



# The generalized Empirical Interpolation Method: stability theory on Hilbert spaces with an application to the Stokes equation

Yvon Maday, Olga Mula, Anthony T. Patera, Masayuki Yano

## ► To cite this version:

Yvon Maday, Olga Mula, Anthony T. Patera, Masayuki Yano. The generalized Empirical Interpolation Method: stability theory on Hilbert spaces with an application to the Stokes equation. *Computer Methods in Applied Mechanics and Engineering*, 2015, 287, pp.310 - 334. 10.1016/j.cma.2015.01.018 . hal-01032168v2

**HAL Id: hal-01032168**

**<https://hal.sorbonne-universite.fr/hal-01032168v2>**

Submitted on 5 Dec 2014

**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.

# The Generalized Empirical Interpolation Method: stability theory on Hilbert spaces with an application to the Stokes equation

Maday, Y.<sup>a,b,c,e</sup>, Mula, O.<sup>a,d,e,\*</sup>, Patera, A.T.<sup>f</sup>, Yano, M.<sup>f</sup>

<sup>a</sup>*Sorbonne Universités, UPMC Univ Paris 06 and CNRS UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France.*

<sup>b</sup>*Institut Universitaire de France*

<sup>c</sup>*Division of Applied Mathematics, Brown University, Providence RI, USA.*

<sup>d</sup>*CEA Saclay, DEN/DANS/DM2S/SERMA/LLPR, 91191 Gif-Sur-Yvette CEDEX - France*

<sup>e</sup>*LRC MANON – CEA/DEN/DANS/DM2S and UPMC-CNRS/LJLL.*

<sup>f</sup>*Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

---

## Abstract

The Generalized Empirical Interpolation Method (GEIM) is an extension first presented by Maday and Mula in [1] in 2013 of the classical empirical interpolation method (presented in 2004 by Barrault, Maday, Nguyen and Patera in [2]) where the evaluation at interpolating points is replaced by the more practical evaluation at interpolating continuous linear functionals on a class of Banach spaces. As outlined in [1], this allows to relax the continuity constraint in the target functions and expand both the application domain and the stability of the approach. In this paper, we present a thorough analysis of the concept of stability condition of the generalized interpolant (the Lebesgue constant) by relating it to an inf-sup problem in the case of Hilbert spaces. In the second part of the paper, it will be explained how GEIM can be employed to monitor in real time physical experiments by providing an online accurate approximation of the phenomenon that is computed by combining the acquisition of a minimal number, optimally placed, measurements from the processes with their mathematical models (parameter-dependent PDEs). This idea is illustrated through a parameter dependent Stokes problem in which it is shown that the pressure and velocity fields can efficiently be reconstructed with a relatively low-dimensional interpolation space.

**Keywords:** empirical interpolation; generalized empirical interpolation; reduced basis; model order reduction; stability; Stokes equations

---

## Introduction

Let  $X$  be a Banach space of functions defined over a domain  $\bar{\Omega} \subset \mathbb{R}^d$  (or  $\mathbb{C}^d$ ). Let  $(X_n)_{n \in \mathbb{N}^*}$ ,  $X_n \subset X$ , be a family of finite dimensional spaces,  $\dim X_n = n$ , and let  $(S_n)_{n \in \mathbb{N}^*}$  be an associated family of sets of points:  $S_n = \{x_i^n\}_{i=1}^n$ , with  $x_i^n \in \bar{\Omega}$ . The problem of interpolating any function  $f \in X$  has traditionally been stated as:

$$\text{''Find } f_n \in X_n \text{ such that } f_n(x_i^n) = f(x_i^n), \forall i \in \{1, \dots, n\}\text{''}, \quad (1)$$

where we note that it is implicitly required that  $X$  is a Banach space of continuous functions. The most usual approximation in this sense is the Lagrangian interpolation, where the interpolating spaces  $X_n$  are of polynomial nature (spanned by plain polynomials, rational functions, Fourier series...) and the question on how to appropriately select the interpolating points in this case has broadly been explored. Although there exists still nowadays open issues on Lagrangian interpolation (see, e.g. [3]), it is also interesting to look for extensions of this procedure in which the interpolating spaces  $X_n$  are not necessarily of polynomial nature. The search for new interpolating spaces  $X_n$  is therefore

---

\*Corresponding author: O. Mula (Tel: +33 6 11 18 73 89 or +49 24 18 09 65 92)

Email address: mula-hernandez@ann.jussieu.fr (Mula, O.)

linked with the question on how to optimally select the interpolating points in this case and how to obtain a process that is at least stable and close to the best approximation in some sense.

Although several procedures have been explored in this direction (we refer to [4], [5] and also to the kriging studies in the stochastic community such as [6]), of particular interest for the present work is the Empirical Interpolation Method (EIM, [2], [7], [8]) that has been developed in the broad framework where the functions  $f$  to approximate belong to a compact set  $F$  of continuous functions ( $X = C^0(\Omega)$ ). The structure of  $F$  is supposed to make any  $f \in F$  be approximable by finite expansions of small size. This is quantified by the Kolmogorov  $n$ -width  $d_n(F, X)$  of  $F$  in  $X$  (see definition (2) below) whose smallness measures the extent to which  $F$  can be approximated by some finite dimensional space  $X_n \subset X$  of dimension  $n$ . Unfortunately, in general, the best approximation  $n$ -dimensional space is not known and, in this context, the Empirical Interpolation Method aims to build a family of interpolating spaces  $X_n$  with satisfactory approximation properties together with sets of interpolating points  $S_n$  such that the interpolation is well posed. This is done by a greedy algorithm on both the interpolating points and the interpolating selected functions  $\varphi_i$  (see [2]). This procedure has the main advantage of being constructive, i.e. the sequence of interpolating spaces ( $X_n$ ) and interpolating points ( $S_n$ ) are hierarchically defined and the procedure can easily be implemented by recursion.

A recent extension of this interpolation process consists in generalizing the evaluation at interpolating points by application of a class of interpolating continuous linear functionals chosen in a given dictionary  $\Sigma \subset \mathcal{L}(X)$ . This gives rise to the so-called Generalized Empirical Interpolation Method (GEIM). In this new framework, the particular case where the space  $X = L^2(\Omega)$  was first studied in [1]. We also mention the preliminary works of [9] in which the authors introduced the use of linear functionals in EIM in a finite dimensional framework. In the present paper, we will start by revisiting the foundations of the theory in order to show that GEIM holds for Banach spaces  $X$  (Section 1). The concept of stability condition (Lebesgue constant,  $\Lambda_n$ ) of the generalized interpolant will also be introduced.

In the particular case where  $X$  is a Hilbert space, we will provide an interpretation of the generalized interpolant of a function as an oblique projection. This will shed some light in the understanding of GEIM from an approximation theory perspective (Section 2.1). This point of view will be the key to show that the Lebesgue constant is related to an inf-sup problem (Section 2.2) that can be easily computed (Section 3). The derived formula can be seen as an extension of the classical formula for Lagrangian interpolation to Hilbert spaces. It will also be shown that the Greedy algorithm aims to minimize the Lebesgue constant in a sense that will be made precise in Section 2.3. Furthermore, the inf-sup formula that will be introduced will explicitly show that there exists an interaction between the dictionary  $\Sigma$  of linear functionals and the Lebesgue constant. Although it has so far not been possible to derive a general theory about the impact of  $\Sigma$  on the behavior of the Lebesgue constant, we present in Section 4 a first simple example in which this influence is analyzed through numerical simulation.

The last part of the paper (Section 5) will allow to present some more elaborate potential applications of the method with respect to what is presented in [1]. In particular, we will explain how GEIM can be used to build a tool for the real-time monitoring of a physical or industrial process. This will be achieved thanks to the online computation of a generalized interpolant that will approximate the phenomenon under consideration. Its derivation will combine measurements collected on the fly from the process itself with a mathematical model (a parameter dependent PDE) that represents the physical understanding of the process. It will also be explained how the proposed methodology can be helpful for the minimization of the number of sensors required to reconstruct the field variable and also their optimal selection and placement, which are very important issues in engineering. These ideas will be illustrated through a parameter dependent Stokes problem for  $X = (H^1(\Omega))^2 \times L_0^2(\Omega)$ , where  $L_0^2(\Omega)$  is the space of the  $L^2(\Omega)$  functions with zero mean over  $\Omega$ .

Taking advantage of this idea, we will outline in the conclusion how the method could be used to build an adaptive tool for the supervision of experiments that could distinguish between normal and accidental conditions. We believe that this tool could help in taking real-time decisions regarding the security of processes.

## 1. The Generalized Empirical Interpolation Method

Let  $X$  be a Banach space of functions defined over a domain  $\Omega \subset \mathbb{R}^d$ , where  $d = 1, 2, 3$ . Its norm is denoted by  $\|\cdot\|_X$ . Let  $F$  be a compact set of  $X$ . With  $M$  being some given large number, we assume that the dimension of the vectorial space spanned by  $F$  (denoted as  $\mathcal{F} = \text{span}\{F\}$ ) is of dimension larger than  $M$ . Our goal is to build a family

of  $n$ -dimensional subspaces of  $\mathcal{X}$  that approximate well enough any element of  $F$ . The rationale of this approach is linked to the notion of  $n$ -width following Kolmogorov [10]:

**Definition 1.1.** Let  $F$  be a subset of some Banach space  $\mathcal{X}$  and  $Y_n$  be a generic  $n$ -dimensional subspace of  $\mathcal{X}$ . The deviation between  $F$  and  $Y_n$  is

$$E(F; Y_n) := \sup_{x \in F} \inf_{y \in Y_n} \|x - y\|_{\mathcal{X}}.$$

The Kolmogorov  $n$ -width of  $F$  in  $\mathcal{X}$  is given by

$$\begin{aligned} d_n(F, \mathcal{X}) &:= \inf\{E(F; Y_n) : Y_n \text{ a } n\text{-dimensional subspace of } \mathcal{X}\} \\ &= \inf_{\substack{Y_n \subset \mathcal{X} \\ \dim Y_n = n}} \sup_{x \in F} \inf_{y \in Y_n} \|x - y\|_{\mathcal{X}}. \end{aligned} \quad (2)$$

The smallness of the  $n$ -width of  $F$  thus measures to what extent the set  $F$  can be approximated by an  $n$ -dimensional subspace of  $\mathcal{X}$ . Several reasons can account for a rapid decrease of  $d_n(F, \mathcal{X})$ : if  $F$  is a set of functions defined over a domain, we can refer to regularity, or even to analyticity, of these functions with respect to the domain variable (as analyzed in the example in [10]). Another possibility — that will actually be used in our numerical application — is when  $F = \{u(\mu, \cdot), \mu \in D\}$ , where  $D$  is a compact set of  $\mathbb{R}^p$  and  $u(\mu, \cdot)$  is the solution of a PDE parametrized by  $\mu$ . The approximation of any element  $u(\mu, \cdot) \in F$  by finite expansions is a classical problem addressed by, among others, reduced basis methods and the regularity of  $u$  in  $\mu$  can also be a reason for having a small  $n$ -width as the results of [11] and [12] show.

Finally, let us also assume that we have at our disposal a dictionary of linear functionals  $\Sigma \subset \mathcal{L}(\mathcal{X})$  with the following properties:

P1:  $\forall \sigma \in \Sigma, \|\sigma\|_{\mathcal{L}(\mathcal{X})} = 1$ .

P2: *Unisolvence property*: If  $\varphi \in \text{span}\{F\}$  is such that  $\sigma(\varphi) = 0, \forall \sigma \in \Sigma$ , then  $\varphi = 0$ .

Given this setting, GEIM aims at building  $M$ -dimensional interpolating spaces  $X_M$  spanned by suitably chosen functions  $\{\varphi_1, \varphi_2, \dots, \varphi_M\}$  of  $F$  together with sets of  $M$  selected linear functionals  $\{\sigma_1, \sigma_2, \dots, \sigma_M\}$  coming from  $\Sigma$  such that any  $\varphi \in F$  is well approximated by its generalized interpolant  $\mathcal{J}_M[\varphi] \in X_M$  defined by the following interpolation property:

$$\forall \varphi \in \mathcal{X}, \quad \mathcal{J}_M[\varphi] \in X_M \text{ such that } \sigma_i(\mathcal{J}_M[\varphi]) = \sigma_i(\varphi), \quad \forall i = 1, \dots, M. \quad (3)$$

**Remark 1.2.** Since only some elements of the dictionary  $\Sigma$  are going to be selected, note that  $\Sigma$  consists of "candidate" functionals and only the selected elements will actually be implemented.

The definition of GEIM in the sense of (3) raises several questions:

- is there an optimal selection for the linear functionals  $\sigma_i$  within the dictionary  $\Sigma$ ?
- is there a constructive optimal selection for the functions  $\varphi_i \in F$ ?
- given a set of linearly independent functions  $\{\varphi_i\}_{i \in [1, M]}$  and a set of continuous linear functionals  $\{\sigma_i\}_{i \in [1, M]}$ , does the interpolant exist in the sense of (3)?
- is the interpolant unique?
- under what hypothesis can we expect the GEIM approximation to converge rapidly to  $\varphi$ ?

In what follows, we provide answers to these questions either with rigorous proofs or with numerical evidences.

The construction of the generalized interpolation spaces  $X_M$  and the selection of the suitable associated linear functionals is recursively performed by following a greedy procedure very similar to the one of the classical EIM. The first selected function is, e.g.,

$$\varphi_1 = \arg \sup_{\varphi \in F} \|\varphi\|_{\mathcal{X}},$$

that defines  $X_1 = \text{span}\{\varphi_1\}$ . The first interpolating linear functional is

$$\sigma_1 = \arg \sup_{\sigma \in \Sigma} |\sigma(\varphi_1)|.$$

The interpolation operator  $\mathcal{J}_1 : \mathcal{X} \mapsto X_1$  is defined such that (3) is true for  $M = 1$ , i.e.  $\sigma_1(\mathcal{J}_1[\varphi]) = \sigma_1(\varphi)$ , for any  $\varphi \in \mathcal{X}$ . To facilitate the practical computation of the generalized interpolant, we express it in terms of

$$q_1 = \frac{\varphi_1}{\sigma_1(\varphi_1)},$$

which will be the basis function that will be employed for  $X_1$ . In this basis, the interpolant reads

$$\mathcal{J}_1[\varphi] = \sigma_1(\varphi)q_1, \quad \forall \varphi \in \mathcal{X}.$$

We then proceed by induction. For  $M_{\max} < M$  an upper bound fixed *a priori*, assume that, for a given  $1 \leq M < M_{\max}$ , we have selected a set of functions  $\{\varphi_1, \varphi_2, \dots, \varphi_M\}$  and the associated basis functions  $\{q_1, q_2, \dots, q_M\}$  that span  $X_M$ , as well as the interpolating linear functionals  $\{\sigma_1, \sigma_2, \dots, \sigma_M\}$ . The generalized interpolant is assumed to be well defined by (3), i.e.,

$$\mathcal{J}_M[\varphi] = \sum_{j=1}^M \alpha_j^M(\varphi)q_j, \quad \varphi \in \mathcal{X},$$

where the coefficients  $\alpha_j^M(\varphi)$ ,  $j = 1, \dots, M$  are given by the interpolation problem

$$\begin{cases} \text{Find } \{\alpha_j^M(\varphi)\}_{j=1}^M \text{ such that:} \\ \sum_{j=1}^M \alpha_j^M(\varphi)B_{i,j}^M = \sigma_i(\varphi), \quad \forall i = 1, \dots, M. \end{cases}$$

where  $B_{i,j}^M$  are the coefficients of the  $M \times M$  matrix  $B^M := (\sigma_i(q_j))_{1 \leq i, j \leq M}$ . We now define

$$\forall \varphi \in F, \quad \varepsilon_M(\varphi) = \|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}}.$$

At the  $M + 1$ -th stage of the greedy algorithm, we choose  $\varphi_{M+1}$  such that

$$\varphi_{M+1} = \arg \sup_{\varphi \in F} \varepsilon_M(\varphi) \tag{4}$$

and

$$\sigma_{M+1} = \arg \sup_{\sigma \in \Sigma} |\sigma(\varphi_{M+1} - \mathcal{J}_M[\varphi_{M+1}])|. \tag{5}$$

The next basis function is then

$$q_{M+1} = \frac{\varphi_{M+1} - \mathcal{J}_M[\varphi_{M+1}]}{\sigma_{M+1}(\varphi_{M+1} - \mathcal{J}_M[\varphi_{M+1}])}.$$

We finally set  $X_{M+1} \equiv \text{span}\{\varphi_j, 1 \leq j \leq M + 1\} = \text{span}\{q_j, 1 \leq j \leq M + 1\}$ . The interpolation operator  $\mathcal{J}_{M+1} : \mathcal{X} \mapsto X_{M+1}$  is given by

$$\mathcal{J}_{M+1}[\varphi] = \sum_{j=1}^{M+1} \alpha_j^{M+1}(\varphi)q_j, \quad \forall \varphi \in \mathcal{X},$$

so as to satisfy (3). The coefficients  $\alpha_j^{M+1}(\varphi)$ ,  $j = 1, \dots, M + 1$ , are therefore given by the interpolation problem

$$\begin{cases} \text{Find } \{\alpha_j^{M+1}(\varphi)\}_{j=1}^{M+1} \text{ such that:} \\ \sum_{j=1}^{M+1} \alpha_j^{M+1}(\varphi)B_{i,j}^{M+1} = \sigma_i(\varphi), \quad \forall i = 1, \dots, M + 1, \end{cases}$$

where  $B^{M+1} = (\sigma_i(q_j))_{1 \leq i, j \leq M+1}$ .

By following exactly the same guidelines as in [1] where the particular case  $\mathcal{X} = L^2(\Omega)$  was addressed, it can be proven that, in the general case where  $\mathcal{X}$  is a Banach space, the generalized interpolation is well-posed: for any  $1 \leq M < \mathcal{M}$ , the set of functions  $\{q_j, j \in [1, M]\}$  is linearly independent and therefore the space  $X_M$  is of dimension  $M$ . Furthermore, the matrix  $B^M$  is lower triangular with unity diagonal (hence invertible) with off-diagonal entries in  $[-1, 1]$ .

Note that GEIM reduces to EIM if  $\mathcal{X} \subset C^0(\Omega)$  and  $\Sigma$  is composed of Dirac masses. Also, if the cardinality  $\#F$  of  $F$  is finite, then the Greedy algorithm is exact in the sense that  $F \subset X_{\#F}$ . This type of property does not hold in traditional Lagrangian interpolation due to the fact that the interpolating polynomial spaces are used to interpolate continuous functions that are not necessarily of polynomial nature. Finally, note also that the approach can be shortcut if the basis functions are available, in which case the interpolating linear functionals/points are the only output of GEIM/EIM.

It is also important to point out that the current extension of EIM presents two major advantages: first, it allows the interpolation of functions of weaker regularity than  $C^0(\Omega)$ . The second interest is related to the potential applications of GEIM: the use of linear functionals can model in a more faithful manner real sensors involved in physical experiments (indeed, these are in practice no point evaluations as it is usually supposed but rather local averages of some quantity of interest). The potentialities of these two aspects will be illustrated in the numerical application presented in Section 5.

We now state a first result about the interpolation error of GEIM.

**Theorem 1.3** (Interpolation error on a Banach space).  *$\forall \varphi \in \mathcal{X}$ , the interpolation error satisfies*

$$\|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq (1 + \Lambda_M) \inf_{\psi_M \in X_M} \|\varphi - \psi_M\|_{\mathcal{X}}, \quad (6)$$

where

$$\Lambda_M := \|\mathcal{J}_M\|_{\mathcal{L}(\mathcal{X})} = \sup_{\varphi \in \mathcal{X}} \frac{\|\mathcal{J}_M[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}} \quad (7)$$

is the Lebesgue constant in the  $\mathcal{X}$  norm.

*Proof.* The desired result easily follows since for any  $\varphi \in \mathcal{X}$  and any  $\psi_M \in X_M$  we have

$$\begin{aligned} \|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} &= \|[\varphi - \psi_M] - \mathcal{J}_M[\varphi - \psi_M]\|_{\mathcal{X}} \\ &\leq \|I_F - \mathcal{J}_M\|_{\mathcal{L}(\mathcal{X})} \|\varphi - \psi_M\|_{\mathcal{X}} \\ &\leq (1 + \|\mathcal{J}_M\|_{\mathcal{L}(\mathcal{X})}) \|\varphi - \psi_M\|_{\mathcal{X}}, \end{aligned}$$

which yields the desired inequality.  $\square$

The last term in the right hand side of equation (6) is known as the best fit of  $\varphi$  by elements in the space  $X_M$ . However,  $X_M$  does not in general coincide with the optimal  $M$ -dimensional space in the sense that  $X_M \neq X_M^{\text{opt}}$ , with

$$X_M^{\text{opt}} = \arg \inf_{\substack{Y_M \subset \mathcal{X} \\ \dim(Y_M) = M}} E(F, Y_M).$$

This raises the question of the quality of the finite dimensional subspaces  $X_M$  provided by the Greedy selection procedure. It has been proven first in [13] in the case of  $\mathcal{X} = L^2(\Omega)$  and then in [14] in a general Banach space that the interpolating spaces  $X_M$  coming from the Greedy selection procedure of GEIM are quite optimal and that the lack of optimality comes from the Lebesgue constant. The main results are the following (see [14]):

**Theorem 1.4** (See Corollary 3.13 of [14]).

*i) If  $d_M(F, \mathcal{X}) \leq C_0 M^{-\alpha}$ ,  $M = 1, 2, \dots$  and that  $(1 + \Lambda_M) \leq C_{\zeta} M^{\zeta}$ , for any  $M = 1, 2, \dots$ , then, for all  $\beta > 1/2$ , the interpolation error satisfies for any  $\varphi \in F$  the inequality  $\|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq C_{\zeta} C_1 M^{-\alpha+2\zeta+\beta}$ , where*

$$C_1 := \max \left\{ C_0 2^{\frac{2\alpha^2}{\zeta}} \left( \frac{\zeta + \beta}{\beta - \frac{1}{2}} \right)^{\alpha} \max \left( 1; C_{\zeta}^{\frac{\zeta + \beta}{\zeta}} \right); \max_{M=1, \dots, 2\lfloor 2(\zeta + \beta) \rfloor + 1} M^{\alpha - \zeta - \beta} \right\}.$$

ii) If  $(\Lambda_M)$  is a monotonically increasing sequence and if  $d_M(F, X) \leq C_0 e^{-c_1 M^\alpha}$  for any  $M \geq 1$  and with  $C_0 \geq 1$ , then, for any  $\varphi \in F$ , the interpolation error can be bounded as

$$\|\varphi - \mathcal{J}_M[\varphi]\|_X \leq \begin{cases} 4C_0(1 + \Lambda_1), & \text{if } M = 1. \\ \sqrt{2}C_0(1 + \Lambda_M)^2 \sqrt{M} e^{-c_1 2^{-1-2\alpha} M^\alpha}, & \text{if } M \geq 2. \end{cases}$$

As a consequence of this result, the interpolation error of GEIM will converge if the Lebesgue constant "does not increase too fast" in the sense that it allows that the previous upper bounds tend to zero as the dimension  $M$  increases. By following the same lines as in [1], it can be proven that when  $X$  is a Banach space, the Lebesgue constant has the exponential upper-bound

$$\Lambda_M \leq 2^{M-1} \max_{i \in [1, M]} \|q_i\|_X, \quad (8)$$

which implies that the decay of  $d_M(F, X)$  should be exponential in order to converge. However, the behavior of  $(\Lambda_M)_M$  observed in numerical applications (see Section 5) is rather linear and leads us to expect that the upper bound of (8) is far from being optimal in a class of set  $F$  of small Kolmogorov  $n$ -width.

## 2. Further results in the case of a Hilbert space

In this section  $X$  is a Hilbert space of functions where the norm  $\|\cdot\|_X$  is induced by the inner product  $(\cdot, \cdot)_X$ . We will see that in this case the generalized interpolant can be seen as an oblique projection. It will also be proven that we can derive a sharp interpolation error bound in this case. An explicit (and easily computable) formula for the Lebesgue constant will also be obtained and this formula will be used to show that the Greedy algorithm aims at minimizing the Lebesgue constant.

### 2.1. Interpretation of GEIM as an oblique projection

For  $1 \leq j \leq M$ , if  $\sigma_j$  is the  $j^{\text{th}}$ -linear functional selected by the greedy algorithm, let  $w_j$  be its Riesz representation in  $X$ , i.e.  $w_j$  is such that

$$\forall f \in X, \quad \sigma_j(f) = (w_j, f)_X. \quad (9)$$

It follows from the well posedness of the generalized interpolation that  $\{\sigma_1, \dots, \sigma_M\}$  are linearly independent and therefore  $\{w_1, \dots, w_M\}$  are also linearly independent. We will denote by  $W_M$  the  $M$ -dimensional space

$$W_M = \text{span}\{w_1, \dots, w_M\}.$$

For any  $f \in X$ , let  $\Pi_{W_M}[f]$  be the orthogonal projection of  $f$  on  $W_M$ , i.e.

$$\begin{cases} \Pi_{W_M}[f] \in W_M \\ (f - \Pi_{W_M}[f], w)_X = 0, \quad \forall w \in W_M. \end{cases} \quad (10)$$

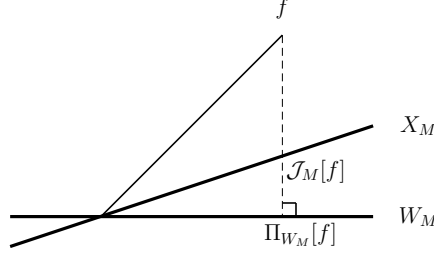
With these notations, we can provide the following interpretation of the generalized interpolant of a function (see Figure 1 for a schematic representation):

**Lemma 2.1.**  $\forall f \in X$ ,  $\mathcal{J}_M[f]$  is an oblique projection onto the space  $X_M$  orthogonal to the space  $W_M$ , i.e.

$$\begin{cases} \mathcal{J}_M[f] \in X_M \\ (\mathcal{J}_M[f] - f, w)_X = 0, \quad \forall w \in W_M. \end{cases} \quad (11)$$

*Proof.* For any  $f \in X$ , the interpolation property reads  $\sigma_j(f) = \sigma_j(\mathcal{J}_M[f])$ , for  $1 \leq j \leq M$ . It is then clear that  $(w_j, f)_X = (w_j, \mathcal{J}_M[f])_X$  and the result easily follows from the fact that  $\{w_1, \dots, w_M\}$  are a basis of  $W_M$ .  $\square$

A direct consequence of Lemma 2.1 is the following result:



**Figure 1:** Interpretation of  $\mathcal{J}_M[f]$  as an oblique projection.

**Corollary 2.2.** *In the particular case where*

$$\Sigma = \left\{ \sigma \in \mathcal{L}(X), \text{ s.t. } \forall f \in \text{span}\{F\}^\perp, \sigma(f) = 0 \right\},$$

*then  $W_M = X_M$  and the resulting generalized interpolant is the orthogonal projection of  $f$  onto the space  $X_M$ , i.e.  $\mathcal{J}_M[f] = \Pi_{W_M}[f]$ .*

*Proof.* First of all, note that, in this setting, there is a bijective mapping between  $\Sigma$  and  $\text{span}\{F\}$  (because the Riesz representation of any  $\sigma \in \Sigma$  is an element of  $\text{span}\{F\}$  and vice-versa). Now, from the arg max definition of  $\sigma_k$  in the greedy algorithm, the Riesz representation of  $\sigma_k$  is the function  $w_k = \varphi_k - \mathcal{J}_{k-1}[\varphi_k]$  for  $k \geq 2$  and  $w_1 = \varphi_1$  if  $k = 1$ . The interpolation property  $\sigma_k(f - \mathcal{J}_M[f]) = 0$  implies in this case that  $(w_k, f - \mathcal{J}_M[f])_X = 0$  for any  $k \in \{1, \dots, M\}$ . But since the family  $\{w_1, \dots, w_M\}$  is a basis of  $X_M$  in this particular case, it follows that  $(f - \mathcal{J}_M[f], w)_X = 0$  for all  $w \in X_M$ .  $\square$

**Remark 2.3.** *The case  $\Sigma = \{\sigma \in \mathcal{L}(X), \text{ s.t. } \forall f \in \text{span}\{F\}^\perp, \sigma(f) = 0\}$  is a theoretical situation that does not usually hold in practical applications. Corollary 2.2 is however a first step towards the theoretical understanding of the impact of the dictionary  $\Sigma$  on the interpolation procedure.*

From Lemma 2.1, note that  $\mathcal{J}_M[f]$  can also be seen as a particular Petrov-Galerkin approximation of the function  $f$  in the case where the approximation space is  $X_M$  and the trial space is  $W_M$ . Indeed, the search for the generalized interpolant can be stated as

$$\begin{cases} \text{Given } f \in X, \text{ find } \mathcal{J}_M[f] \in X_M \text{ such that} \\ (\mathcal{J}_M[f], w)_X = (f, w)_X, \quad \forall w \in W_M. \end{cases} \quad (12)$$

This formulation leads to the classical error estimation:

$$\|f - \mathcal{J}_M[f]\|_X \leq \left(1 + \frac{1}{\beta_M}\right) \inf_{\psi_M \in X_M} \|f - \psi_M\|_X, \quad (13)$$

where  $\beta_M$  is the inf-sup constant

$$\beta_M := \inf_{x \in X_M} \sup_{w \in W_M} \frac{(x, w)_X}{\|x\|_X \|w\|_X}. \quad (14)$$

It will be proven in the next section that the parameter  $1/\beta_M$  is, in fact, equal to the Lebesgue constant  $\Lambda_M$ . We will also see that the error bound provided in relation (13) is slightly suboptimal due to the presence of the coefficient 1 before the parameter  $1/\beta_M$ .

## 2.2. Interpolation error

The interpretation of the generalized interpolant as an oblique projection is useful to derive the following result about the interpolation error:



**Theorem 2.4** (Interpolation error on a Hilbert space).  $\forall \varphi \in \mathcal{X}$ , the interpolation error satisfies the sharp upper bound:

$$\|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq \Lambda_M \inf_{\psi_M \in X_M} \|\varphi - \psi_M\|_{\mathcal{X}} \quad (15)$$

where  $\Lambda_M := \|\mathcal{J}_M\|_{\mathcal{L}(\mathcal{X})} = \sup_{\varphi \in \mathcal{X}} \frac{\|\mathcal{J}_M[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}}$  is the Lebesgue constant in the  $\mathcal{X}$  norm. Furthermore,  $\Lambda_M = \frac{1}{\beta_M}$ , where

$$\beta_M := \inf_{x \in X_M} \sup_{w \in W_M} \frac{(x, w)_{\mathcal{X}}}{\|x\|_{\mathcal{X}} \|w\|_{\mathcal{X}}}. \quad (16)$$

*Proof.* Let  $\nu_M := \inf_{w^\perp \in W_M^\perp} \sup_{\psi^\perp \in X_M^\perp} \frac{(w^\perp, \psi^\perp)_{\mathcal{X}}}{\|w^\perp\|_{\mathcal{X}} \|\psi^\perp\|_{\mathcal{X}}}$ . It immediately follows that

$$\forall w^\perp \in W_M^\perp, \quad \nu_M \|w^\perp\|_{\mathcal{X}} \leq \sup_{\psi^\perp \in X_M^\perp} \frac{(w^\perp, \psi^\perp)_{\mathcal{X}}}{\|\psi^\perp\|_{\mathcal{X}}}.$$

Furthermore, for any  $\varphi \in \mathcal{X}$ , it follows from Lemma 2.1 that  $\varphi - \mathcal{J}_M[\varphi] \in W_M^\perp$ . Then:

$$\nu_M \|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq \sup_{\psi^\perp \in X_M^\perp} \frac{(\varphi - \mathcal{J}_M[\varphi], \psi^\perp)_{\mathcal{X}}}{\|\psi^\perp\|_{\mathcal{X}}}. \quad (17)$$

Besides, for any  $\psi \in X_M$  and any  $\psi^\perp \in X_M^\perp$ :

$$(\varphi - \mathcal{J}_M[\varphi], \psi^\perp)_{\mathcal{X}} = (\varphi - \psi, \psi^\perp)_{\mathcal{X}}. \quad (18)$$

The Cauchy-Schwarz inequality applied to (17) combined with relation (18) yields:

$$\nu_M \|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq \inf_{\psi \in X_M} \|\varphi - \psi\|_{\mathcal{X}}. \quad (19)$$

Next, it can be proven (see the proposition of Appendix A.1) that  $\nu_M = \beta_M$ , which yields the inequality

$$\|\varphi - \mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq \frac{1}{\beta_M} \inf_{\psi \in X_M} \|\varphi - \psi\|_{\mathcal{X}}. \quad (20)$$

The end of the proof consists in showing that

$$\frac{1}{\beta_M} = \Lambda_M = \sup_{\varphi \in \mathcal{X}} \frac{\|\mathcal{J}_M[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}}. \quad (21)$$

This is done by noting first of all that formula (16) implies that

$$\forall \varphi \in \mathcal{X}, \quad \beta_M \|\mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq \sup_{w \in W_M} \frac{(\mathcal{J}_M[\varphi], w)_{\mathcal{X}}}{\|w\|_{\mathcal{X}}} \leq \|\varphi\|_{\mathcal{X}},$$

where we have used the fact that  $(\mathcal{J}_M[\varphi], w)_{\mathcal{X}} = (\varphi, w)_{\mathcal{X}}$  for all  $w \in W_M$  and the Cauchy-Schwarz inequality. Therefore,

$$\forall \varphi \in \mathcal{X}, \quad \|\mathcal{J}_M[\varphi]\|_{\mathcal{X}} \leq \frac{1}{\beta_M} \|\varphi\|_{\mathcal{X}},$$

which yields

$$\Lambda_M \leq \frac{1}{\beta_M}. \quad (22)$$

Let us now denote by  $\psi$  an element of  $X_M$  with norm  $\|\psi\|_X = 1$  such that

$$\sup_{w_M \in W_M} \frac{(\psi, w_M)_X}{\|w_M\|_X} = \beta_M.$$

Using the notation introduced in (10), we denote by  $\Pi_{W_M}[\psi]$  the orthogonal projection of  $\psi$  on  $W_M$ . Similarly,  $\Pi_{W_M^\perp}[\psi]$  is the orthogonal projection of  $\psi$  on  $W_M^\perp$  so that  $(\Pi_{W_M}[\psi], \Pi_{W_M^\perp}[\psi])_X = 0$  and

$$\psi = \Pi_{W_M}[\psi] + \Pi_{W_M^\perp}[\psi].$$

Note that

$$\mathcal{J}_M[\Pi_{W_M}[\psi]] = \psi. \quad (23)$$

Indeed, since  $\psi \in X_M$ , we have that

$$\psi = \mathcal{J}_M[\psi] = \mathcal{J}_M[\Pi_{W_M}[\psi]] + \mathcal{J}_M[\Pi_{W_M^\perp}[\psi]] = \mathcal{J}_M[\Pi_{W_M}[\psi]]$$

because  $\mathcal{J}_M[\Pi_{W_M^\perp}[\psi]] = 0$  in virtue of the interpolation property given in relation (11). In addition to this,

$$\sup_{w_M \in W_M} \frac{(\psi, w_M)_X}{\|w_M\|_X}$$

is achieved for  $w_M = \Pi_{W_M}[\psi]$  and thus

$$\beta_M = \|\Pi_{W_M}[\psi]\|_X. \quad (24)$$

From relations (23) and (24), we infer that  $\Pi_{W_M}[\psi]$  is such that

$$\frac{\|\mathcal{J}_M[\Pi_{W_M}[\psi]]\|_X}{\|\Pi_{W_M}[\psi]\|_X} = \frac{1}{\beta_M} \leq \sup_{\varphi \in X} \frac{\|\mathcal{J}_M[\varphi]\|_X}{\|\varphi\|_X} = \Lambda_M. \quad (25)$$

Relations (22) and (25) yield the final equality  $\frac{1}{\beta_M} = \Lambda_M$ . □

**Remark 2.5.** The link between the Lebesgue constant  $\Lambda_M$  and the inf-sup quantity  $\beta_M$  introduced in Theorem 2.4 shows that  $\Lambda_M$  depends on the dictionary of linear functionals  $\Sigma$  and also on the interpolating space  $X_M$ . Although no theoretical analysis of the impact of these elements has been possible so far, we present in Section 4 a numerical study about the influence of the dictionary  $\Sigma$  in  $\Lambda_M$ .

**Remark 2.6.** Note that, since Theorem 2.4 holds only in Hilbert spaces, formula (16) does not apply to the Lebesgue constant of the classical EIM given that it is defined in the  $L^\infty(\Omega)$  norm. The Hilbertian framework allows nevertheless to consider Dirac masses as linear functionals like in EIM if we place ourselves, e.g., in  $H^2(\Omega)$ .

### 2.3. The Greedy algorithm aims at optimizing the Lebesgue constant

If we look in detail at the steps followed by the Greedy algorithm, once  $X_{M-1}$  and  $W_{M-1}$  have been derived, the construction of  $X_M$  and  $W_M$  starts by adding an element  $\varphi$  to  $X_{M-1}$ . In the Greedy process, this is done following formula (4), but let us analyze what happens when we add any  $\varphi_M \in F$ . The first consequence of its addition is that the resulting inf-sup constant becomes zero:

$$\inf_{\varphi \in \text{span}\{X_{M-1}, \varphi_M\}} \sup_{w \in W_{M-1}} \frac{(\varphi, w)_X}{\|\varphi\|_X \|w\|_X} = 0. \quad (26)$$

Indeed, the addition of  $\varphi_M$  to the interpolating basis functions has the consequence of adding the element  $\tilde{\varphi}_M = \varphi_M - \mathcal{J}_{M-1}[\varphi_M]$  that, by definition, satisfies  $(\tilde{\varphi}_M, w)_X = 0$ ,  $\forall w \in W_{M-1}$ . We thus need to add an element to  $W_{M-1}$  in order to stabilize the inf-sup condition.

Let us denote by  $W$  the set of Riesz representations in  $\mathcal{X}$  of the elements of our dictionary  $\Sigma$ . Since

$$\inf_{\varphi \in \text{span}\{X_{M-1}, \varphi_M\}} \sup_{w \in W_{M-1}} \frac{(\varphi, w)_\mathcal{X}}{\|\varphi\|_\mathcal{X} \|w\|_\mathcal{X}}$$

is reached by  $\tilde{\varphi}_M$ , the aim is to add an element  $w_M$  of  $W$  that maximizes

$$\max_{w \in W} \frac{(\tilde{\varphi}_M, w)_\mathcal{X}}{\|w\|_\mathcal{X}}. \quad (27)$$

Since the elements of the dictionary are of norm 1 (see property P1 above), this corresponds exactly to one of the steps performed by the Greedy algorithm (see equation (5)). Furthermore, from the unisolvence property of our dictionary, the application

$$\begin{aligned} \|\cdot\|_* : \mathcal{X} &\mapsto \mathbb{R} \\ \varphi &\mapsto \max_{w \in W} (\varphi, w)_\mathcal{X} \end{aligned}$$

defines a norm in  $\mathcal{X}$ . Then, formula (27) reads:

$$\max_{w \in W} \frac{(\tilde{\varphi}_M, w)_\mathcal{X}}{\|w\|_\mathcal{X}} = \|\varphi_M - \mathcal{J}_{M-1}[\varphi_M]\|_*.$$

It is thus clear that the choice of  $\varphi_M$  that maximizes the value of  $\beta_M$  is the one that maximizes  $\varphi_M - \mathcal{J}_{M-1}[\varphi_M]$  in the  $\|\cdot\|_*$  norm. However, since in practice we do not have access to the entire knowledge of this norm,  $\|\cdot\|_*$  is replaced by the ambient norm  $\|\cdot\|_\mathcal{X}$ :

$$\varphi_M = \arg \max_{\varphi \in F} \|\varphi - \mathcal{J}_{M-1}[\varphi]\|_* \sim \arg \max_{\varphi \in F} \|\varphi - \mathcal{J}_{M-1}[\varphi]\|_\mathcal{X}, \quad (28)$$

which is exactly what the Greedy algorithm does (see (4)). Hence, as a conclusion, with the practical tools that can be implemented, the choice of  $\varphi_M$  aims at minimizing the Lebesgue constant with the approximation explained in (28).

### 3. Practical implementation of the Greedy algorithm and the Lebesgue constant

In the present section, we discuss some practical issues regarding the implementation of the Greedy algorithm and the Lebesgue constant  $\Lambda_M$ .

Since the cardinality of  $F$  is usually infinite, the practical implementation of the Greedy algorithm is carried out in a large enough sample subset  $\mathcal{S}_F$  of finite cardinality  $\#\mathcal{S}_F$  much larger than the dimension of the discrete spaces  $X_M$  and  $W_M$  we plan to use. For example, if  $F = \{u(\mu, \cdot), \mu \in \mathcal{D}\}$ , we choose  $\mathcal{S}_F = \{u(\mu, \cdot), \mu \in \Xi_\mu \subset \mathcal{D}\}$  and  $\Xi_\mu$  consists of  $\#\mathcal{S}_F$  parameter sample points  $\mu$ . We assume that this sample subset is representative enough of the entire set  $F$  in the sense that

$$\sup_{x \in F} \left\{ \inf_{y \in \text{span}\{\mathcal{S}_F\}} \|x - y\|_\mathcal{X} \right\}$$

is much smaller than the accuracy we envision through the interpolation process. This assumption is valid for small dimension of  $F$ , or, more precisely, for small dimension of the parameter set  $\mathcal{D}$ . In case it cannot be implemented directly, we can follow two strategies that have been introduced on greedy approaches for reduced basis approximations either based on (parameter) domain decomposition like in [15] or [16] based on an adaptive construction of the sample subset, starting from a very coarse definition as in [17]. These approaches have not been implemented here but we do not foresee any difficulty in adopting them to the GEIM framework.

The following lemma shows that the generalized interpolant can be recursively computed.

**Lemma 3.1.** *For any function  $f \in \mathcal{X}$ , we have the following recursion for  $M \geq 1$*

$$\begin{cases} \mathcal{J}_M[f] = \mathcal{J}_{M-1}[f] + \sigma_M(f - \mathcal{J}_{M-1}[f])q_M \\ \mathcal{J}_0[f] = 0 \end{cases} \quad (29)$$

and the generalized interpolant of  $f$  can be recursively computed.

*Proof.* Using the fact that the spaces  $X_M$  are hierarchically defined, both sides of (29) belong to  $X_M$ . Using the fact that  $\sigma_i(q_M) = 0$  for  $i < M$  and the definition of  $\mathcal{J}_M$  and  $\mathcal{J}_{M-1}$ , we infer that

$$\sigma_i(\mathcal{J}_M[f]) = \sigma_i(\mathcal{J}_{M-1}[f] + \sigma_M(f - \mathcal{J}_{M-1}[f])q_M), \quad \forall i < M.$$

Finally, it is clear that the right and left hand sides have the same image through  $\sigma_M$ . The equality holds by uniqueness of the generalized interpolation procedure.  $\square$

**Remark 3.2.** *This result also holds for the classical EIM case.*

The greedy algorithm is in practice a very time-consuming task whose computing time could significantly be reduced by the use of parallel architectures and the use of formula (29) as is outlined in Algorithm 1.

---

**Algorithm 1** Practical implementation of the Greedy procedure

---

```

1: Input:  $\Sigma$ ,  $\mathcal{S}_F = \{f_k \in F\}_{k=1}^{\#S_F}$ ,  $\varepsilon_{tol}$ ,  $M_{\max}$ ,  $M = 0$ 
2: Assign a set of functions  $\{f_{k_{p,start}}, \dots, f_{k_{p,stop}}\}$  to each processor  $p$ .
3: repeat
4:    $M \leftarrow M + 1$ 
5:    $\varepsilon_{p,\max} = 0$  ▷ parallel
6:   for  $k = \{k_{p,start}, \dots, k_{p,stop}\}$  do
7:      $f = f_k$ 
8:     Compute and store  $\sigma_M(f - \mathcal{J}_M(f))$ .
9:     Assemble  $\mathcal{J}_{M+1}(f)$  following formula (29)
10:    Compute  $\varepsilon_{M+1} = \|f - \mathcal{J}_{M+1}(f)\|_X$ 
11:    if  $\varepsilon_{M+1} > \varepsilon_{p,\max}$  then
12:       $k_{p,\max} = k$  and  $\varepsilon_{p,\max} = \varepsilon_{M+1}$ 
13:    end if
14:  end for ▷ end parallel
15:  Gather  $\{(\varepsilon_{p,\max}, k_{p,\max})\}_{p=1}^{N_{proc}}$  and find  $(\varepsilon_{\max}, k_{\max}) = \arg \max_{p \in \{1, \dots, N_{proc}\}} (\varepsilon_{p,\max}, k_{p,\max})$ .
16:   $r_{M+1} = f_{k_{\max}} - \mathcal{J}_M(f_{k_{\max}})$ 
17:   $\tilde{\varepsilon}_{p,\max} = 0$  ▷ parallel
18:  for  $j = \{j_{p,start}, \dots, j_{p,stop}\}$  do
19:     $\sigma = \sigma_j$ 
20:    Compute  $\tilde{\varepsilon}_{M+1} = |\sigma(r_{M+1})|$ 
21:    if  $\tilde{\varepsilon}_{M+1} > \tilde{\varepsilon}_{p,\max}$  then
22:       $j_{p,\max} = j$  and  $\tilde{\varepsilon}_{p,\max} = \tilde{\varepsilon}_{M+1}$ 
23:    end if
24:  end for ▷ end parallel
25:  Gather  $\{(\tilde{\varepsilon}_{p,\max}, j_{p,\max})\}_{p=1}^{N_{proc}}$  and find  $(\tilde{\varepsilon}_{\max}, j_{\max}) = \arg \max_{p \in \{1, \dots, N_{proc}\}} (\tilde{\varepsilon}_{p,\max}, j_{p,\max})$ .
26:  Compute and store  $q_{M+1} = \frac{r_{M+1}}{\sigma_{j_{\max}}(r_{M+1})}$ .
27:  Store  $\sigma_{M+1} = \sigma_{j_{\max}}$ .
28:  Compute and store  $w_{M+1}$  (Riesz representation of  $\sigma_{M+1}$ ).
29: until  $\varepsilon_{\max} < \varepsilon_{tol}$  or  $M > M_{\max}$ 
30: Output:  $\{\sigma_1, \dots, \sigma_{M+1}\}$ ,  $W_{M+1} = \text{span}\{w_1, \dots, w_{M+1}\}$ ,  $X_{M+1} = \text{span}\{q_1, \dots, q_{M+1}\}$ .

```

---

Once  $X_M$  and  $W_M$  have been constructed thanks to Algorithm 1, the Lebesgue constant can be computed by the resolution of an eigenvalue problem as is explained in

**Lemma 3.3.** *If  $\{\tilde{q}_1, \dots, \tilde{q}_M\}$  and  $\{\tilde{w}_1, \dots, \tilde{w}_M\}$  are an orthonormal basis of  $X_M$  and  $W_M$  respectively, then*

$$\beta_M = 1/\Lambda_M = \sqrt{\lambda_{\min}(A^T A)}, \quad (30)$$

where  $A$  is the  $M \times M$  matrix whose entries are  $A_{i,j} = (\tilde{w}_i, \tilde{q}_j)_X$  and  $\lambda_{\min}(A^T A)$  denotes the minimum eigenvalue of the positive definite matrix  $A^T A$ .

*Proof.* Since

$$\beta_M = \inf_{x \in X_M} \sup_{w \in W_M} \frac{(x, w)_X}{\|x\|_X \|w\|_X} = \inf_{x \in \mathbb{R}^M} \sup_{w \in \mathbb{R}^M} \frac{(Ax, w)_2}{\|x\|_2 \|w\|_2} = \inf_{x \in \mathbb{R}^M} \frac{\|Ax\|_2}{\|x\|_2},$$

the result easily follows because  $\frac{\|Ax\|_2^2}{\|x\|_2^2}$  is the Rayleigh quotient of  $A^T A$  whose infimum is achieved by  $\lambda_{\min}(A^T A)$ .  $\square$

**Remark 3.4.** Note that  $\beta_M$  corresponds to the minimum singular value of the matrix  $A$ , which is a matrix of small size  $M \times M$ . Its computation can be easily performed by, e.g., the inverse power method.

#### 4. A numerical study about the impact of the dictionary $\Sigma$ of linear functionals in the Lebesgue constant

As outlined in Remark 2.5, the explicit expression of the Lebesgue constant presented in formula (16) shows that  $\Lambda_M$  is intimately linked to the dictionary of linear functionals  $\Sigma$  that is used in the Greedy algorithm to build the interpolation process. With the exception of the trivial case considered in Corollary 2.2, no theoretical analysis of the impact of  $\Sigma$  on the behavior of the Lebesgue constant has been possible so far. For this reason, we present here some numerical results on this issue as a first illustration of this connection. The same computations will also let us numerically validate the formula (16) for  $\Lambda_M$ , whose original definition is given by (7).

We place ourselves in  $\Omega = [0, 1]$  and consider the numerical approximation in  $L^2(\Omega)$  or  $H^1(\Omega)$  of the following compact set:

$$F = \{f(\cdot, \mu_1, \mu_2) \mid (\mu_1, \mu_2) \in [0.01, 24.9] \times [0, 15]\}, \quad (31)$$

where

$$f(x, \mu_1, \mu_2) = \frac{1}{\sqrt{1 + (25 + \mu_1 \cos(\mu_2 x))x^2}}, \quad \forall x \in \Omega.$$

We recall that  $L^2(\Omega) = \{f \mid \|f\|_{L^2(\Omega)} < \infty\}$ , where the norm  $\|\cdot\|_{L^2(\Omega)}$  is induced by the inner product  $(w, v)_{L^2(\Omega)} = \int_{\Omega} w(x)v(x)dx$ . Also,  $H^1(\Omega) = \{f \mid \|f\|_{H^1(\Omega)} < \infty\}$ , where the norm  $\|\cdot\|_{H^1(\Omega)}$  is induced by the inner product  $(w, v)_{H^1(\Omega)} = \int_{\Omega} w(x)v(x)dx + \int_{\Omega} \nabla w(x) \cdot \nabla v(x)dx$ .

Any  $f \in F$  will be approximated by its generalized interpolant at dimension  $M$ . For this purpose, the practical construction of the interpolating space  $X_M$  and the selection of the linear functionals is done through the Greedy algorithm described in Section 3. The following dictionary of linear functionals has been employed:

$$\Sigma = \{\sigma_k \in \mathcal{L}(X), k \in \{1, \dots, N_{\text{sensor}}\}\}, \quad (32)$$

where  $N_{\text{sensor}} = 150$ , and

$$\sigma_k(\varphi) = \int_{x \in \Omega} c_{k,s}(x) \varphi(x) dx, \quad \forall \varphi \in X. \quad (33)$$

The function  $c_{k,s}$  reads:

$$c_{k,s}(x) = \frac{m_{k,s}(x)}{\|m_{k,s}(\cdot)\|_{L^1(\Omega)}}, \quad \forall x \in \Omega,$$

where

$$m_{k,s}(x) := e^{-(x-x_k)^2/(2s^2)}, \quad \forall x \in \Omega$$

and  $x_k \in \Omega$ . We will explore the variation of the coefficient  $s \in \mathbb{R}_+$  in order to understand the influence of the dictionary  $\Sigma$  on  $\Lambda_M$ .

#### 4.1. Validation of the inf-sup formula

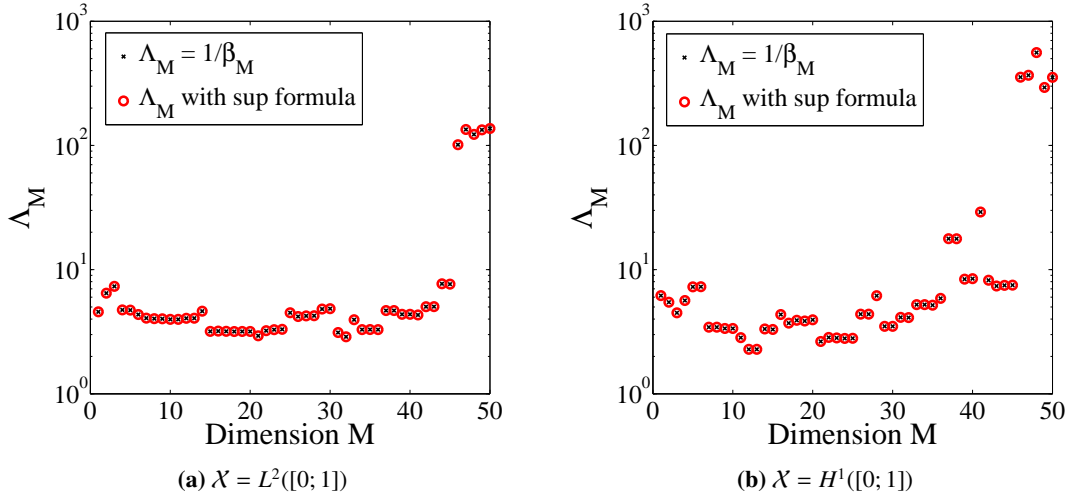
We will first start by fixing  $s$  to a value of 0.005 and by numerically validating formula (16) of the Lebesgue constant by comparing it to the value given by the original formula (7).

Regarding the computation of (16), the quantity  $\beta_M$  has been derived using formula (30) of Lemma 3.3. It suffices to evaluate the scalar products of the matrix  $A$  defined in that lemma and obtain the minimum eigenvalue of  $A^T A$ . For the practical computations, a  $\mathbb{P}_1$  finite element approximation of the functions  $\tilde{q}_i$  and  $\tilde{w}_i$  has been used in order to simplify the scalar product evaluation in the  $L^2$  and  $H^1$  spaces. For the same reason and as a matter of global coherence, the Lebesgue constant

$$\sup_{\varphi \in \mathcal{X}} \frac{\|\mathcal{J}_M[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}}$$

is also approximated in the same  $\mathbb{P}_1$  finite element approximation of the elements of  $\mathcal{X}$ . This approach leads to the computation of a discrete Raleigh quotient, whose derivation is explained in detail in appendix B.

The results of the computation are given in Figure 2 and show an excellent agreement between both values in  $L^2$  and  $H^1$ . The same agreement holds for any value of the parameter  $s$  of the linear functionals, but, as will be presented in the next section, the behavior of  $\Lambda_M$  varies depending on this parameter.



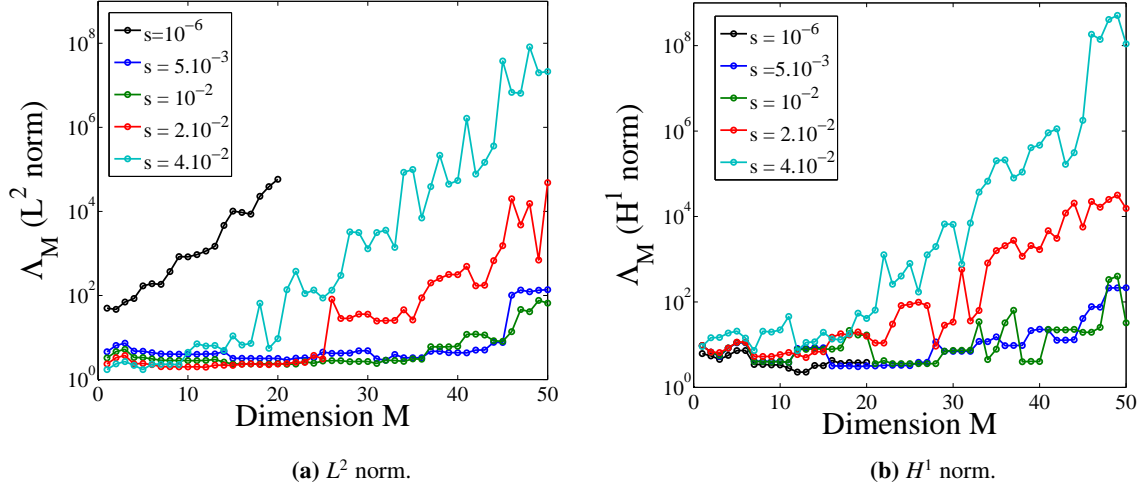
**Figure 2:** Numerical validation of the inf-sup formula: comparison between formulae (7) and (16).

In the particular case presented here, the behavior of the Lebesgue constant does not significantly change if we place ourselves in  $L^2$  or in  $H^1$  and  $\Lambda_M$  remains constant (the degradation in the behavior for  $M \geq 44$  is due to numerical round-off errors).

#### 4.2. Impact of the dictionary of linear functionals

We now study the impact of  $s$  on the evolution of the Lebesgue constant through our example in one dimension. For this purpose, we present in Figures 3a and 3b the behavior in  $L^2$  and in  $H^1$  of  $\Lambda_M$  for different values of  $s$ .

To begin with, we will focus on the behavior for sufficiently large values of  $s$  and analyze the range  $s \geq 5 \cdot 10^{-3}$ . It can be observed that, as  $s$  increases, the Lebesgue constant progressively degrades in both norms. The sequence  $(\Lambda_M)$  starts to diverge at dimensions that are lower and lower as  $s$  increases (compare, e.g., the behaviors between the case  $s = 2 \cdot 10^{-2}$  and  $s = 4 \cdot 10^{-2}$ ). An intuitive manner to interpret this observation is as follows: the dictionary under consideration in this example (see formula (32)) consists on local averages operations whose "range" is controlled by  $s$ . As  $s$  increases, the range increases and a limit will be reached in which the addition of more linear functionals will result in a redundant addition of information because of an overlap of the domains where the local averages are acting. As a result, the larger  $s$ , the sooner this redundancy will appear and the more unstable the process.



**Figure 3:** Impact on  $\Lambda_M$  of the parameter  $s$  of the linear functionals.

It is also important to understand the behavior when the parameter  $s$  tends to zero. In this case, the linear functionals tend to Dirac masses, that, in 1D, are elements of  $H^{-1}$  but not of  $L^2$ . Hence, in the limit  $s = 0$ , the definition of the space  $W_M$  will be possible in  $H^1$  but not in  $L^2$  because the problem:

$$\begin{cases} \text{Find } w_i \in \mathcal{X} \text{ such that:} \\ \sigma_i(\varphi) = (w_i, \varphi)_X = \delta_{x_i}(\varphi), \quad \forall \varphi \in \mathcal{X} \end{cases} \quad (34)$$

is well-defined in  $H^1$  but not in  $L^2$ . This observation helps to understand first of all why  $\Lambda_1$  remains roughly constant in  $H^1$  as  $s$  decreases whereas it behaves as  $s^{-1/2}$  in the  $L^2$  norm (see Figure 4). Indeed, in the  $H^1$  case, we have the inequality

$$\frac{\|\mathcal{J}_1[\varphi]\|_{H^1(\Omega)}}{\|\varphi\|_{H^1(\Omega)}} = |\sigma_1(\varphi)| \frac{\|q_1\|_{H^1(\Omega)}}{\|\varphi\|_{H^1(\Omega)}} \leq \|\varphi\|_{L^\infty(\Omega)} \frac{\|q_1\|_{H^1(\Omega)}}{\|\varphi\|_{H^1(\Omega)}}, \quad \forall \varphi \in H^1(\Omega),$$

which is bounded for any  $s \in \mathbb{R}_+$ . However, in the case of  $L^2(\Omega)$ , it can be inferred that

$$\frac{\|\mathcal{J}_1[\varphi]\|_{L^2(\Omega)}}{\|\varphi\|_{L^2(\Omega)}} = |\sigma_1(\varphi)| \frac{\|q_1\|_{L^2(\Omega)}}{\|\varphi\|_{L^2(\Omega)}} \leq \frac{\|m_{1,s}\|_{L^2(\Omega)}}{\|m_{1,s}\|_{L^1(\Omega)}} \|q_1\|_{L^2(\Omega)}, \quad \forall \varphi \in H^1(\Omega)$$

where we have applied the Cauchy-Schwarz inequality to  $|\sigma_1(\varphi)|$ . A simple change of variable  $u = \frac{x - x_1}{s}$  in the evaluation of  $\frac{\|m_{1,s}\|_{L^2(\Omega)}}{\|m_{1,s}\|_{L^1(\Omega)}}$  leads to the bound

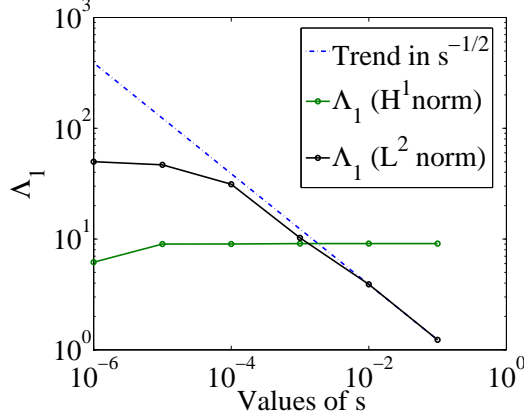
$$\frac{\|\mathcal{J}_1[\varphi]\|_{L^2(\Omega)}}{\|\varphi\|_{L^2(\Omega)}} \leq C \|q_1\|_{L^2(\Omega)} s^{-1/2}, \quad \forall \varphi \in L^2(\Omega), \quad (35)$$

where

$$C = \frac{\int_{\Omega} e^{-u^2} du}{\int_{\Omega} e^{-u^2/2} du}.$$

In Figure 4, note that for values  $s \leq 10^{-4}$ , the behavior of  $\Lambda_1$  no longer follows  $s^{-1/2}$  but this is due to computer limitations. Indeed, the computations have been carried out with a maximum number of  $10^4$  degrees of freedom in

the  $\mathbb{P}_1$  approximation because of memory storage issues. As a result, for  $s \leq 10^{-4}$ , we no longer capture enough information with this finite element precision.



**Figure 4:** Behavior of  $\Lambda_1$  as a function of  $s$  ( $H^1$  and  $L^2$  norms). Remark: the scale of the figure is log-log.

As a consequence of the diverging behavior of  $\Lambda_1$  in  $L^2$  as the parameter  $s$  decreases, it is reasonable to expect that the sequence  $(\Lambda_M)$  quickly diverges as  $s \rightarrow 0$  in  $L^2$  but that it remains bounded in  $H^1$ . This behavior is indeed illustrated in Figures 3a and 3b through the example of  $s = 10^{-6}$ , in which it is possible to observe the phenomenon.

## 5. Application of GEIM to the real-time monitoring of a physical experiment

The main purpose of this section is to illustrate that GEIM can be used as a tool for the real-time monitoring of a physical or industrial process. The rationale is to provide an online accurate approximation of the phenomenon under consideration thanks to the computation of a generalized interpolating function that will be derived on the fly by combining in an appropriate manner measurements from the experiment with mathematical models (a parameter dependent PDE). Such a tool could be employed for keeping track of the process in the whole domain of the experiment and not only at locations where the sensors are placed.

We will also show that the proposed method can be helpful in design of experiments: the method minimizes the number of sensors required to reconstruct the field variable and informs the optimal sensor types and placement.

### 5.1. The general method

Let us assume that we want to monitor in real time a field  $u_{\text{true}}$  appearing as an input for some quantities of interest in a given experiment that involves sensor measurements. We assume that the conditions of the experiment are described by a vector of parameters  $\mu_{\text{true}} \in E$ , where  $E$  is a compact set of  $\mathbb{R}^p$ , and that  $u_{\text{true}}$  is the solution of a parameter dependent PDE

$$D_\mu u = g_\mu \quad \mu \in E \quad (36)$$

when  $\mu = \mu_{\text{true}}$  (in other words  $u_{\text{true}} = u_{\mu_{\text{true}}}$ ). The vector  $\mu_{\text{true}}$  will be unknown in general so the computation of  $u_{\text{true}}$  cannot be done by traditional discretization techniques like finite elements. Besides, even if  $\mu_{\text{true}}$  was known, its computation could not be performed in real-time with classical techniques. For all these reasons, we propose to compute the generalized interpolant  $\mathcal{J}_M[u_{\text{true}}]$  as an approximation of  $u_{\text{true}}$ .

Such an approximation requires that the set of solutions  $\{u_\mu, \forall \mu \in E\}$  is included in some compact set  $F$  of  $X$  that is of small Kolmogorov  $n$ -width in  $X$  ([12]). A dictionary  $\Sigma \subset \mathcal{L}(X)$  is also required. Each element  $\sigma \in \mathcal{L}(X)$  will mathematically model a sensor in the experiment. The location and type of the sensors might be fixed *a priori* (by the experts running the experiment), but it is also possible to consider that  $\Sigma$  is composed by elements that represent potential locations and potential types of sensors.



Since we need to define the generalized interpolating spaces  $X_M = \text{span}\{q_1, \dots, q_M\}$  together with the suitable interpolating linear functionals  $\{\sigma_1, \dots, \sigma_M\}$ , a greedy algorithm has to be performed beforehand and therefore the computation of  $\mathcal{J}_M[u_{\text{true}}]$  is divided into two steps:

- In an *offline phase* (i.e. before the experiment takes place):
  - We define a finite subset  $\mathcal{S}_F = \{u(\mu, \cdot), \mu \in \Xi_\mu \subset E\} \subset F$  and solve (36) for each element of  $\mathcal{S}_F$  with an accurate enough discretization strategy. This can be done with traditional approximation tools like, e.g., finite elements or a reduced basis strategy.
  - Following the steps of Algorithm 1, a greedy algorithm over the set  $\mathcal{S}_F$  is performed to build an  $M$ -dimensional reduced basis  $X_M = \text{span}\{q_j \in F, j \in [1, M]\}$  together with the suitable linear functionals  $\{\sigma_1, \dots, \sigma_M\}$  coming from the dictionary  $\Sigma$ . The selection of the linear functionals means that, among all the sensors in the experiment that constitute our dictionary  $\Sigma$ , we select the  $M$  most suitable according to the greedy criterion. These  $M$  linear functionals/sensors will be the only ones required to reconstruct the field so the present methodology could be viewed as a strategy to minimize the number of sensors involved in the experiment as well as to find their optimal type and placement.
- In an *online phase* (i.e. when the experiment is running), we collect in real time the measurements

$$\{\sigma_1(u_{\mu_{\text{true}}}), \dots, \sigma_M(u_{\mu_{\text{true}}})\}$$

from the  $M$  selected sensors (that are placed in the experiment). The generalized interpolant  $\mathcal{J}_M[u_{\mu_{\text{true}}}]$  can then be computed following formula (3). It has been observed so far (see the numerical example below and [1]) that the interpolation error decreases very quickly as the dimension  $M$  increases and therefore relatively small values of  $M$  are required to reach a good accuracy in the approximation of  $u_{\mu_{\text{true}}}$  by  $\mathcal{J}_M[u_{\mu_{\text{true}}}]$ . Thanks to this, the computation of  $\mathcal{J}_M[u_{\mu_{\text{true}}}]$  can be performed in real-time (or almost).

**Remark 5.1.** Note that our strategy supposes that the physical experiment  $u_{\text{true}}$  is perfectly described by the solution  $u_\mu$  of (36) for  $\mu = \mu_{\text{true}}$ . This is a very strong hypothesis because the model might not perfectly describe the experiment under consideration. Besides, it is here assumed that there is no noise in the measurements, which is also a strong assumption. In [1], some preliminary analysis has been presented to take into account the presence of noise in the measurements. Regarding the model bias, in the recent works of [18, 19], the authors are able to take it into account under several hypothesis in the so called "Parametrized-Background Data-Weak Formulation" for variational data assimilation. In fact, GEIM is a particular instance of this method for the case (with the notations of [19])  $N = M$  and this latter choice is appropriate for situations in which the bias is small.

**Remark 5.2.** In the strategy proposed in this section, sensor measurements are incorporated in the interpolation procedure through the space  $W_M$  (which is spanned by the Riesz representations of the linear functionals of the sensors). In the reference [20], one can find an early work in oceanography in which data assimilation is also incorporated through the construction of the space  $W_M$ . However, in the case of [20], no a priori error analysis was provided in the computational procedure that was proposed.

## 5.2. A numerical application involving the Stokes equation

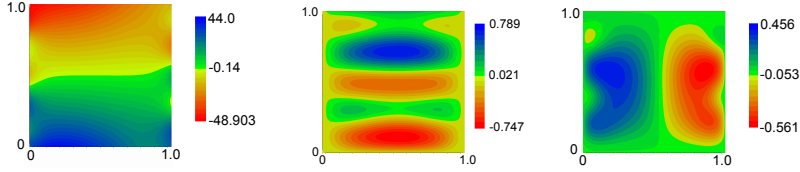
We are going to illustrate the procedure in the case where the experiment corresponds to a lid-driven cavity problem that takes place in the spatial domain  $\Omega = [0; 1] \times [0; 1] \subset \mathbb{R}^2$ . We consider two parameters  $\mu = (\mu_1, \mu_2) \in [1; 8] \times [1; 8]$  such that, for a given  $\mu$ , the parametrized PDE reads

$$\left\{ \begin{array}{l} \text{Find the solution } (\mathbf{u}_\mu, p_\mu) \in (H^1(\Omega))^2 \times L_0^2(\Omega) \text{ of :} \\ -\Delta \mathbf{u}_\mu + \mathbf{grad}(p_\mu) = \mathbf{f}_\mu, \text{ a.e. in } \Omega \\ \text{div}(\mathbf{u}_\mu) = 0, \text{ a.e. in } \Omega \\ \mathbf{u}_\mu = \begin{pmatrix} x(1-x) \\ 0 \end{pmatrix}, \text{ a.e. on } \Gamma_1 \\ \mathbf{u}_\mu = \mathbf{0}, \text{ a.e. on } \partial\Omega \setminus \Gamma_1 \end{array} \right. \quad (37)$$

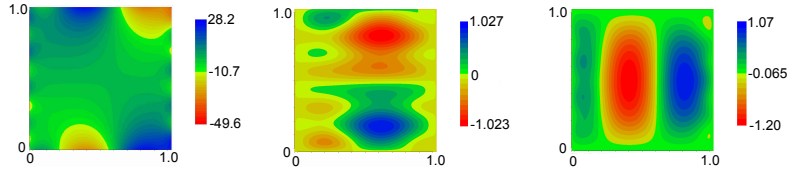
where the forcing term  $\mathbf{f}_\mu = \begin{pmatrix} 100\sin(\mu_1\Pi y) \\ -100\sin(\mu_2\Pi\frac{1-x}{2}) \end{pmatrix}$  and  $\Gamma_1 = \{x \in [0; 1], y = 1\}$ . The space  $L_0^2(\Omega)$  corresponds to the functions in  $L^2(\Omega)$  with zero mean over  $\Omega$ , namely

$$L_0^2(\Omega) = \left\{ f \in L^2(\Omega); \int_{\Omega} f = 0 \right\}.$$

Two examples of solutions are provided on Figures 5 and 6 where finite elements <sup>1</sup>  $\mathbb{P}1$ -bubble/ $\mathbb{P}1$  have been used to approximate the velocity/pressure fields respectively (this approximation is the one that will be used to compute the off-line snapshots that form  $\mathcal{S}_F$ ).



**Figure 5:** From left to right: pressure, horizontal and vertical velocity solutions for the parameter  $\mu = (5; 1)$ .



**Figure 6:** From left to right: pressure, horizontal and vertical velocity solutions for the parameter  $\mu = (8; 5)$ .

We assume that

- the set of solutions  $\{(\mathbf{u}, p)(\mu), \forall \mu\} \subset F$  and  $F$  is of small Kolmogorov  $n$ -width in  $(H^1(\Omega))^2 \times L_0^2(\Omega)$ . This assumption is made *a priori* and will be verified *a posteriori* in a convergence study of the interpolation errors.
- we have velocity and pressure sensors at our disposal which mathematically means that we have:
  - a dictionary for the velocity:  $\Sigma^{\mathbf{u}} = \{\sigma^{\mathbf{u}}\} \subset \mathcal{L}(H^1(\Omega)^2)$
  - a dictionary for the pressure:  $\Sigma^p = \{\sigma^p\} \subset \mathcal{L}(L^2(\Omega))$

In our numerical example, the linear functionals that have been used consist of local averages of the same form as (32) and (33) but adapted to the 2D case. The parameter  $s$  has been fixed to  $s = 10^{-3}$  and we will have  $N_{\text{sensor}} = 100$  sensors for the pressure and other  $N_{\text{sensor}} = 100$  sensors for the velocity. The centers of these local averages are located on a  $10 \times 10$  equispaced grid over  $\Omega$ .

Given an experiment corresponding to the vector of parameters  $\mu_{\text{exp}}$ , we are going to — quickly and accurately— approximate in  $\Omega$  the vectorial field  $(\mathbf{u}, p)(\mu_{\text{exp}})$  by its generalized interpolant  $\mathcal{J}_M[(\mathbf{u}, p)(\mu_{\text{exp}})]$  using the only knowledge of measurements from sensors. Because we are facing here the reconstruction of a vectorial field, several potential input from  $(\mathbf{u}, p)(\mu)$  can be proposed. In the present paper, three different classes will be considered. They will all fulfill the divergence-free condition for the velocity interpolant  $\text{div}(\mathcal{J}_M[\mathbf{u}(\mu)]) = 0$ .

#### **Reconstruction 1: Independent treatment of $u(\mu)$ and $p(\mu)$ .**

The first possibility is to treat  $(\mathbf{u}, p)(\mu)$  not as a vectorial field but as two independent fields  $\mathbf{u}(\mu)$  and  $p(\mu)$  to interpolate

<sup>1</sup>Our computations have been carried out with Freefem++ [21], [22].

independently with velocity measurements for  $\mathbf{u}(\mu)$  and pressure measurements for  $p(\mu)$ . In other words, the generalized interpolant is defined in this case as  $\mathcal{J}_{M_u, M_p}[(\mathbf{u}, p)(\mu)] = (\mathcal{J}_{M_u}^u[\mathbf{u}(\mu)]; \mathcal{J}_{M_p}^p[p(\mu)])$ . This requires the offline execution of two greedy algorithms: one for the velocity and another for the pressure. Each one respectively provides

- a velocity basis  $\{\mathbf{u}(\mu_i)\}_{i=1}^{M_u}$  and a set of  $M_u$  velocity sensors  $\{\sigma_i^u\}_{i=1}^{M_u}$  chosen among the dictionary  $\Sigma^u$ . The interpolant for the velocity will be  $\mathcal{J}_{M_u}^u[\mathbf{u}(\mu)] = \sum_{i=1}^{M_u} \alpha_i \mathbf{u}(\mu_i)$  where the  $\alpha_i$  are given by the interpolating conditions  $\sigma_i^u(\mathcal{J}_{M_u}^u[\mathbf{u}(\mu)]) = \sigma_i^u(\mathbf{u}(\mu))$ ,  $\forall i \in \{1, \dots, M_u\}$ .
- a pressure basis  $\{p(\mu_j)\}_{j=1}^{M_p}$  and a set of pressure sensors  $\{\sigma_j^p\}_{j=1}^{M_p}$  chosen among  $\Sigma^p$ . The interpolant for the pressure will be  $\mathcal{J}_M^p[p(\mu)] = \sum_{j=1}^{M_p} \gamma_j p(\mu_j)$  where the  $\gamma_j$  are given by the interpolating conditions  $\sigma_j^p(\mathcal{J}_M^p[p(\mu)]) = \sigma_j^p(p(\mu))$ ,  $\forall j \in \{1, \dots, M_p\}$ .

Note that in this approximation, the construction of  $\mathcal{J}_{M_u, M_p}[(\mathbf{u}, p)(\mu)]$  involves  $M_p$  pressure sensors and  $M_u$  velocity sensors, i.e.  $M_p + M_u$  coefficients. In Figure 7, we have represented the locations of the sensors in the order given by the greedy algorithm.

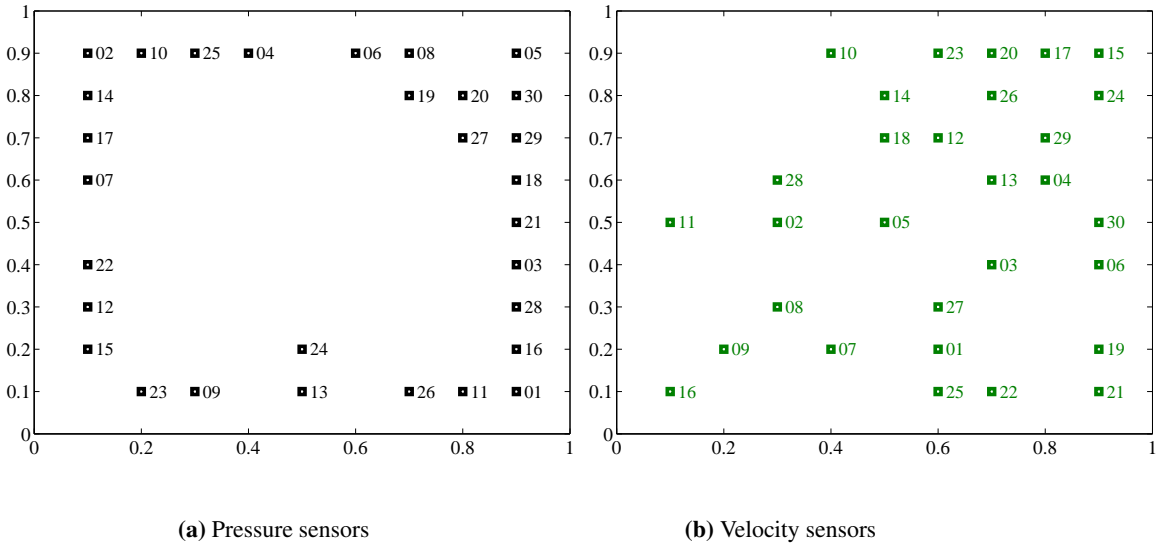


Figure 7: Locations of the sensors for reconstruction 1.

The performances of the method are plotted in Figure 8 where a numerical estimation of the behavior of the interpolating errors for the reconstruction of  $u$  and  $p$  have been represented. These values have been obtained by the interpolation of 196 configurations coming from different parameter values  $\mu_i$  following formula:

$$\begin{cases} e_{M_p}^p = \max_{i \in \{1, \dots, 196\}} \frac{\|p(\mu_i) - \mathcal{J}_{M_p}^p[p(\mu_i)]\|_{L^2(\Omega)}}{\|p(\mu_i)\|_{L^2(\Omega)}} \\ e_{M_u}^u = \max_{i \in \{1, \dots, 196\}} \frac{\|\mathbf{u}(\mu_i) - \mathcal{J}_{M_u}^u[\mathbf{u}(\mu_i)]\|_{H^1(\Omega)^2}}{\|\mathbf{u}(\mu_i)\|_{H^1(\Omega)^2}}. \end{cases} \quad (38)$$

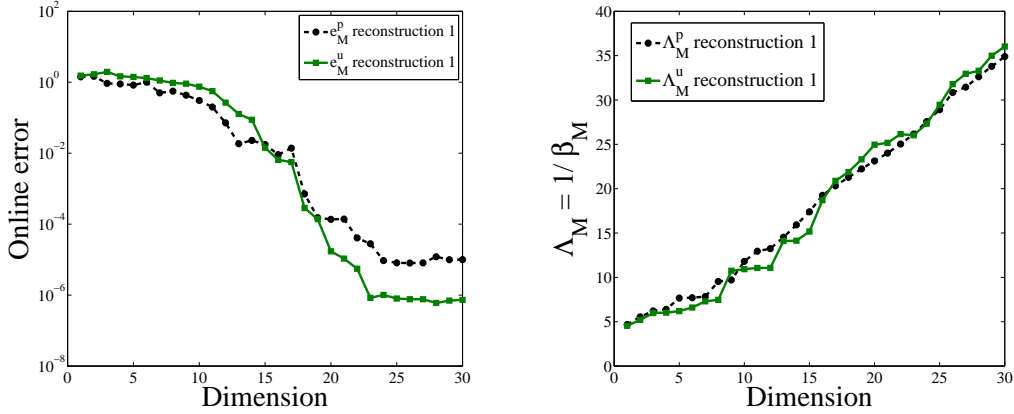
In this figure, we can observe the convergence of the interpolation errors for both the velocity and pressure fields. After a preasymptotic stage for interpolating spaces of small dimension, an exponential convergence of the error is observed. After about dimension  $M = 25$ , the error stagnates due to the fact that we have reached the finite element

accuracy used for the computation of the offline snapshots. The computation of the Lebesgue constants

$$\Lambda_{M_p}^p := \sup_{p \in L^2(\Omega)} \frac{\|\mathcal{J}_{M_p}^p(p)\|_{L^2(\Omega)}}{\|p\|_{L^2(\Omega)}}, \quad \Lambda_{M_u}^u = \sup_{u \in H^1(\Omega)^2} \frac{\|\mathcal{J}_{M_u}^u(u)\|_{H^1(\Omega)^2}}{\|u\|_{H^1(\Omega)^2}} \quad (39)$$

has also been performed following formula (16). Its behavior seems linear with the dimension of interpolation and is therefore far from the crude theoretical upper bound given in formula (8). From the results presented in Section 4, an idea to improve the behavior of the Lebesgue constant could be to consider a smaller value for  $s$ . However, this does not seem necessary here given the good convergence properties observed. Furthermore, in a real case, the parameter  $s$  of the linear functionals is difficult to change as it is fixed by the filter characteristics of the sensors involved in the experiment.

**Remark 5.3.** *In the case of a quickly diverging Lebesgue constant, the convergence properties of the method could be severely degraded (see Theorem 1.4). As Section 4 illustrates, one of the reasons for the divergence could be related to a lack filter-width of the sensors (parameter  $s$  too large) and, in this case, it would be necessary to restart the procedure with sensors with a smaller filter-width.*



**Figure 8:** Reconstruction 1: A numerical estimation of the behavior of the interpolation error (left) and the Lebesgue constant (right) as a function of the dimension of the interpolating spaces  $X_M$

### Reconstructions 2 and 3: Vectorial treatment for $u(\mu)$ and $p(\mu)$ .

An alternative to the first reconstruction is to consider  $(u, p)(\mu)$  as a vectorial field and define its generalized interpolant as  $\mathcal{J}_M[(u, p)(\mu)] := \sum_{i=1}^M \gamma_i (u, p)(\mu_i)$ , where now only  $M$  coefficients  $\gamma_i$  are involved. The joint basis  $\{(u, p)(\mu_i)\}_{i=1}^M$  is provided by a greedy algorithm in the offline stage together with a set of  $M$  linear functionals  $\{\sigma_i^{(u,p)}\}_{i=1}^M$ . Each of these linear functionals involve pressure and velocity measurements at a given spatial location and are defined as  $\sigma_i^{(u,p)} := \sigma_i^u(u) + \sigma_i^p(p)$ . The interpolating conditions for the inference of the coefficients  $\gamma_i$  are now the following:

$$\sigma_i^{(u,p)}((u, p)(\mu)) = \sigma_i^{(u,p)}(\mathcal{J}_M[(u, p)(\mu)]) = \sum_{j=1}^M \gamma_j \sigma_i^{(u,p)}((u, p)(\mu_j)), \quad \forall i \in \{1, \dots, M\}, \quad (40)$$

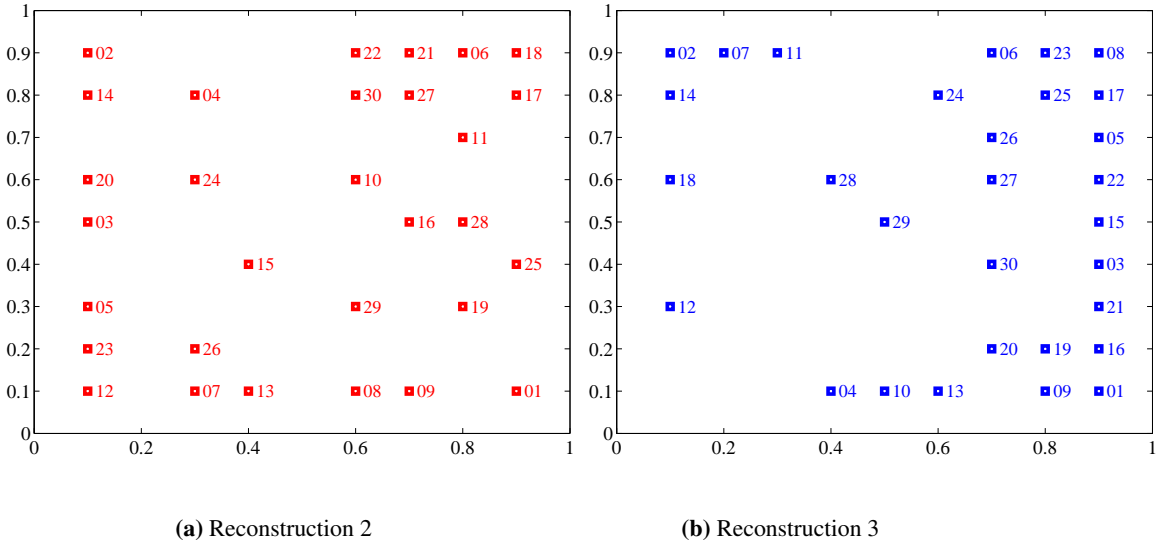
Notice that this definition of the linear functionals  $\sigma^{(u,p)}$  can involve both velocity and pressure measurements or can take into account velocity or pressure measurements only by setting  $\sigma^u = 0$  or  $\sigma^p = 0$ . We have explored this flexibility in the following two reconstructions:

- the interpolation of the pressure and velocity fields with pressure and velocity measurements:  $\sigma_i^{(u,p)} := \sigma_i^u(u) + \sigma_i^p(p)$  (reconstruction 2).

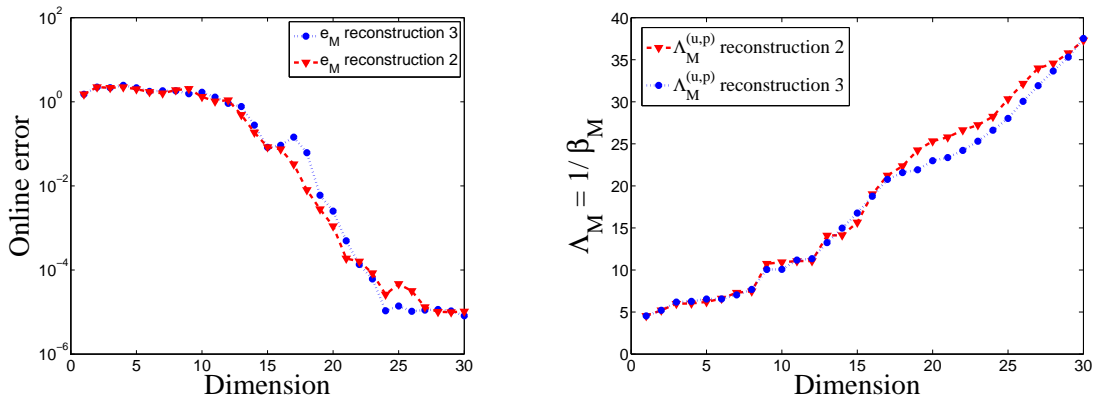
- the interpolation of the pressure and velocity fields with pressure measurements only:  $\sigma_i^{(\mathbf{u},p)} := \sigma_i^p(p)$ . In other words, we are here studying if a velocity field can efficiently be reconstructed with the only knowledge of pressure measurements (reconstruction 3).

The sensor locations provided by the greedy algorithm are shown in Figure 9 and the results are summarized in Figures 10 where an estimation of the interpolation error is plotted according to formula (41).

$$e_M^{(\mathbf{u},p)} = \max_{i \in \{1, \dots, 196\}} \frac{\|(\mathbf{u}, p)(\mu_i) - \mathcal{J}_M^{(\mathbf{u},p)}[(\mathbf{u}, p)(\mu_i)]\|_{H^1(\Omega)^2 \times L^2(\Omega)}}{\|(\mathbf{u}, p)(\mu_i)\|_{H^1(\Omega)^2 \times L^2(\Omega)}}. \quad (41)$$



**Figure 9:** Locations of the sensors for reconstructions 2 and 3.



**Figure 10:** Reconstructions 2 and 3: A numerical estimation of the behavior of the interpolation error (left) and the Lebesgue constant (right) as a function of the dimension of the interpolating spaces  $X_M$ .

The interpolating error of the two types of reconstructions presents a very similar decay behavior in both cases and the convergence is also very similar to reconstruction 1. The most interesting consequence of this is that the velocity can efficiently be reconstructed with only pressure measurements. This result cannot probably be generalized to all types of situations but it proves that in some cases like the current one there is some redundancy in the data and that, in this precise problem, there is no need in having velocity measurements in order to obtain a good accuracy in the approximation of the velocity field.

The Lebesgue constant

$$\Lambda_M^{(u,p)} = \sup_{(u,p) \in H^1(\Omega)^2 \times L^2(\Omega)} \frac{\|\mathcal{J}_M^{(u,p)}(\mathbf{u}, p)\|_{H^1(\Omega)^2 \times L^2(\Omega)}}{\|(\mathbf{u}, p)\|_{H^1(\Omega)^2 \times L^2(\Omega)}} \quad (42)$$

has also been computed for reconstructions 2 and 3 as is shown in Figure 10. Once again, the behavior is linear which is a moderate growth rate.

## 6. Conclusion and perspectives

After revisiting the foundations of GEIM for Banach spaces, the present work has focused on understanding the stability of the process and a relation between  $\Lambda_M$  and an inf-sup problem has been established in the particular case of Hilbert spaces. An interpretation of the generalized interpolant as an oblique projection has also been presented in that case. The derived formula for  $\Lambda_M$  has also allowed us to notice that the Greedy algorithm optimizes in some sense the Lebesgue constant.

A first analysis about the impact of the dictionary of linear functionals  $\Sigma$  on the Lebesgue constant has also been presented through a numerical test case. Furthermore, for a given dictionary  $\Sigma$ , the Lebesgue constant depends on the norm of the ambient space  $\mathcal{X}$  (see formula (7)). A comparison of the behavior of  $(\Lambda_M)$  when  $\mathcal{X} = L^2$  or  $H^1$  has been provided in the case of a dictionary composed of simple local averages.

Beyond these results, there are still many challenging theoretical open questions. Among the most important we mention:

- Can we obtain a general theory for the impact of  $\Sigma$  on the behavior of  $(\Lambda_M)$ ? Can we obtain a tighter upper bound than (8)?
- How can we carry out the offline computation in a reasonable time when the number of parameters is large?
- How can we treat the model bias between  $u_{\text{true}}$  and the manifold of solutions of our parameter dependent PDE? The works of [18] will probably be helpful to carry out this task.
- How can we treat noisy measurements? One can find some preliminary ideas in [1] and the works of [23].

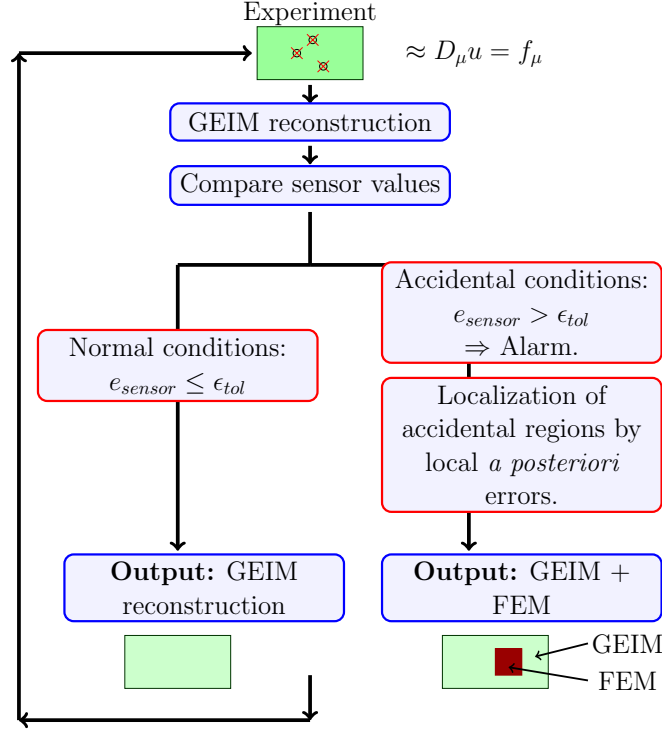
Furthermore, the recent results of [12] lead us to think that it would be interesting to explore non-linear inputs of the form

$$\sigma(t(\varphi)),$$

where  $\sigma \in \mathcal{L}(\mathcal{X})$ ,  $t : \mathcal{X} \rightarrow \mathcal{X}$  is a non-linear mapping and  $\varphi$  is an element of a compact set of small Kolmogorov  $n$ -width in  $\mathcal{X}$ . In an ongoing work, we are exploring this idea in the case of the Navier-Stokes equations.

In the second part of the paper, we have illustrated one of the most straightforward practical applications of GEIM that consists of monitoring in real-time a process. The idea is that GEIM could reconstruct in real-time physical quantities in the whole domain of an experiment by combining the real-time acquisition of measurements from sensors with mathematical models (parameter dependent PDEs).

This scheme has been applied to an example dealing with a parametrized lid-driven Stokes equation. The example shows a fast decrease in the interpolation error, which confirms that it is feasible to use GEIM to monitor experiments in real-time in cases where  $d_n(F, \mathcal{X})$  is small enough (i.e. when the experiment is simple enough). The behavior of the Lebesgue constant seems to be linear and seems to be in accordance with previous works for the classical EIM (see [8]). The linear increase is far from the theoretical exponential upper bound of (8) and suggests that the bound might not be optimal in sets  $F$  of small Kolmogorov  $n$ -width. In the example, two types of sensors have been used (of



**Figure 11:** A tool to supervise in real-time the safety of an experiment.

pressure and velocity) and the idea of introducing different classes of sensors could be extended to make more refined distinctions between them.

By taking this method as a starting point, GEIM could be used to devise a more complete tool capable of supervising the safety of processes (see Figure 11). The idea would be the following: suppose we have  $2M$  measurements at our disposal. We first invoke GEIM to reconstruct the solution field using  $M$  selected measurements. We next predict, based on the GEIM-reconstructed field, the output associated with the remaining  $M$  sensors. We then compare the  $M$  predicted values with the actual experimental measurements. If the values differ too much from each other, then we consider that an abnormal event has occurred in the experiment and an alarm can be sounded to signal the incident.

In addition to this alarm information, we can seek to provide an accurate enough reconstruction of the solution during the incident by using the following strategy: through the computation of an *a posteriori* error estimator in the regions where the sensor measurements are not in accordance, we could imagine to localize the spatial region(s) where the reconstruction is no longer accurate. The domain could then be split into two parts:

- a subdomain with small Kolmogorov  $n$ -width where the reconstruction by GEIM is still accurate enough.
- a subdomain with big Kolmogorov  $n$ -width where the accident is located and GEIM is no longer accurate. The domain is computed by traditional discretization techniques such as finite elements complemented with Dirichlet boundary conditions from the GEIM reconstruction.

Under the hypothesis that the subdomain affected by the accident is small, the reconstruction could still be performed relatively quickly, preserving the real-time aspect of our device. The feasibility of decomposing the domain and coupling GEIM with other approximations has been explored in [1] in a simple Laplace problem.

Last but not least, it would also be interesting to explore the robustness of the method in cases where one or several sensors involved in the GEIM reconstruction fail either by being mute or providing a wrong answer.

## Appendix A.

**Proposition Appendix A.1.** *Let  $\mathcal{X}$  be a Hilbert space and  $E, F$  two subspaces of  $\mathcal{X}$ . Then,  $\beta_{E,F} = \beta_{F^\perp, E^\perp}$ , where:*

$$\beta_{E,F} \equiv \inf_{\substack{e \in E \\ \|e\|=1}} \sup_{\substack{f \in F \\ \|f\|=1}} (e, f) \quad (\text{A.1})$$

$$\beta_{F^\perp, E^\perp} \equiv \inf_{\substack{f \in F^\perp \\ \|f\|=1}} \sup_{\substack{e \in E^\perp \\ \|e\|=1}} (e, f). \quad (\text{A.2})$$

*Proof.* Given  $e \in \mathcal{X}$  of norm unity, we introduce  $f_e^*$  as

$$f_e^* = \arg \sup_{\substack{g \in F \\ \|g\|=1}} (e, g).$$

We can then show from optimality that  $(e, h) = 0$  for all  $h$  in  $\{q \in F \mid (q, f_e^*) = 0\}$  and hence

$$e = \lambda f_e^* + \varepsilon \quad (\text{A.3})$$

for some  $\lambda \in \mathcal{R}$  and  $\varepsilon \in F^\perp$  such that  $\lambda^2 + \|\varepsilon\|^2 = 1$  (from our normalization and orthogonality). We then deduce from (A.3), orthogonality, and Cauchy-Schwarz that

$$\sup_{\substack{p \in F \\ \|p\|=1}} (e, p) = \lambda$$

and

$$\sup_{\substack{p \in F^\perp \\ \|p\|=1}} (e, p) = \|\varepsilon\|.$$

Hence,

$$\left( \sup_{\substack{p \in F \\ \|p\|=1}} (e, p) \right)^2 + \left( \sup_{\substack{p \in F^\perp \\ \|p\|=1}} (e, p) \right)^2 = 1 \quad (\text{A.4})$$

thanks to our normalization.

We may now note from (A.1) and (A.4) that

$$\begin{aligned} \beta_{E,F} &= \inf_{\substack{e \in E \\ \|e\|=1}} \sqrt{1 - \left( \sup_{\substack{p \in F^\perp \\ \|p\|=1}} (e, p) \right)^2} \\ &= \sqrt{1 - \left( \sup_{\substack{e \in E \\ \|e\|=1}} \sup_{\substack{p \in F^\perp \\ \|p\|=1}} (e, p) \right)^2} \\ &= \sqrt{1 - \left( \sup_{\substack{p \in F^\perp \\ \|p\|=1}} \sup_{\substack{e \in E \\ \|e\|=1}} (e, p) \right)^2} \end{aligned} \quad (\text{A.5})$$

as we can exchange the two supremizer operations.

Finally, we define a second inf-sup constant,

$$\beta_{F^\perp, E^\perp} \equiv \inf_{\substack{f \in F^\perp \\ \|f\|=1}} \sup_{\substack{e \in E^\perp \\ \|e\|=1}} (e, f). \quad (\text{A.6})$$

We can repeat the procedure above —  $E$  goes to  $F^\perp$  and  $F$  goes to  $E^\perp$  — to find

$$\beta_{F^\perp, E^\perp} = \sqrt{1 - \left( \sup_{\substack{p \in F^\perp \\ \|p\|=1}} \sup_{\substack{e \in (E^\perp)^\perp \\ \|e\|=1}} (e, p) \right)^2}, \quad (\text{A.7})$$

and hence conclude from (A.5) and (A.7) that

$$\beta_{E,F} = \beta_{F^\perp, E^\perp}$$

since  $(E^\perp)^\perp = E$ . □



## Appendix B.

We propose here a practical method for the computation of

$$\sup_{\varphi \in \mathcal{X}} \frac{\|\mathcal{J}_M[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}}. \quad (\text{B.1})$$

The strategy consists in using a finite element Galerkin projection as an approximation of the elements of  $\mathcal{X}$ . We therefore propose to compute

$$\max_{\varphi \in V_h^k} \frac{\|\mathcal{J}_M[\varphi]\|_{V_h^k}}{\|\varphi\|_{V_h^k}}$$

as a surrogate of (B.1), where  $V_h^k$  is the classical continuous finite element approximating space of mesh size  $h$  that involves piece-wise  $\mathbb{P}_k$  polynomials. Let  $\mathcal{B} = \text{span}\{b_1, \dots, b_N\}$  be a basis of  $V_h^k$  and let  $M$  be the  $N \times N$  mass matrix of entries  $M_{i,j} = (b_i, b_j)_{\mathcal{X}}$ ,  $1 \leq i, j \leq N$ . For any  $\varphi \in V_h^k$ , let

$$\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_N)^T \quad (\text{B.2})$$

be the vector of coordinates of  $\varphi$  in the basis  $\mathcal{B}$ . In coherence with these notations, for any  $1 \leq i \leq M$ , the vectors

$$\mathbf{q}_i = (q_{1,i}, \dots, q_{N,i})^T \quad \text{and} \quad \mathbf{w}_i = (w_{1,i}, \dots, w_{N,i})^T \quad (\text{B.3})$$

will respectively denote the Galerkin projections onto  $V_h^k$  of the interpolating basis functions  $q_i \in \mathcal{X}$  and of the Riesz representation of the  $i$ -th linear functional,  $\sigma_i$ . Furthermore, let  $Q^M$  be the  $N \times M$  matrix such that

$$Q^M = [\mathbf{q}_1, \dots, \mathbf{q}_M],$$

and let  $C^M$  be the  $M \times N$  matrix such that:

$$C_{i,j}^M = \sigma_i(b_j) = (w_i, b_j)_{\mathcal{X}}, \quad \forall 1 \leq i \leq M, 1 \leq j \leq N.$$

Finally, we recall that  $B^M$  is the  $M \times M$  matrix defined in Section 1 whose entries are

$$B_{i,j}^M = \sigma_i(q_j) = (w_i, q_j)_{\mathcal{X}}, \quad \forall 1 \leq i \leq M, 1 \leq j \leq M.$$

An approximation of the entries of  $B^M$  and  $C^M$  can easily be computed by using the finite element Galerkin projections of the involved functions:

$$\begin{cases} C_{i,j}^M \approx \mathbf{w}_i^T M \mathbf{b}_j, & \forall 1 \leq i \leq M, 1 \leq j \leq N \\ B_{i,j}^M \approx \mathbf{w}_i^T M \mathbf{q}_j, & \forall 1 \leq i \leq M, 1 \leq j \leq M. \end{cases}$$

With these notations, we can easily prove

**Lemma Appendix B.1.** *Let  $T$  be the  $N \times N$  symmetric positive definite matrix:*

$$T := (Q^M (B^M)^{-1} C^M)^T M (Q^M (B^M)^{-1} C^M),$$

and let  $\lambda_{\max}(T)$  be the largest eigenvalue of the generalized eigenvalue problem

$$\begin{cases} \text{Find } (\lambda, x) \in \mathbb{R} \times \mathbb{R}^N \text{ such that:} \\ Tx = \lambda Mx. \end{cases} \quad (\text{B.4})$$

Then:

$$\max_{\varphi \in V_h^k} \frac{\|\mathcal{J}_M[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}} = \sqrt{\lambda_{\max}}. \quad (\text{B.5})$$

*Proof.* For any  $\varphi \in V_h^k$  and any  $1 \leq i \leq M$ :

$$\sigma_i(\varphi) = \sum_{j=1}^N \varphi_j \sigma_i(b_j) = \mathbf{e}_i^T C^M \boldsymbol{\varphi},$$

where  $\mathbf{e}_i$  is the  $i$ -th canonical vector of dimension  $M$ . Furthermore, if

$$\mathcal{J}_M[\varphi] = \sum_{i=1}^M \alpha_i^M(\varphi) q_i \quad (\text{B.6})$$

is the generalized interpolant of  $\varphi$  in dimension  $M$ , we have:

$$\sigma_i(\mathcal{J}_M[\varphi]) = \mathbf{e}_i^T B^M \boldsymbol{\alpha}, \quad \forall 1 \leq i \leq M,$$

where  $\boldsymbol{\alpha} = (\alpha_1^M(\varphi), \dots, \alpha_M^M(\varphi))^T$ . From the interpolation property stated in (3), it follows that

$$\boldsymbol{\alpha} = (B^M)^{-1} C^M \boldsymbol{\varphi}.$$

Then, the finite element Galerkin projection of the interpolant of (B.6) can be expressed as:

$$\mathcal{J}_M[\varphi] \approx Q^M \boldsymbol{\alpha} = Q^M (B^M)^{-1} C^M \boldsymbol{\varphi}.$$

Hence,

$$\begin{aligned} \max_{\varphi \in V_h^k} \frac{\|\mathcal{J}_M[\varphi]\|_X}{\|\varphi\|_X} &= \left( \max_{\varphi \in \mathbb{R}^N} \frac{\boldsymbol{\varphi}^T (Q^M (B^M)^{-1} C^M)^T M (Q^M (B^M)^{-1} C^M) \boldsymbol{\varphi}}{\boldsymbol{\varphi}^T M \boldsymbol{\varphi}} \right)^{1/2} \\ &= \sqrt{\lambda_{\max}(T)}. \end{aligned}$$

□

**Remark Appendix B.2.** The computation of  $\Lambda_{\max}$  can easily be performed by, e.g., the power method scheme applied to the matrix  $T$ . However, note that the evaluation of  $\Lambda_M$  with formula (B.5) requires the construction of  $T$ , which is a large dense matrix of dimension  $N \times N$ . In cases where the storage of  $T$  is no longer possible, the Lebesgue constant can still be computed with formula (30), whose evaluation requires the construction of a much smaller matrix of dimension  $M \times M$ .

## Acknowledgments

This work was supported in part by the joint research program MANON between CEA-Saclay and University Pierre et Marie Curie-Paris 6. This work was also supported by Fondation Sciences Mathématiques de Paris, OSD/AFOSR/MURI Grant FA9550-09-1-0613, and ONR Grant N00014-11-1-0713.

We would like to thank Prof. F. Hecht for very valuable advices regarding our numerical simulations with FreeFem++ [21, 22].

- [1] Y. Maday, O. Mula, A generalized empirical interpolation method: Application of reduced basis techniques to data assimilation, in: F. Brezzi, P. Colli Franzone, U. Gianazza, G. Gilardi (Eds.), Analysis and Numerics of Partial Differential Equations, Vol. 4 of Springer INdAM Series, Springer Milan, 2013, pp. 221–235.
- [2] M. Barrault, Y. Maday, N. Nguyen, A. Patera, An empirical interpolation method: Application to efficient reduced-basis discretization of partial differential equations., C. R. Acad. Sci. Paris, Série I. 339 (2004) 667–672.
- [3] A. Chkifa, A. Cohen, C. Schwab, High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDE's, Foundations of Computational Mathematics (2013) 1–33.

- [4] I. H. Sloan, Nonpolynomial interpolation, *J. Approx. Theory* 39 (2) (1983) 97 – 117.
- [5] M. Floater, K. Hormann, Barycentric rational interpolation with no poles and high rates of approximation, *Numer. Math.* 107 (2) (2007) 315–331.
- [6] J. Kleijnen, W. van Beers, Robustness of kriging when interpolating in random simulation with heterogeneous variances: Some experiments, *European Journal of Operational Research* 165 (3) (2005) 826 – 834.
- [7] M. Grepl, Y. Maday, N. Nguyen, A. Patera, Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations., *ESAIM, Math. Model. Numer. Anal.* 41(3) (2007) 575–605.
- [8] Y. Maday, N. Nguyen, A. Patera, G. Pau, A general multipurpose interpolation procedure: the magic points, *Comm. Pure Appl. Anal.* 8(1) (2009) 383–404.
- [9] M. Drogman, B. Haasdonk, M. Ohlberger, Reduced basis approximation for non-linear parametrized evolution equations based on Empirical Operator Interpolation, *SIAM J. Sci. Comput.* 34 (2012) A937–A969.
- [10] A. Kolmogoroff, Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse, *Annals of Mathematics* 37 (1936) 107–110.
- [11] Y. Maday, A. Patera, G. Turinici, A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations, *Journal of Scientific Computing* 17 (1-4) (2002) 437–446.
- [12] A. Cohen, R. DeVore, Kolmogorov widths under holomorphic mappings, Submitted.
- [13] Y. Maday, O. Mula, G. Turinici, A priori convergence of the Generalized Empirical Interpolation Method, in: 10th international conference on Sampling Theory and Applications (SampTA 2013), 2013, pp. 168–171.
- [14] Y. Maday, O. Mula, G. Turinici, Convergence analysis of the Generalized Empirical Interpolation Method, Submitted – (2014) –.
- [15] J. Eftang, A. Patera, E. Rønquist, An” hp” certified reduced basis method for parametrized elliptic partial differential equations, *SIAM J. Sci. Comput.* 32 (6) (2010) 3170–3200.
- [16] J. Eftang, B. Stamm, Parameter multi-domain ”hp” empirical interpolation, *Int. J. Numer. Methods Eng.* 90 (4) (2012) 412–428.
- [17] Y. Maday, B. Stamm, Locally adaptive greedy approximations for anisotropic parameter reduced basis spaces., *SIAM J. Sci. Comput.* 35 (6).
- [18] M. Yano, J. Penn, A. Patera, A model-data weak formulation for simultaneous estimation of state and model bias, *Comptes Rendus Mathématique* 351 (2324) (2013) 937 – 941.
- [19] Y. Maday, A. T. Patera, J. D. Penn, M. Yano, A parameterized-background data-weak approach to variational data assimilation: formulation, analysis, and application to acoustics, *Int. J. Num. Meth. Eng.* [doi:10.1002/nme.4747](https://doi.org/10.1002/nme.4747).
- [20] A. Bennett, Array design by inverse methods, Vol. 15, 1985.
- [21] F. Hecht, New developments in Freefem++, *J. Numer. Math.* 20 (3-4) (2012) 251–265.
- [22] <http://www.freefem.org/ff++/>.
- [23] A. Patera, E. Rønquist, Regression on parametric manifolds: Estimation of spatial fields, functional outputs, and parameters from noisy data, *C.R. Acad. Sci. Paris Series I* 350 (9-10) (2012) 543–547.