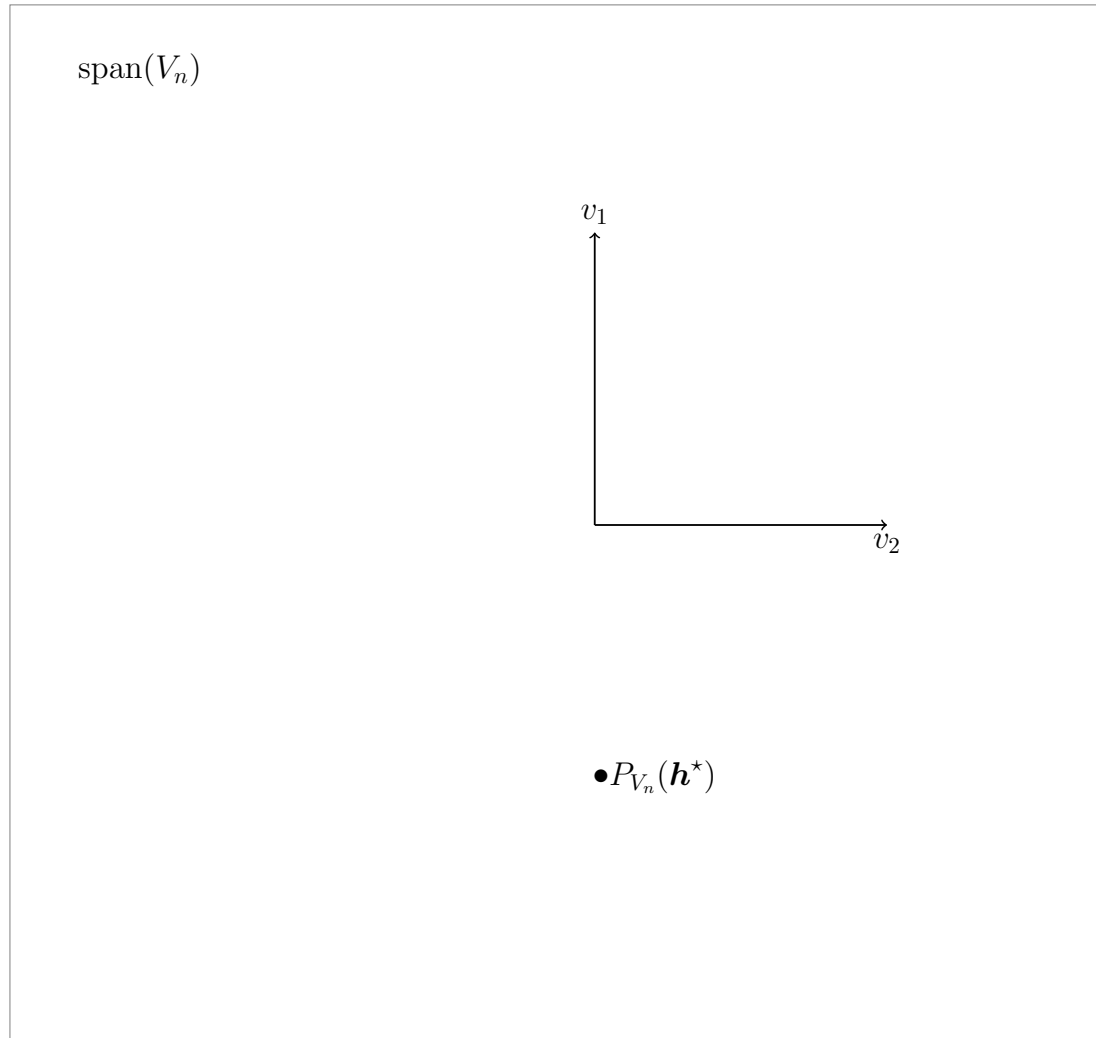
The proposed « multi-space » decoder adds new constraints to the variational problem

$$\hat{\mathbf{h}}_{\text{MS}} = \arg \min_{\mathbf{h} \in V_n} \sum_{j=1}^n (b_j - \langle \mathbf{a}_j, \mathbf{h} \rangle)^2,$$

subject to $\text{dist}(\mathbf{h}, V_k) \leq \hat{\epsilon}_k, \quad k = 0 \dots n - 1.$

See [*Binev et al., SIAM JUQ 17*] for a related work.

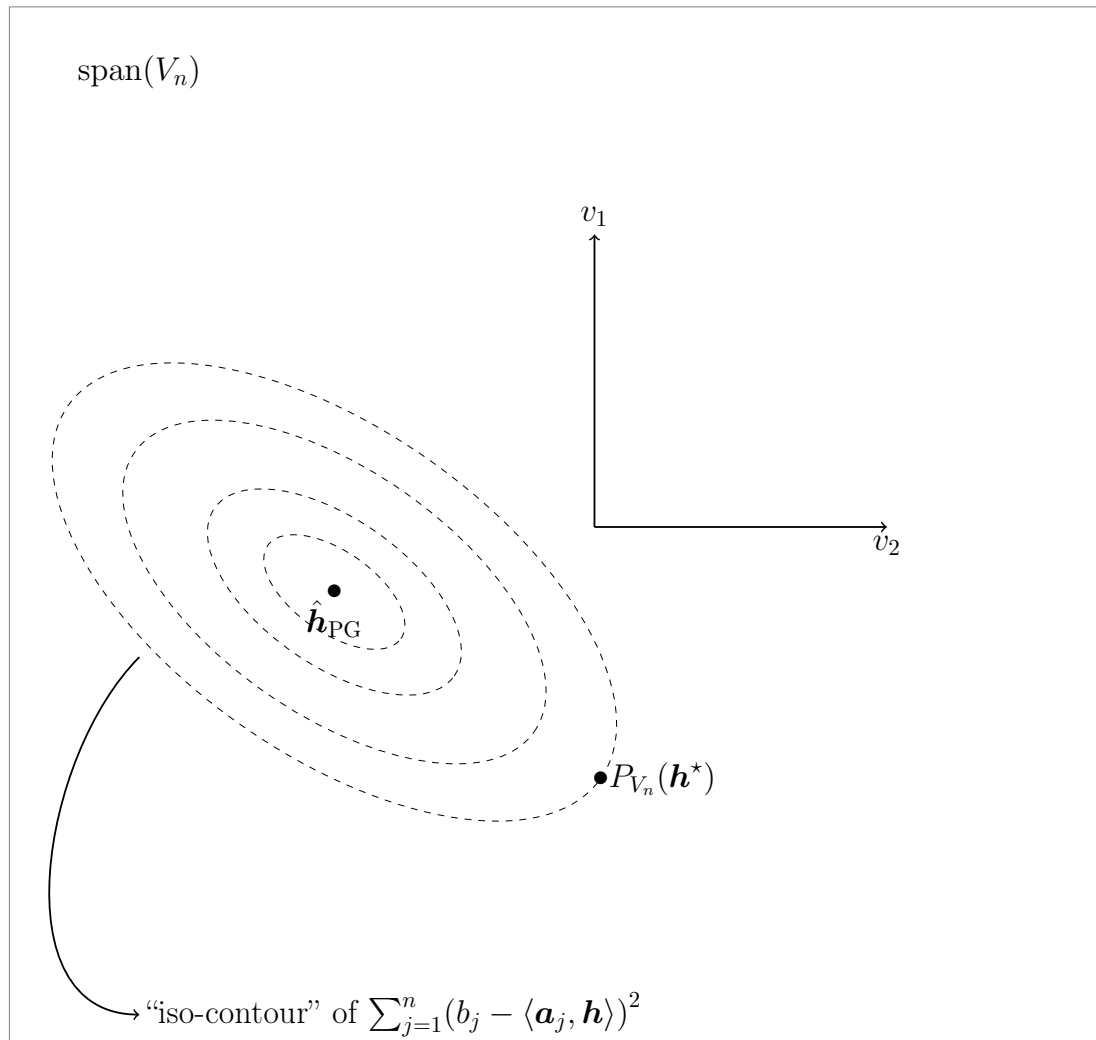
A graphical representation of the problem



$$n = 2$$

$$V_k = \text{span} \left(\{\mathbf{v}_i\}_{i=1}^k \right)$$

A graphical representation of the problem

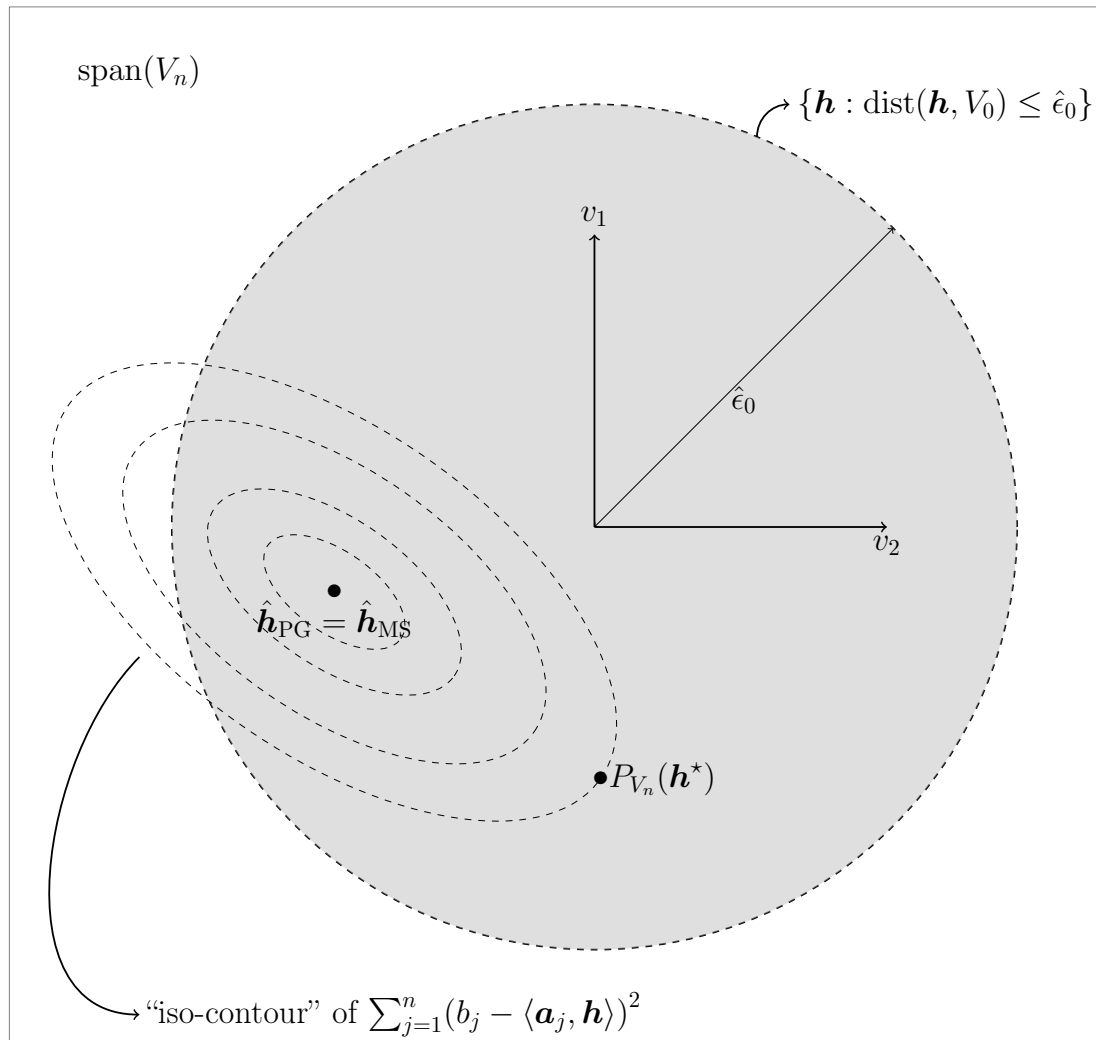


$$n = 2$$

$$V_k = \text{span} \left(\{\mathbf{v}_i\}_{i=1}^k \right)$$

The shape of the iso-contours depends on $\mathbf{G} = [\langle \mathbf{a}_i, \mathbf{v}_j \rangle]_{ij}$

A graphical representation of the problem

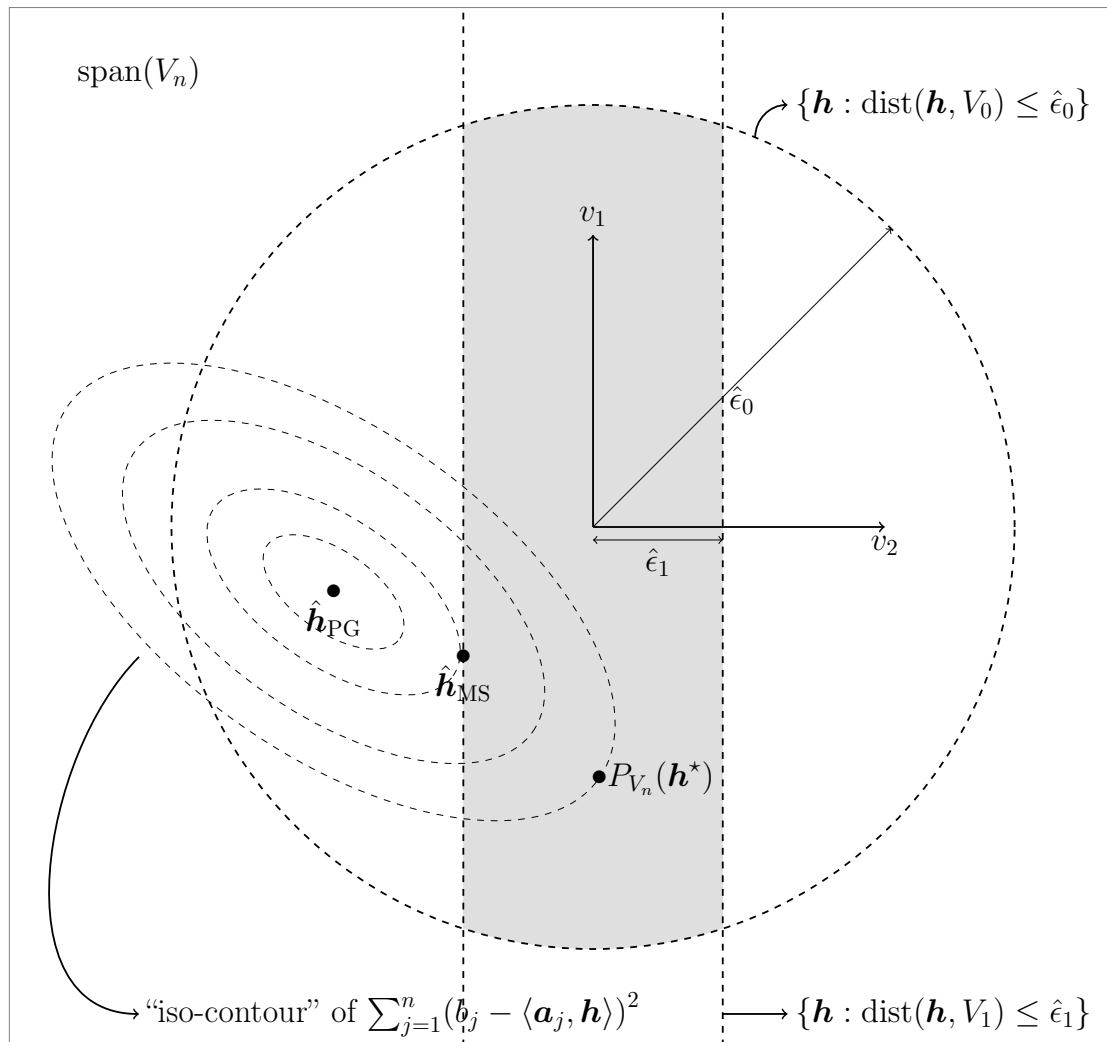


$$n = 2$$

$$V_k = \text{span} \left(\{\mathbf{v}_i\}_{i=1}^k \right)$$

The shape of the iso-contours depends on $\mathbf{G} = [\langle \mathbf{a}_i, \mathbf{v}_j \rangle]_{ij}$

A graphical representation of the problem



$$n = 2$$

$$V_k = \text{span} \left(\{\mathbf{v}_i\}_{i=1}^k \right)$$

The shape of the iso-contours depends on $\mathbf{G} = [\langle \mathbf{a}_i, \mathbf{v}_j \rangle]_{ij}$

The feasibility region depends on $\{V_k\}_{k=1}^n$ and $\{\hat{\epsilon}_k\}_{k=1}^n$



Can we give some guarantee on the performance of the « multi-space » decoder?

Instance optimality properties

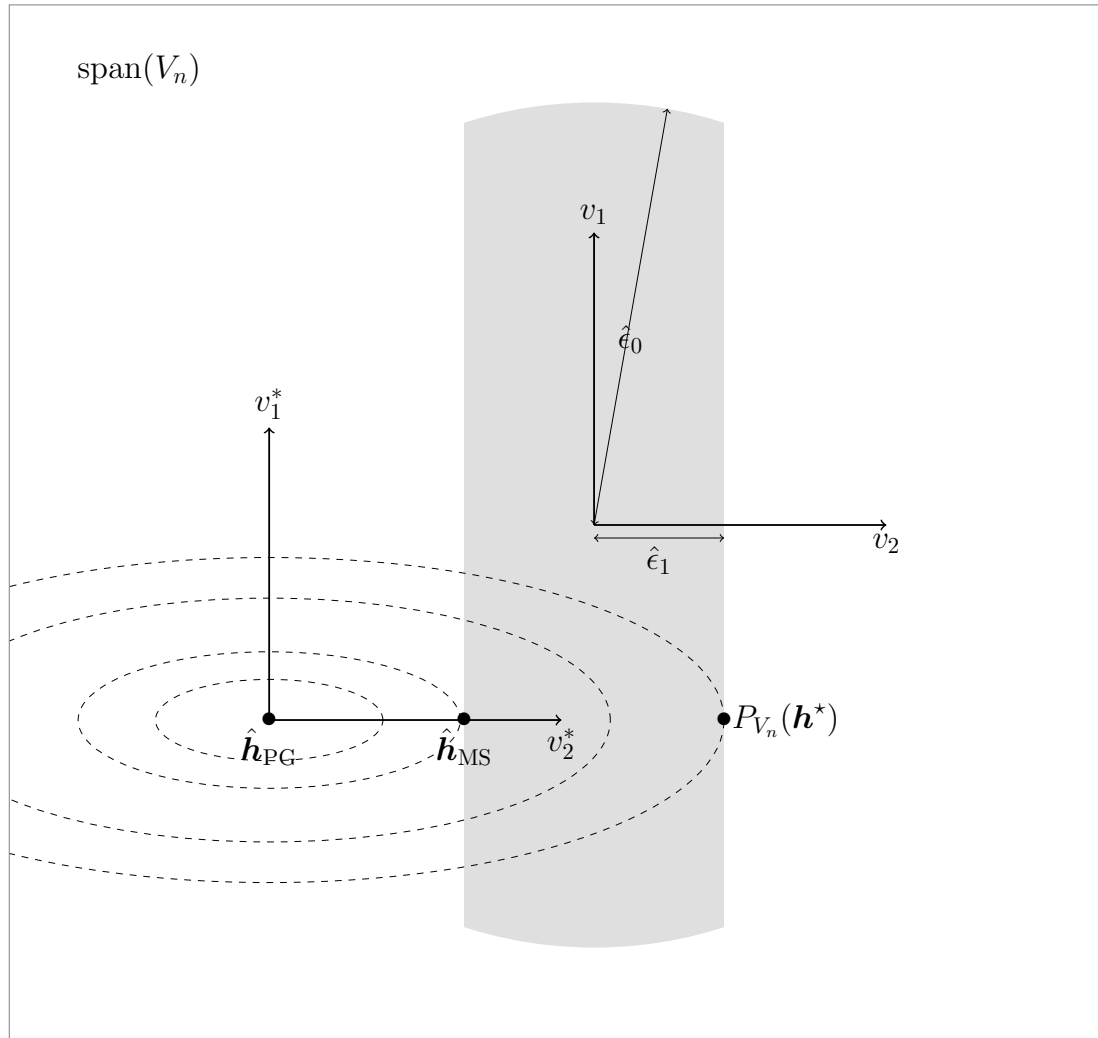
- *Petrov-Galerkin*: $\left\| \mathbf{h}^\star - \hat{\mathbf{h}}_{\text{PG}} \right\| \leq C(\mathbf{G}) \text{dist}(\mathbf{h}^\star, V_n),$

- « *Multi-space* » decoder:

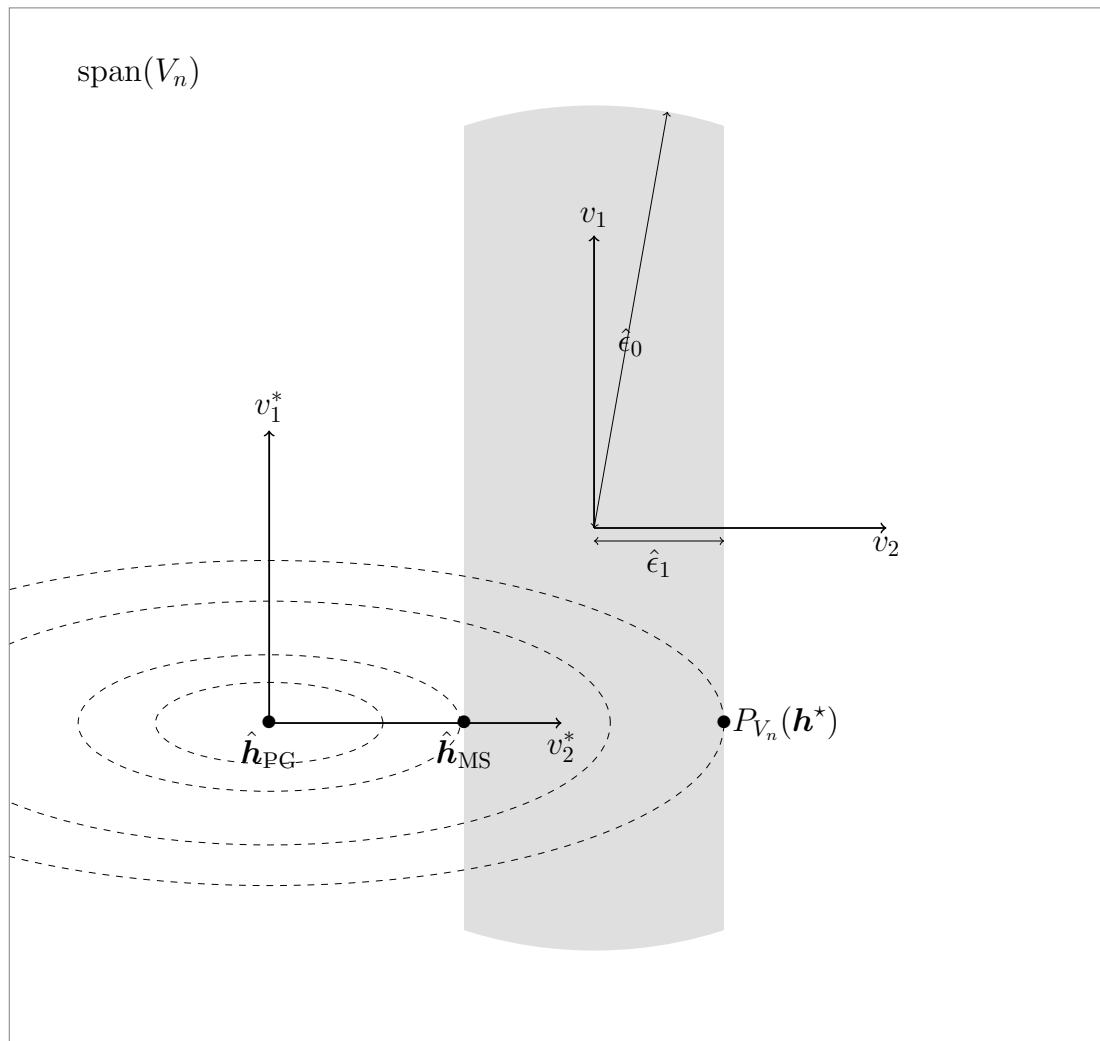
$$\left\| \mathbf{h}^\star - \hat{\mathbf{h}}_{\text{MS}} \right\| \leq \left(\sum_{j=\ell+1}^n \delta_j^2 + \rho \delta_\ell^2 + (\text{dist}(\mathbf{h}^\star, V_n))^2 \right)^{\frac{1}{2}}$$

where ℓ and δ_j 's are “easily-computable” quantities only
depending on $\mathbf{G} = [\langle \mathbf{a}_i, \mathbf{v}_j \rangle]_{i,j}$, $\{\hat{\epsilon}_k\}_{k=1}^{n-1}$ and $\{\text{dist}(\mathbf{h}^\star, V_k)\}_{k=1}^n$

Particularization of the instance optimality bound to some examples



Particularization of the instance optimality bound to some examples



$$\hat{\epsilon}_j = \begin{cases} 1 & j = 0 \dots n-3, \\ \epsilon^{\frac{1}{2}} & j = n-2, n-1, \\ \epsilon & j = n, \end{cases}$$

$$\text{dist}(\mathbf{h}^*, V_k) = \hat{\epsilon}_j$$

$$\sigma_j = \text{sing. values of } \mathbf{G}$$

$$= \begin{cases} 1 & j = 1 \dots n-3, \\ \epsilon^{\frac{1}{2}} & j = n-2, n-1, \\ \epsilon & j = n. \end{cases}$$

$$\|\mathbf{h}^* - \hat{\mathbf{h}}_{PG}\| \leq 1$$

$$\|\mathbf{h}^* - \hat{\mathbf{h}}_{MS}\| \leq 3\epsilon^{\frac{1}{2}}$$



Quid of the computational complexity?

PG projection can be carried out efficiently with a complexity $\mathcal{O}(n^2)$ per iteration

$$\hat{\mathbf{h}}_{\text{PG}} = \arg \min_{\mathbf{h} \in V_n} \sum_{j=1}^n (b_j - \langle \mathbf{a}_j, \mathbf{h} \rangle)^2$$

« Least square » problem: can be solved efficiently via gradient-based methods with a complexity $\mathcal{O}(n^2)$ per iteration.

Our decoder can also be implemented with a complexity $\mathcal{O}(n^2)$ per iteration

Our problem can be rewritten as:

$$\begin{aligned} \hat{\mathbf{h}}_{\text{MS}} = \arg \min_{\mathbf{h} \in V_n} & \sum_{j=1}^n (b_j - \langle \mathbf{a}_j, \mathbf{h} \rangle)^2, \\ \text{subject to } & \|P_{V_k}^\perp(\mathbf{h})\| \leq \hat{\epsilon}_k, \quad k = 0 \dots n-1. \end{aligned}$$

We use the « Alternating Directions Method of Multipliers » to solve this convex problem with a complexity $\mathcal{O}(n^2)$ per iteration

Some results

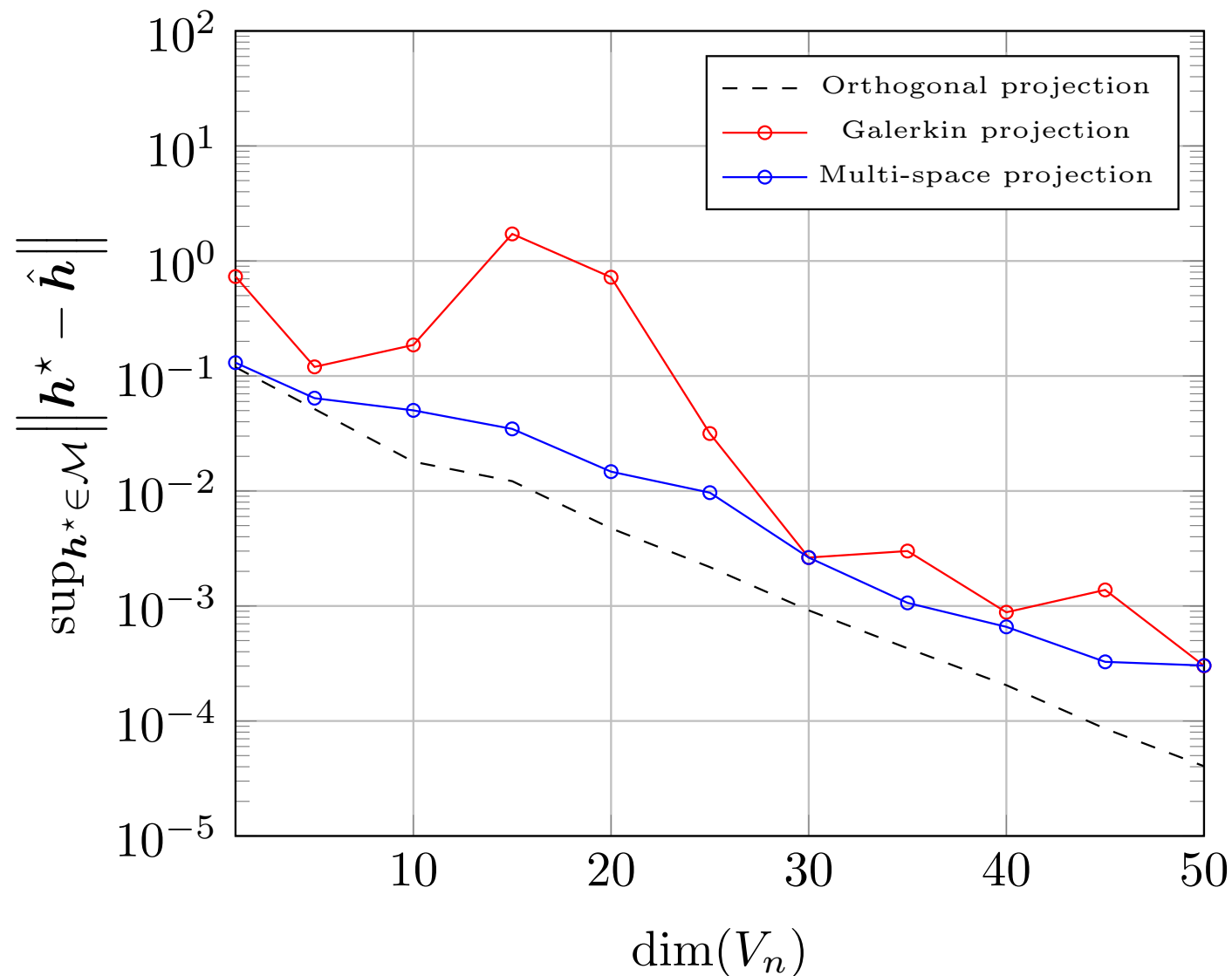
We consider a parametric mass transfert problem

$$\begin{aligned}\mu_1 \Delta \mathbf{h} + \mathbf{b}(\mu_2) \cdot \nabla \mathbf{h} &= \mathbf{s} && \text{in } \Omega \\ \mu_1 \nabla \mathbf{h} \cdot \mathbf{n} &= 0 && \text{in } \partial\Omega\end{aligned}$$

where

$$\begin{aligned}\mathbf{b}(\mu_2) &= [\cos(\mu_2) \sin(\mu_2)]^T \\ \mathbf{s} &= \exp\left(-\frac{\|\mathbf{x} - \mathbf{m}\|_2^2}{2\sigma^2}\right) \\ \mu_1 &= [0.03, 0.05] \\ \mu_2 &= [0, 2\pi]\end{aligned}$$

The multi-space decoder improves the worst-case approximation error over PG



Our contributions

- We exploit additional information in the Petrov-Galerkin projection
- We derive an instance optimal guarantee for our « multi-space » decoder
- We propose an efficient implementation for the « multi-space » decoder

Thank you for your attention.

Appendix

Main Theorem (IOP for multi-space decoder)

Theorem 1. *The solution of the “multi-space” decoder satisfies:*

$$\|\mathbf{h}^* - \hat{\mathbf{h}}_{\text{MS}}\| \leq \left(\sum_{j=\ell+1}^n \delta_j^2 + \rho \delta_\ell^2 + (\text{dist}(\mathbf{h}^*, V_n))^2 \right)^{\frac{1}{2}},$$

where ℓ is the largest integer such that

$$\sum_{j=\ell}^n \sigma_j^2 \delta_j^2 \geq 4\gamma^2 (\text{dist}(\mathbf{h}^*, V_n))^2,$$

and $\rho \in [0, 1]$ is defined as

$$\rho \sigma_\ell^2 \delta_\ell^2 + \sum_{j=\ell+1}^n \sigma_j^2 \delta_j^2 = 4\gamma^2 (\text{dist}(\mathbf{h}^*, V_n))^2.$$

$$\gamma = \sup_{\mathbf{h} \in V^\perp, \|\mathbf{h}\|=1} \left(\sum_{j=1}^n \langle \mathbf{a}_j, \mathbf{h} \rangle^2 \right)^{\frac{1}{2}},$$

$$\mathbf{G} = [\langle \mathbf{a}_i, \mathbf{v}_j \rangle]_{i,j}$$

$\mathbf{X}$ = right singular vectors of $\mathbf{G}$,

σ_k = k th singular value of $\mathbf{G}$,

$$\delta_j = \sum_{k=1}^n |x_{kj}| (\hat{\epsilon}_{k-1} + \text{dist}(\mathbf{h}^*, V_k)).$$

ADMM (first step): problem reformulation

$$\hat{\mathbf{h}}_{\text{MS}} = \arg \min_{\mathbf{h}_n \in V_n} \sum_{j=1}^n (b_j - \langle \mathbf{a}_j, \mathbf{h}_n \rangle)^2,$$

subject to $\begin{cases} \|P_{V_k}^\perp(\mathbf{h}_k)\| \leq \hat{\epsilon}_k, & k = 0 \dots n-1, \\ \mathbf{h}_0 = \mathbf{h}_1 = \dots = \mathbf{h}_n. \end{cases}$



We add n new variables and n new constraints.

ADMM (second step): main recursions

$$\mathbf{h}_n^{(l+1)} = \arg \min_{\mathbf{h}_n \in V_n} \sum_{j=1}^n (b_j - \langle \mathbf{a}_j, \mathbf{h}_n \rangle)^2 + \frac{\rho}{2} \sum_{k=1}^{n-1} \left\| \mathbf{h}_n - \mathbf{h}_k^{(l)} + \mathbf{u}_k^{(l)} \right\|^2,$$

$$\mathbf{h}_k^{(l+1)} = \arg \min_{\mathbf{h}_k} \left\| \mathbf{h}_k - \mathbf{h}_n^{(l+1)} + \mathbf{u}_k^{(l)} \right\|^2 \text{ subject to } \|P_{V_k}^\perp(\mathbf{h}_k)\| \leq \hat{\epsilon}_k,$$

$$\mathbf{u}_k^{(l+1)} = \mathbf{u}_k^{(l)} + \mathbf{h}_n^{(l+1)} - \mathbf{h}_k^{(l+1)}$$



$\mathbf{h}_k^{(l)} \in V_n$ and $\mathbf{u}_k^{(l)} \in V_n \ \forall k, l$

They can thus be represented by n -dimensional vectors.