Convergence analysis of the Generalized Empirical Interpolation Method

Yvon Maday, Olga Mula, Gabriel Turinici

CONVERGENCE ANALYSIS OF THE GENERALIZED EMPIRICAL INTERPOLATION METHOD *

Y. MADAY[†‡§‖], O. MULA[†¶‖], AND G. TURINICI[**‡]

Abstract. The Generalized Empirical Interpolation Method (GEIM, [12]) is an extension first presented in [12] of the classical Empirical Interpolation Method (see [1], [8], [15]). It replaces values at interpolation points by evaluations from continuous linear forms, which allows, in particular, to relax the classical continuity constraint in the functions to interpolate. These functions are members of a compact subset F of a Banach or Hilbert space with a small Kolmogorov n-width and the quality of the approximation strongly depends on the choice of the interpolating functions and linear forms. For this reason, the purpose of this work is to provide *a priori* convergence rates for the GEIM that proposes a greedy algorithm to choose these interpolation couples. We show that, when the Kolmogorov n-width of F decays polynomially or exponentially, the interpolation error has the same behavior modulo the norm of the interpolation operator of GEIM. Sharper results will also be obtained in the situation when the ambient space is a Hilbert.

Key words. interpolation; empirical interpolation; generalized empirical interpolation; convergence rates; reduced basis; reduced order model;

AMS subject classifications.

1. Introduction.

Let $\mathcal{X}$ be a Banach space of functions defined over a domain $\overline{\Omega} \in \mathbb{R}^d$ or $\mathbb{C}^d$, $X_n \in \mathcal{X}$ be a sequence of finite n-dimensional spaces and $S_n = \{x_1, \dots, x_n\}$ be a set of n points in $\overline{\Omega}$. The problem of interpolating any function $f \in \mathcal{X}$ has traditionally been stated as:

$$\text{“Find } f_n \in X_n \text{ such that } f_n(x_i) = f(x_i),\ \forall i \in \{1, \dots, n\}\text{”,} \tag{1.1}$$

where we note that it is implicitly assumed that $\mathcal{X}$ is a Banach space of continuous functions. Given X_n and S_n, among the most important issues raised by interpolation stand questions of existence and uniqueness of the interpolant of any $f \in \mathcal{X}$ and also about the stability of the process (via the study of the behavior of the Lebesgue constant – see [5] for this notion–). This, in turn, leads to an even more fundamental question related to the optimal choice of the interpolating space X_n together with the set of points S_n that provide the best interpolation properties. The difficulty of the task has usually led to restrict the study to lagrangian type approximations where the interpolating space X_n is spanned by algebraic polynomials, rational functions, Fourier series, etc. This approach is rather well documented and understood, especially in the case of polynomial interpolation where we know that, in one dimension, an almost optimal location for the interpolating points is given by the Gauss-Chebyshev nodes. More involved conditions are also known in higher dimensions in order for a polynomial interpolation to be well defined and we refer to [5] for more details on this topic.

*This work was supported in part by the joint research program MANON between CEA-Saclay and University Pierre et Marie Curie-Paris 6.

†Sorbonne Universités, UPMC Univ Paris 06 and CNRS UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France (maday@ann.jussieu.fr ; mula-hernandez@ann.jussieu.fr)

‡Institut Universitaire de France

§Division of Applied Mathematics, Brown University, Providence RI, USA.

¶CEA Saclay, DEN/DANS/DM2S/SERMA/LLPR, 91191 Gif-Sur-Yvette CEDEX - France

‖LRC MANON – CEA/DEN/DANS/DM2S and UPMC-CNRS/LJLL.

**CEREMADE, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris, France (gabriel.turinici@dauphine.fr)

Although the extension of the Lagrangian interpolation has already been explored in the literature (see, e.g. [17], [7] and also the activity concerning the kriging [9], [11] in the stochastic community), the question still remains on how to extend the concept of interpolation stated in (1.1) to general functions. One step in this direction is the Empirical Interpolation Method (EIM, [1], [8], [15]) that aims at interpolating continous functions belonging to a compact set $F \subset \mathcal{X}$ by interpolating spaces X_n spanned by functions that are not necessarily of polynomial type. This is achieved by the construction of suitable sets of interpolating spaces and the selection of suitable interpolating points S_n thanks to a greedy selection procedure.

The empirical interpolation process is, by construction, problem dependent given the fact that the constructed X_n and S_n depend on F. Furthermore, it is clear that the successful approximation of any function in F by this method requires to suppose that the set F is approximable by linear combinations of small size. In particular, this is the case when the Kolmogorov n-width $d_n(F, \mathcal{X})$ of F in $\mathcal{X}$ is small. Indeed, $d_n(F, \mathcal{X})$ is defined by

$$d_n(F, \mathcal{X}) := \inf_{\substack{X_n \subset \mathcal{X} \\ dim(X_n)=n}} \sup_{x \in F} \inf_{y \in X_n} \|x - y\|_{\mathcal{X}}$$

and measures the extent to which F can be approximated by finite dimensional spaces $X_n \subset \mathcal{X}$ of dimension n (see [10]). Several reasons can account for the rapid decrease of the Kolmogorov n-width: 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 is when $F = \{u(\mu, .),\ \mu \in D\}$, where D is a compact set of $\mathbb{R}^p$ and $u(\mu, .)$ is the solution of a PDE parametrized by μ. The approximation of any element $u(\mu, .) \in F$ by finite expansions is a classical problem addressed by reduced basis and the regularity of u in μ can also be a reason for having a small n-width as the results of [4] show.

In order to deal with functions that may not be continuous in space and also to account for experimental framework where data are acquired from sensors, an extension of this Lagrangian interpolation process has been proposed and is called GEIM as for Generalized Empirical Interpolation Method (see also [16], for another, though related approach to the problem of data assimilation). The method was first presented in [12] and consists in replacing the evaluation at interpolating points by application of a class of interpolating continuous linear forms chosen in a given dictionary $\Sigma \subset \mathcal{L}(\mathcal{X})$. In [13], it has been explained how GEIM can be extended to the frame of Banach spaces $\mathcal{X}$ and that EIM is a particular instance of it in the case where $\mathcal{X} = \mathcal{C}(\Omega)$ and the dictionary is composed of Dirac masses.

In this context, the present paper is a contribution to the understanding of the quality of this type of interpolation procedure through the analysis of the behavior of the interpolation error in GEIM in a framework of rapidly enough decreasing Kolmogorov n-width. To this purpose, the accuracy of the approximation in X_n of the elements of F will be compared to the best possible performance in an n-dimensional space which is measured by the Kolmogorov $n-$width $d_n(F, \mathcal{X})$. The present work is not the first contribution that studies the convergence rates of approximations of functions on spaces X_n constructed by greedy algorithms. Pioneer results in the case that $\mathcal{X}$ is a Hilbert can be found in [3] and [2]. An important extension of these works is [6] where the previous results were not only improved for the Hilbert framework but they were also extended to the case of Banach spaces. By employing the methodology proposed in [6], convergence rates for the generalized empirical interpolation were first presented in [14] when $\mathcal{X} = L^2(\Omega)$. As a sequel of [14] and still following the guidelines proposed in [6], we derive in this paper convergence rates for GEIM in the case of Banach spaces.

The document is organized as follows: in section 2 it will be shown that, under several hypothesis, the greedy algorithm of GEIM is of a weak greedy type (weak greedy algorithms are a category of greedy algorithms first identified in [2]). This observation is a preliminary step to analyze the convergence decay rates of the interpolation error. Section 3 provides these results in the case where $\mathcal{X}$ is a Banach space and in section 4 improved results will be derived in the particular case of Hilbert spaces.

2. The Generalized Empirical Interpolation Method. Let $\mathcal{X}$ 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 $\|.\|_{\mathcal{X}}$. Let F be a compact set of $\mathcal{X}$ whose elements $f \in F$ are such that $\|f\|_{\mathcal{X}} \leq 1$. With $\mathcal{N}$ being some given large number, we assume that the dimension of the vectorial space spanned by F is larger than $\mathcal{N}$. Our goal is to build, for all $n < \mathcal{N}$, a sequence of n-dimensional subspaces of $\mathcal{X}$ that approximate well enough any element of F. Assume also that we have at our disposal a dictionary of linear forms $\Sigma \subset \mathcal{L}(\mathcal{X})$ 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 n-dimensional interpolating spaces X_n spanned by functions $\{\varphi_0, \varphi_1, \dots, \varphi_{n-1}\}$ of F together with sets of n selected linear forms $\{\sigma_0, \sigma_1, \dots, \sigma_{n-1}\}$ coming from Σ such that any $\varphi \in F$ is well approximated by its generalized interpolant $\mathcal{J}_n[\varphi] \in X_n$. $\mathcal{J}_n[\varphi]$ has the following interpolation property:

$$\mathcal{J}_n[\varphi] = \sum_{j=0}^{n-1} \beta_j \varphi_j, \text{ such that } \quad \sigma_i(\mathcal{J}_n[\varphi]) = \sigma_i(\varphi), \ \forall i = 0, \dots, n-1. \tag{2.1}$$

The construction of the interpolation spaces X_n and the selection of the suitable associated elements of the dictionary is recursively carried out by a greedy algorithm. The search for the functions φ_i should ideally be done on F but this a too demanding task in practical applications. Hence, the search is in practice carried out over a discrete subset $\Xi_F \subset F$. For a fixed accuracy parameter $0 < \eta < 1$, there exists a discrete subset $\Xi_F^\eta \subset F$ such that the algorithm is of a weak greedy type as defined in section 1.3 of [2]. In the following, Ξ_F will denote this subset Ξ_F^η. Before proving its existence in lemma 2.1, let us momentarily assume this fact in order to explain how the search of the interpolating basis functions is carried out:

The first interpolating function φ_0 is chosen such that:

$$\|\varphi_0\|_{\mathcal{X}} = \max_{\varphi \in \Xi_F} \|\varphi\|_{\mathcal{X}} \geq \eta \sup_{\varphi \in F} \|\varphi\|_{\mathcal{X}},$$

the last inequality being a consequence of the definition of $\Xi_F \equiv \Xi_F^\eta$. The first interpolating linear form is

$$\sigma_0 = \arg\sup_{\sigma \in \Sigma} |\sigma(\varphi_0)|.$$

We then define the first basis function as $q_0 = \dfrac{\varphi_0}{\sigma_0(\varphi_0)}$ and the interpolation operator $\mathcal{J}_1 : \mathcal{X} \mapsto \text{span}\{q_0\}$ such that $\sigma_0(\varphi) = \sigma_0(\mathcal{J}_1[\varphi])$, for any $\varphi \in \mathcal{X}$. This yields the following expression:

$$\forall \varphi \in \mathcal{X}, \quad \mathcal{J}_1[\varphi] = \sigma_0(\varphi) q_0. \tag{2.2}$$

The second interpolating function φ_1 is chosen such that

$$\|\varphi_1 - \mathcal{J}_1[\varphi_1]\|_{\mathcal{X}} = \max_{\varphi \in \Xi_F} \|\varphi - \mathcal{J}_1[\varphi]\|_{\mathcal{X}} \geq \eta \sup_{\varphi \in F} \|\varphi - \mathcal{J}_1[\varphi]\|_{\mathcal{X}}.$$

The second interpolating linear form is

$$\sigma_1 = \arg\sup_{\sigma \in \Sigma} |\sigma(\varphi_1 - \mathcal{J}_1[\varphi_1])|,$$

and the second basis function is defined as

$$q_1 = \frac{\varphi_1 - \mathcal{J}_1[\varphi_1]}{\sigma_1(\varphi_1 - \mathcal{J}_1[\varphi_1])}.$$

We then proceed by induction. With $N_{\max} < \mathcal{N}$ being an upper bound fixed *a priori*, assume that, for a given $1 \leq n < N_{\max}$, we have built the set of interpolating functions $\{q_0, q_1, \ldots, q_{n-1}\}$ and the set of associated interpolating linear forms $\{\sigma_0, \sigma_1, \ldots, \sigma_{n-1}\}$ such that

$$\forall \varphi \in \mathcal{X}, \quad \mathcal{J}_n[\varphi] = \sum_{j=0}^{n-1} \alpha_j^n(\varphi) q_j$$

is well defined and the coefficients $\alpha_j^n(\varphi)$, $j = 0, \ldots, n-1$, are given by the interpolation problem

$$\begin{cases} \text{Find } \left(\alpha_j^n(\varphi)\right)_{j=0}^{n-1} \text{ such that:} \\ \sum_{j=0}^{n-1} \alpha_j^n(\varphi)\sigma_i(q_j) = \sigma_i(\varphi), \quad \forall i = 0, \ldots, n-1. \end{cases}$$

We now define

$$\forall \varphi \in \Xi_F, \quad \varepsilon_n(\varphi) = \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}}.$$

We choose φ_n such that

$$\varepsilon_n(\varphi_n) = \max_{\varphi \in \Xi_F} \varepsilon_n(\varphi) \geq \eta \sup_{\varphi \in F} \varepsilon_n(\varphi)$$

and $\sigma_n = \arg\sup_{\sigma \in \Sigma} |\sigma(\varphi_n - \mathcal{J}_n[\varphi_n])|$. The next basis function is then

$$q_n = \frac{\varphi_n - \mathcal{J}_n[\varphi_n]}{\sigma_n(\varphi_n - \mathcal{J}_n[\varphi_n])}.$$

We finally set $X_{n+1} \equiv \text{span}\{q_j,\ j \in [0, n]\} = \text{span}\{\varphi_j,\ j \in [0, n]\}$. The interpolation operator $\mathcal{J}_{n+1} : \mathcal{X} \mapsto X_{n+1}$ is given by

$$\forall \varphi \in \mathcal{X}, \quad \mathcal{J}_{n+1}[\varphi] = \sum_{j=0}^{n} \alpha_j^{n+1}(\varphi) q_j \tag{2.3}$$

and the coefficients $\alpha_j^{n+1}(\varphi)$, $j = 0, \dots, n$, are given by the interpolation problem

$$\begin{cases} \text{Find } \left(\alpha_j^{n+1}(\varphi)\right)_{j=0}^{n} \text{ such that:} \\ \sum_{j=0}^{n} \alpha_j^{n+1}(\varphi)\sigma_i(q_j) = \sigma_i(\varphi), \quad \forall i = 0, \dots, n. \end{cases}$$

It has been proven in [15] (for EIM) and [13] (for GEIM) that for any $1 \leq n \leq N_{\max}$, the set $\{q_j,\ j \in [0, n-1]\}$ is linearly independent and that the generalized empirical interpolation procedure is well-posed in $\mathcal{X}$. It has also been proven that the interpolation error satisfies:

$$\forall \varphi \in F, \quad \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1 + \Lambda_n) \inf_{\psi_n \in X_n} \|\varphi - \psi_n\|_{\mathcal{X}}, \tag{2.4}$$

where Λ_n is the Lebesgue constant in the $\mathcal{X}$ norm:

$$\Lambda_n := \sup_{\varphi \in \mathcal{X}} \frac{\|\mathcal{J}_n[\varphi]\|_{\mathcal{X}}}{\|\varphi\|_{\mathcal{X}}}. \tag{2.5}$$

Note that the parameter η quantifies the optimality of the greedy search: $\eta = 1$ will be the ideal case where $\Xi_F = F$ and the smaller the η, the worse Ξ_F will capture the interpolation behavior of the whole set F. Note also that Ξ_F^η cannot be easily determined in practice because its evaluation would require the computation of supremizers over the whole set F, which is not entirely possible in practice. The following lemma shows the existence of the discrete subset $\Xi_F = \Xi_F^\eta$, for any given η.

Lemma 2.1. *Let F be a compact subset of $\mathcal{X}$. Then, for any $0 < \eta < 1$, there exits a discrete subset Ξ_F^η such that*

$$\begin{cases} \max_{\varphi \in \Xi_F^\eta} \|\varphi\|_{\mathcal{X}} \geq \eta \sup_{\varphi \in F} \|\varphi\|_{\mathcal{X}}, \\ \max_{\varphi \in \Xi_F^\eta} \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \geq \eta \sup_{\varphi \in F} \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}}, \quad \forall\, n \in \{1, \dots, N_{\max}\}. \end{cases} \tag{2.6}$$

Proof. For a given $0 < \eta < 1$ and from the finite open cover property of the compact set F, there exists a discrete subset $\Xi_0^\eta \subset F$ and a function $\tilde{\varphi}_0 \in F$ such that:

$$\sup_{\varphi \in F} \inf_{\psi \in \Xi_0^\eta} \|\varphi - \psi\|_{\mathcal{X}} \leq (1 - \eta) \|\tilde{\varphi}_0\|_{\mathcal{X}}.$$

Let $\varphi_0 = \arg\max_{\psi \in \Xi_0^\eta} \|\psi\|_{\mathcal{X}}$ and $\varphi_0^{\sup} = \arg\sup_{\varphi \in F} \|\varphi\|_{\mathcal{X}}$. Then, for any $\psi \in \Xi_0^\eta$:

$$\|\varphi_0\|_{\mathcal{X}} \geq \|\psi\|_{\mathcal{X}} \geq -\|\psi - \varphi_0^{\sup}\|_{\mathcal{X}} + \|\varphi_0^{\sup}\|_{\mathcal{X}} \geq -(1-\eta)\|\tilde{\varphi}_0\|_{\mathcal{X}} + \|\varphi_0^{\sup}\|_{\mathcal{X}} \geq \eta \|\varphi_0^{\sup}\|_{\mathcal{X}}.$$

This completes the proof of the first inequality of (2.6). The second inequality is derived following the same guidelines: for any $1 \leq n \leq N_{\max}$, the application

$$\begin{aligned} r_n : \mathcal{X} &\mapsto \mathcal{X} \\ \varphi &\mapsto \varphi - \mathcal{J}_n[\varphi] \end{aligned}$$

is clearly continuous (with a norm that depends on Λ_n) and $r_n(F)$ is a compact subset of $\mathcal{X}$. From the finite open cover property of $r_n(F)$, there exists a discrete subset $\Xi_n^\eta \subset F$ and $\tilde{\varphi}_n \in F$ such that:

$$\sup_{\varphi \in F} \inf_{\psi \in \Xi_n^\eta} \|r_n[\varphi] - r_n[\psi]\|_{\mathcal{X}} \leq (1 - \eta) \|r_n[\tilde{\varphi}_n]\|_{\mathcal{X}}.$$

Let $\varphi_n = \arg\max_{\psi\in\Xi_n^\eta} \|r_n[\psi]\|_{\mathcal{X}}$ and $\varphi_n^{\sup} = \arg\sup_{\varphi\in F} \|r_n[\varphi]\|_{\mathcal{X}}$. Then, for any $\psi \in \Xi_n^\eta$:

$$\begin{aligned}\|\varphi_n - \mathcal{J}_n[\varphi_n]\|_{\mathcal{X}} &\geq \|r_n[\psi]\|_{\mathcal{X}} \\ &\geq -\|r_n[\psi] - r_n[\varphi_n^{\sup}]\|_{\mathcal{X}} + \|r_n[\varphi_n^{\sup}]\|_{\mathcal{X}} \\ &\geq -(1-\eta)\|r_n[\tilde{\varphi}_n]\|_{\mathcal{X}} + \|r_n[\varphi_n^{\sup}]\|_{\mathcal{X}} \\ &\geq \eta \, \|\varphi_n^{\sup} - \mathcal{J}_n[\varphi_n^{\sup}]\|_{\mathcal{X}}.\end{aligned}$$

The proof follows by taking

$$\Xi_F^\eta = \bigcup_{j=0}^{N_{\max}} \Xi_j^\eta.$$

□

Remark 2.2. *Note that the construction done in the proof is actually constructive in an adaptive and recursive way. Indeed, starting from the Ξ_0^η, that allows to define φ_0, the first interpolating function, the recursive update of the set Ξ^η can be done by adding a set Ξ_n^η defined similarly as $\Xi_0^{\eta_n}$, with $1-\eta_n = \frac{(1-\eta)}{(1+\Lambda_n)}\|r_n[\tilde{\varphi}_n]\|_X$, the evaluation of Λ_n being explained in [13] in the Hilbertian context.*

Remark 2.3. *In a similar manner as in the case where F is an infinite set of functions, if the dictionary Σ is not a finite set of linear forms, the greedy search is in practice carried out over a discrete subset $\widetilde{\Sigma} \subset \Sigma$. The choice of the subset $\widetilde{\Sigma}$ will have an impact on the definition of the sequence of subsets $(\Xi_F^\eta)_{j=0}^{N_{\max}}$ described in the proof of lemma 2.1. The "coarser" the choice on $\tilde{\Sigma}$, the "finer" the subsets $(\Xi_F^\eta)_{j=0}^{N_{\max}}$ must be in order to satisfy relation (2.6).*

Remark 2.4. *The Lebesgue constant Λ_n defined in our interpolation procedure depends both on the set F and on the choice of the dictionary of continuous linear forms Σ. In the case of Hilbert spaces, a formula for Λ_n has been given in [13] where the impact of the selected linear forms is expressed more explicitly than in formula (2.5) and allows for an easier implementation. Although no theoretical analysis about the impact of F or Σ on the behavior of (Λ_n) has been possible so far, one can find in [13] an illustration of these interactions in a simple numerical example. In the same reference, it is also outlined how the generalized interpolant of a function can be efficiently computed in practice by a recursion formula.*

3. Convergence rates of GEIM in a Banach space. In order to have a consistent notation in what follows, we define $\varphi_n = 0$ and $X_n = X_{N_{\max}}$ for $n > N_{\max}$.

3.1. Preliminary notations and properties. We remind that $\mathcal{X}$ is a Banach space. To fix some notations, let K be a nonempty subset of $\mathcal{X}$. For every $\varphi \in \mathcal{X}$, the distance between φ and the set K is denoted by $\mathrm{dist}(\varphi, K)$ and is defined by the following minimum equation:

$$\mathrm{dist}(\varphi, K) = \inf_{y\in K} \|\varphi - y\|_{\mathcal{X}}.$$

For any $\varphi \in \mathcal{X}$, the metric projection of φ onto K is given by the set

$$P_K(\varphi) = \{z \in K : \|\varphi - z\|_{\mathcal{X}} = \mathrm{dist}(\varphi, K)\}.$$

In general, this set can be empty or composed of one or more than one element. However, in the particular case where K is a finite dimensional space, $P_K(\varphi)$ is not empty. For any $n \geq 1$, the non empty set

$$P_n(\varphi) = \{z \in X_n : \|\varphi - z\|_{\mathcal{X}} = \operatorname{dist}(\varphi, X_n)\} \tag{3.1}$$

will denote the metric projection of $\varphi \in \mathcal{X}$ onto X_n. Since, the uniqueness of the metric projection onto X_n is not necessarily ensured, in the following, $P_n(\varphi)$ will denote one of the elements of the set (3.1). We also define for any $1 \leq n \leq N_{\max}$:

$$\tau_n(F)_{\mathcal{X}} := \max_{f \in F} \|f - P_n(f)\|_{\mathcal{X}}, \quad n = 1, 2, \dots \tag{3.2}$$

and

$$\gamma_n = \frac{\eta}{1 + \Lambda_n}, \quad \forall\ 1 \leq n \leq N_{\max}. \tag{3.3}$$

We will use the abbreviation τ_n and d_n for $\tau_n(F)_{\mathcal{X}}$ and $d_n(F, \mathcal{X})$. Likewise, (τ_n) and (d_n) will denote the sequences $(\tau_n(F)_{\mathcal{X}})_{n=1}^{\infty}$ and $(d_n(F, \mathcal{X}))_{n=1}^{\infty}$ respectively. We finish this section by proving the following lemma:

Lemma 3.1. *For any $n \geq 1$, $\|\varphi_n - P_n(\varphi_n)\|_{\mathcal{X}} \geq \gamma_n \tau_n$.*

Proof. From equation (2.4) applied to $\varphi = \varphi_n$ we have $\|\varphi_n - P_n(\varphi_n)\|_{\mathcal{X}} \geq \dfrac{1}{1+\Lambda_n}\|\varphi_n - \mathcal{J}_n(\varphi_n)\|_{\mathcal{X}}$. But $\|\varphi_n - \mathcal{J}_n(\varphi_n)\|_{\mathcal{X}} \geq \eta\|\varphi - \mathcal{J}_n(\varphi)\|_{\mathcal{X}}$ for any $\varphi \in F$ according to the definition of φ_n. Thus $\|\varphi_n - P_n(\varphi_n)\|_{\mathcal{X}} \geq \gamma_n\|\varphi - \mathcal{J}_n(\varphi)\|_{\mathcal{X}} \geq \gamma_n\|\varphi - P_n(\varphi)\|_{\mathcal{X}}$. ◻

Thanks to lemma 3.1, we have proven that the weak greedy algorithm of GEIM has very similar properties as the abstract weak greedy algorithm analyzed in [6]. The difference is that, in our case, the parameter γ depends on the dimension n whereas in [6] γ was a constant. This observation will be the key to derive convergence decay rates in the sequence (τ_n) by extending the proofs of [6]. The main two lemmas that were derived in [6] (with γ independent of n) are recalled in lemmas 3.2 and 3.3 and section 3.2 presents their extension when γ depends on n. Then, by using equation (2.4), the results on the convergence of the interpolation error will easily follow (section 3.3).

Lemma 3.2 (Corollary 4.2 – *(ii)* of [6] – Polynomial decay rates for (τ_n) when $\gamma_n = \gamma$).

If, for $\alpha > 0$, we have $d_n \leq C_0 n^{-\alpha}$, $n = 1, 2, \dots$, then for any $0 < \beta < \min\{\alpha, 1/2\}$, we have $\tau_n \leq C_1 n^{-\alpha+1/2+\beta}$, $n = 1, 2, \dots$, with

$$C_1 := \max\left\{C_0 4^{4\alpha+1}\gamma^{-4}\left(\frac{2\beta+1}{2\beta}\right)^{\alpha};\ \max_{n=1,\dots,7} n^{\alpha-\beta-1/2}\right\}.$$

Lemma 3.3 (Corollary 4.2 – *(iii)* of [6] – Exponential decay rates for (τ_n) when $\gamma_n = \gamma$).

If, for $\alpha > 0$, $d_n \leq C_0 e^{-c_1 n^{\alpha}}$, $n = 1, 2, \dots$, then $\tau_n < \sqrt{2C_0}\gamma^{-1}\sqrt{n}e^{-c_2 n^{\alpha}}$, $n = 1, 2, \dots$, where $c_2 = 2^{-1-2\alpha}c_1$. The factor $\sqrt{n}$ can be deleted by reducing the constant c_2.

3.2. Convergence rates for (τ_n) in the case where (γ_n) is not constant. We look for an upper bound of the sequence (τ_n) that involves the sequence of Kolmogorov n-widths (d_n). The case $n = 1$ is addressed in

Lemma 3.4. *In the case where $n = 1$, we have the following upper bound for τ_1:*

$$\tau_1 \leq 2\left(1 + \frac{1}{\eta}\right) d_1.$$

Proof. Given the parameter η coming from the GEIM greedy algorithm, let $\beta > \dfrac{1}{\eta}$.
We begin by recalling and defining some notations:

- φ_0 is the first interpolating function chosen by the greedy algorithm and $X_1 = \text{span}\{\varphi_0\}$.
- For any φ, $P_1(\varphi)$ is the metric projection of φ onto X_1.
- Let $\|\varphi_0^{\text{sup}}\|_{\mathcal{X}} = \sup_{\varphi \in F} \|\varphi\|_{\mathcal{X}}$. From the greedy selection procedure: $\|\varphi_0\|_{\mathcal{X}} \geq \eta \|\varphi_0^{\text{sup}}\|_{\mathcal{X}}$.
- Let X_μ be the one dimensional subspace associated to d_1. In other words,
$$X_\mu = \underset{\substack{X_1 \subset \mathcal{X} \\ dim(X_1)=1}}{\arg\inf} \sup_{x \in F} \inf_{y \in X_1} \|x - y\|_{\mathcal{X}}$$
and
$$\forall\ \varphi \in \mathcal{X},\ \|\varphi - P_{X_\mu}(\varphi)\|_{\mathcal{X}} \leq d_1.$$
- Let $\varphi_\mu^{\text{sup}} = \arg\max_{\varphi \in F} \|P_{X_\mu}(\varphi)\|_{\mathcal{X}}$.

We now divide the proof by considering two complementary cases of values of $\|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}}$. If $\|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}} \leq \dfrac{1+\eta}{\eta - \frac{1}{\beta}}\, d_1$, we easily derive that

$$\begin{aligned}
\forall \varphi \in F,\ \|\varphi - P_1(\varphi)\|_{\mathcal{X}} &\leq \|\varphi\|_{\mathcal{X}} \\
&\leq \|\varphi - P_{X_\mu}(\varphi)\|_{\mathcal{X}} + \|P_{X_\mu}(\varphi)\|_{\mathcal{X}} \\
&\leq d_1 + \|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}} \\
&\leq \left(1 + \frac{1+\eta}{\eta - \frac{1}{\beta}}\right) d_1.
\end{aligned}$$

If $\|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}} \geq \dfrac{1+\eta}{\eta - \frac{1}{\beta}}\, d_1$, we start by deriving the following inequality for $\|P_{X_\mu}(\varphi_0)\|_{\mathcal{X}}$:

$$\begin{aligned}
\|P_{X_\mu}(\varphi_0)\|_{\mathcal{X}} &\geq \|\varphi_0\|_{\mathcal{X}} - d_1 \\
&\geq \eta \|\varphi_0^{\text{sup}}\|_{\mathcal{X}} - d_1 \\
&\geq \eta \|\varphi_\mu^{\text{sup}}\|_{\mathcal{X}} - d_1 \\
&\geq \eta \left(\|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}} - d_1\right) - d_1 \\
&\geq \eta \|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}} - (1+\eta) d_1 \\
&\geq \frac{\|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}}}{\beta}.
\end{aligned} \tag{3.4}$$

From inequality (3.4), it follows that $\|P_{X_\mu}(\varphi_0)\|_{\mathcal{X}} > 0$ given that $\|P_{X_\mu}(\varphi_\mu^{\text{sup}})\|_{\mathcal{X}}$ is strictly positive. Furthermore, for any $\varphi \in \mathcal{X}$, there exits $\lambda_\varphi \in \mathbb{R}_+$ such that:

$$P_{X_\mu}(\varphi) = \lambda_\varphi P_{X_\mu}(\varphi_0). \tag{3.5}$$

Hence the decomposition:

$$\begin{aligned}
\varphi &= P_{X_\mu}(\varphi) + \varphi - P_{X_\mu}(\varphi) \\
&= \lambda_\varphi P_{X_\mu}(\varphi_0) + \varphi - P_{X_\mu}(\varphi) \\
&= \lambda_\varphi (P_{X_\mu}(\varphi_0) - \varphi_0) + \lambda_\varphi \varphi_0 + \varphi - P_{X_\mu}(\varphi)
\end{aligned} \tag{3.6}$$

Equation (3.6) yields:

$$\begin{aligned}\|\varphi - P_1(\varphi)\|_{\mathcal{X}} &\leq \|\varphi - \lambda_\varphi \varphi_0\|_{\mathcal{X}} \\ &\leq |\lambda_\varphi| \|P_{X_\mu}(\varphi_0) - \varphi_0\|_{\mathcal{X}} + \|\varphi - P_{X_\mu}(\varphi)\|_{\mathcal{X}} \\ &\leq (1 + |\lambda_\varphi|) d_1,\end{aligned}$$

Furthermore, given that $\|P_{X_\mu}(\varphi_\mu^{\sup})\|_{\mathcal{X}} \geq \|P_{X_\mu}(\varphi)\|_{\mathcal{X}}$ for any $\varphi \in F$, we have

$$\|P_{X_\mu}(\varphi_\mu^{\sup})\|_{\mathcal{X}} \geq |\lambda_\varphi| \|P_{X_\mu}(\varphi_0)\|_{\mathcal{X}}, \tag{3.7}$$

where we have used equality (3.5). Inequalities (3.4) and (3.7) yield $|\lambda_\varphi| \leq \beta$ and therefore

$$\|\varphi - P_1(\varphi)\|_{\mathcal{X}} \leq (1+\beta) d_1.$$

Hence, we have proven that for any $\beta > 1/\eta$ and any $\varphi \in F$, we have

$$\|\varphi - P_1(\varphi)\|_{\mathcal{X}} \leq \max\left(1+\beta; 1 + \frac{1+\eta}{\eta - \frac{1}{\beta}}\right) d_1.$$

If we define

$$\forall \beta > 1/\eta, \quad g_\eta(\beta) := \max\left(1+\beta; 1 + \frac{1+\eta}{\eta - \frac{1}{\beta}}\right),$$

it follows that $\|\varphi - P_1(\varphi)\|_{\mathcal{X}} \leq \min_{\beta > 1/2} g_\eta(\beta) d_1 = 2\left(1 + \frac{1}{\eta}\right) d_1$. ☐

For higher dimensions ($n > 1$), we first begin by proving

THEOREM 3.5. *For any $N \geq 0$, consider a weak greedy algorithm with the property of lemma 3.1 and constant γ_N. We have the following inequalities between τ_N and d_N: for any $K \geq 1$, $1 \leq m < K$*

$$\prod_{i=1}^{K} \tau_{N+i}^2 \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^2} 2^K K^{K-m} \left(\sum_{i=1}^{K} \tau_{N+i}^2\right)^m d_m^{2(K-m)} \tag{3.8}$$

Proof. This result is an extension of theorem 4.1 of [6] to the case where the parameter of the weak greedy algorithm (γ_N) depends on the dimension of the reduced space X_N. Its proof consists in a slight modification of the demonstration presented in [6]. The complete proof is given in the appendix section for the self-consistency of this paper. ☐

This theorem easily yields the following useful corollaries.

COROLLARY 3.6. *For any $n \geq 1$, we have:*

$$\tau_n \leq \frac{1}{\prod_{i=1}^{n} \gamma_i^{1/n}} \sqrt{2} \min_{1 \leq m < n} \left\{ n^{\frac{n-m}{2n}} \left(\sum_{i=1}^{n} \tau_i^2\right)^{\frac{m}{2n}} d_m^{\frac{n-m}{n}} \right\} \tag{3.9}$$

In particular, for any $\ell \geq 1$:

$$\tau_{2\ell} \leq 2 \frac{1}{\prod_{i=1}^{2\ell} \gamma_i^{1/2\ell}} \sqrt{\ell d_\ell}. \tag{3.10}$$

Proof. We take $N = 0$, $K = n$ and any $1 \leq m < n$ in (3.8) and use the monotonicity of (τ_n) to obtain:

$$\tau_n^{2n} \leq \prod_{i=1}^{n} \tau_i^2 \leq \frac{1}{\prod_{i=1}^{n} \gamma_i^2} 2^n n^{n-m} \left(\sum_{i=1}^{n} \tau_i^2 \right)^m d_m^{2(n-m)}.$$

If we take the $2n$-th root on both sides, we arrive at (3.9). In particular, if $n = 2\ell$ and $m = \ell$, we have:

$$\tau_{2\ell} \leq \frac{1}{\prod_{i=1}^{2\ell} \gamma_i^{1/2\ell}} \sqrt{2}(2\ell)^{1/4} \left(\sum_{i=1}^{2\ell} \tau_i^2 \right)^{1/4} \sqrt{d_\ell} \leq \frac{1}{\prod_{i=1}^{2\ell} \gamma_i^{1/2\ell}} \sqrt{2}(2\ell)^{1/4} (2\ell)^{1/4} \sqrt{d_\ell} = 2 \frac{1}{\prod_{i=1}^{2\ell} \gamma_i^{1/2\ell}} \sqrt{\ell d_\ell},$$

where we have used the fact that all $\tau_i \leq 1$. ☐

Corollary 3.7. *For $N \geq 0$, $K \geq 1$ and $1 \leq m < K$:*

$$\tau_{N+K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2K} \tau_{N+1}^{m/K} d_m^{1-m/K} \tag{3.11}$$

Proof. Given that (τ_n) is a monotonically decreasing sequence as is obtained by following the same lines as above, we derive from inequality (3.8) that:

$$\tau_{N+K}^{2K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^2} 2^K K^{K-m} \left(\sum_{i=1}^{K} \tau_{N+i}^2 \right)^m d_m^{2(K-m)}$$

Therefore,

$$\tau_{N+K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2} K^{\frac{K-m}{2K}} \left(K \tau_{N+1}^2 \right)^{m/2K} d_m^{1-m/K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2K} \tau_{N+1}^{m/K} d_m^{1-m/K}.$$

☐

We now derive convergence rates in (τ_n) for polynomial or exponential convergence of the Kolmogorov n-width. In the first two lemmas 3.8 and 3.9, the result is derived without making any assumption on the behavior of the sequence (γ_n) (that depends on the Lebesgue constant of GEIM).

Lemma 3.8 (Polynomial decay of (d_n)). *For any $n \geq 1$, let $n = 4\ell + k$ (where $\ell \in \{0, 1, \dots\}$ and $k \in \{0, 1, 2, 3\}$). Assume that there exists a constant $C_0 > 0$ such that $\forall n \geq 1$, $d_n \leq C_0 n^{-\alpha}$, then*

$$\tau_n \leq C_0 \beta_n n^{-\alpha}, \tag{3.12}$$

where

$$\beta_n = \beta_{4\ell+k} := \begin{cases} 2\left(1 + \dfrac{1}{\eta}\right) & \text{if } n = 1 \\ \dfrac{1}{\prod_{i=1}^{\ell_2} \gamma_{\ell_1 - \lceil \frac{k}{4} \rceil + i}^{\frac{1}{\ell_2}}} \sqrt{2\ell_2 \beta_{\ell_1}} (2\sqrt{2})^\alpha & \text{if } n \geq 2 \end{cases}$$

and $\ell_1 = 2\ell + \lfloor \frac{2k}{3} \rfloor$, $\ell_2 = 2\left(\ell + \lceil \frac{k}{4} \rceil\right)$, *where* $\lfloor . \rfloor$ *and* $\lceil . \rceil$ *are the floor and ceiling functions respectively.*

Proof. The proof is done by recurrence over n and the case $n = 1$ directly follows from lemma 3.4. In the case $n \geq 2$:, we write $n = N + K$ with $N \geq 0$ and $K \geq 2$. Thanks to corollary 3.7, we have for any $1 \leq m < K$:

$$\tau_n = \tau_{N+K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2K} \tau_{N+1}^{m/K} d_m^{1-m/K} \tag{3.13}$$

We now use that $d_m \leq C_0 m^{-\alpha}$ and the recurrence hypothesis $\tau_{N+1} \leq C_0 \beta_{N+1} (N+1)^{-\alpha}$ which yield:

$$\tau_{N+K} \leq C_0 \sqrt{2K} \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{\frac{1}{K}}} \beta_{N+1}^{\frac{m}{K}} \xi(N, K, m)^{\alpha} (N + K)^{-\alpha}, \tag{3.14}$$

where $\xi(N, K, m) = \dfrac{N + K}{m} \left(\dfrac{m}{N + 1} \right)^{\frac{m}{K}}$ for any $1 \leq m < K$ and any given index $n = N + K \geq 2$, where $N \geq 0$ and $K \geq 2$.

Furthermore, any $n \geq 2$ can be written as $n = 4\ell + k$ with $\ell \in \mathbb{N}$ and $k \in \{0, 1, 2, 3\}$. If $k = 1, 2$ or 3, it can easily be proven that the function ξ is bounded by $2\sqrt{2}$ by setting

$$\begin{cases} N = 2\ell - 1,\ K = 2\ell + 2,\ m = \ell + 1 \text{ and } \ell \geq 1 \text{ in the case } k = 1, \\ N = 2\ell,\ K = 2\ell + 2,\ m = \ell + 1 \text{ and } \ell \geq 0 \text{ in the case } k = 2, \\ N = 2\ell + 1,\ K = 2\ell + 2,\ m = \ell + 1 \text{ and } \ell \geq 0 \text{ in the case } k = 3. \end{cases}$$

These choices of N, K and m combined with the upper bound of ξ yield the result $\tau_n \leq C_0 \beta_n n^{-\alpha}$ in the case $k = 1, 2$ or 3.

To deal with the case $n = 4\ell$, we come back to estimate (3.13) and use the fact that $\tau_{N+1} \leq \tau_N$. It follows that:

$$\tau_n \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2K} \tau_N^{m/K} d_m^{1-m/K}. \tag{3.15}$$

If we choose $N = K = 2\ell$ and $m = \ell$, the inequality (3.15) directly yields the desired result

$$\tau_{4\ell} \leq C_0 \sqrt{2} \sqrt{2\ell \beta_{2\ell}} \frac{1}{\prod_{i=1}^{2\ell} \gamma_{2\ell+i}^{\frac{1}{2\ell}}} (2\sqrt{2})^{\alpha} (4\ell)^{-\alpha}.$$

□

Lemma 3.9 (Exponential decay in (d_n)). *Assume that there exist constants* $C_0 \geq 1$ *and* $\alpha > 0$ *such that* $\forall n \geq 1$, $d_n \leq C_0 e^{-c_1 n^{\alpha}}$, *then*

$$\tau_n \leq C_0 \beta_n e^{-c_2 n^{\alpha}},$$

where $c_2 := c_1 2^{-2\alpha-1}$ *and*

$$\beta_n := \begin{cases} 2\left(1+\frac{1}{\eta}\right), & \textit{if } n=1 \\ \sqrt{2}\frac{1}{\prod_{i=1}^{2\lfloor\frac{n}{2}\rfloor} \gamma_i^{\frac{1}{2\lfloor\frac{n}{2}\rfloor}}}\sqrt{n}, & \textit{if } n\geq 2. \end{cases}$$

Proof. The case $n=1$ easily follows from lemma 3.4. For $n=2\ell$ ($\ell\geq 1$), inequality (3.10) directly yields:

$$\tau_{2\ell} \leq 2\frac{1}{\prod_{i=1}^{2\ell}\gamma_i^{1/2\ell}}\sqrt{\ell d_\ell} \leq C_0\sqrt{2}\frac{1}{\prod_{i=1}^{2\ell}\gamma_i^{1/2\ell}}\sqrt{2\ell}e^{-\frac{c_1}{2^{1+\alpha}}(2\ell)^\alpha}, \tag{3.16}$$

where we have used the fact that $d_\ell \leq C_0 e^{-c_1(\ell)^\alpha}$ and that $C_0\geq 1$. For $n=2\ell+1$, by using inequality (3.16) and the fact that $\tau_{2\ell+1}\leq\tau_{2\ell}$, we have:

$$\tau_{2\ell+1} \leq C_0\sqrt{2}\frac{1}{\prod_{i=1}^{2\ell}\gamma_i^{1/2\ell}}\sqrt{2\ell}e^{-\frac{c_1}{2^{1+\alpha}}(2\ell)^\alpha} \leq C_0\sqrt{2}\frac{1}{\prod_{i=1}^{2\ell}\gamma_i^{1/2\ell}}\sqrt{2\ell+1}e^{-\frac{c_1}{2^{1+2\alpha}}(2\ell+1)^\alpha}. \tag{3.17}$$

□

The sequence (γ_n) is directly related to the Lebesgue constant sequence (Λ_n) and, in the particular case where (Λ_n) is monotonically increasing. It is therefore interesting to analyze the convergence rates that lemmas 3.8 and 3.9 provide in this particular case and the following corollary accounts for it.

Corollary 3.10. *In the case where (Λ_n) is monotically increasing (i.e. (γ_n) monotonically decreasing), the following bounds can be derived for τ_n:*

i) If $d_n\leq C_0 n^{-\alpha}$ for any $n\geq 1$, then $\tau_n\leq C_0\tilde\beta_n n^{-\alpha}$, with

$$\tilde\beta_n := \begin{cases} 2\left(1+\frac{1}{\eta}\right), & \textit{if } n=1 \\ 2^{3\alpha+1}\ell_2\left(\frac{1}{\gamma_n}\right)^2, & \forall n\geq 2. \end{cases}$$

If we write n as $n=4\ell+k$ (with $\ell\in\{0,1,\dots\}$ and $k\in\{0,1,2,3\}$), then $\ell_2 = 2\left(\ell+\lceil\frac{k}{4}\rceil\right)$.

ii) If $d_n\leq C_0 e^{-c_1 n^\alpha}$ for $n\geq 1$ and $C_0\geq 1$, then $\tau_n\leq C_0\tilde\beta_n e^{-c_2 n^{-\alpha}}$, with $c_2 = c_1 2^{-2\alpha-1}$

$$\tilde\beta_n := \begin{cases} 2\left(1+\frac{1}{\eta}\right), & \textit{if } n=1 \\ \sqrt{2}\frac{1}{\gamma_n}\sqrt{n}, & \textit{if } n\geq 2. \end{cases}$$

Proof.

i) The proof consists in showing by recursion that $\tilde\beta_n$ is larger than the coefficient β_n defined in lemma 3.8.

In the case $n = 1$, $\tilde{\beta}_1 = \beta_1$. Then, for $n > 1$, given that (γ_n) is monotonically decreasing, we have

$$\beta_n \leq \frac{1}{\gamma_n}\sqrt{2\ell_2\beta_{\ell_1}}\left(2\sqrt{2}\right)^{\alpha} \leq \frac{1}{\gamma_n}\sqrt{2\ell_2\tilde{\beta}_{\ell_1}}\left(2\sqrt{2}\right)^{\alpha},$$

where we have used the recurrence hypothesis $\beta_{\ell_1} \leq \tilde{\beta}_{\ell_1}$ in the second inequality. Furthermore, since

$$\tilde{\beta}_{\ell_1} \leq 2^{3\alpha+1}\ell_2\gamma_n^{-2},$$

it follows that:

$$\beta_n \leq \gamma_n^{-1}\sqrt{2\ell_2 2^{3\alpha+1}\ell_2\gamma_n^{-2}}\left(2\sqrt{2}\right)^{\alpha} = 2^{3\alpha+1}\ell_2\gamma_n^{-2} = \tilde{\beta}_n.$$

ii) The result is straightforward and follows from the definition of β_n given in lemma 3.9.

□

In the case where (γ_n) is constant, corollary 3.10 shows that we obtain exactly the same result as the one derived in [6] for the exponential case (recalled in lemma 3.3 in this paper). In the case of polynomial decay, the result of corollary 3.10 provides a slightly degraded result with respect to the one presented in [6] (recalled in lemma 3.2). The most important difference relies in the fact that the authors get a convergence rate of order $\mathcal{O}(n^{-\alpha+1/2+\varepsilon})$ whereas the present results yields a convergence in $\mathcal{O}(n^{-\alpha+1})$.

It has so far not been possible to derive better convergence rates in the polynomial case for a general behavior of the sequence (Λ_n). Therefore, in an attempt to recover the convergence of order $\mathcal{O}(n^{-\alpha+1/2+\varepsilon})$ in the polynomial case, we propose to assume a particular behavior of the Lebesgue constant. In the case case where (Λ_n) presents a polynomial increasing behavior

$$\Lambda_n = \mathcal{O}(n^{\zeta}),$$

lemma 3.11 shows that the convergence is of order $\mathcal{O}(n^{-\alpha+\zeta+1/2+\varepsilon})$, which is, in some sense, similar to the result of [6].

Lemma 3.11 (Polynomial decay of (d_n) and polynomial increase in (Λ_n)). *If for $\alpha > 0$, we have $d_n \leq C_0 n^{-\alpha}$ and $\gamma_n^{-1} \leq C_\zeta n^{\zeta}$, $n \in \mathbb{N}^*$, then for any $\beta > 1/2$, we have $\tau_n \leq C_1 n^{-\alpha+\zeta+\beta}$, $n \in \mathbb{N}^*$, where*

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

Note that in the above lemma, the constant β has no connection with β_n defined above.

Proof. It follows from the monotonicity of (τ_n) and inequality (3.8) for $N = K = n$ and any $1 \leq m < n$ that:

$$\tau_{2n} \leq \sqrt{2n}\frac{1}{\prod_{i=1}^{n}\gamma_{n+i}^{1/n}}\tau_n^{\delta}d_m^{1-\delta}, \quad \delta := \frac{m}{n} \tag{3.18}$$

Given $\beta > 1/2$, we define $m := \left\lfloor \frac{\beta-\frac{1}{2}}{\zeta+\beta}\right\rfloor + 1$ (so that $m < n$ for $n > 2(\zeta+\beta) > 2\zeta+1$). It follows that

$$\delta = \frac{m}{n} \in \left(\frac{\beta-\frac{1}{2}}{\zeta+\beta}; \frac{\beta-\frac{1}{2}}{\zeta+\beta}+\frac{1}{n}\right) \tag{3.19}$$

We prove our claim by contradiction. Suppose it is not true and M is the first value where $\tau_M > C_1 M^{-\alpha+\zeta+\beta}$. Clearly, because of the definition of C_1 and the fact that $\tau_n \leq 1$, we must have $M > 2\lfloor 2(\zeta+\beta)\rfloor + 1$ (since $M \geq 2\lfloor 2(\zeta+\beta)\rfloor + 2$). We first consider the case $M = 2n$, and therefore $n \geq \lfloor 2(\zeta+\beta)\rfloor + 1$. From (3.18), we have:

$$\begin{aligned} C_1(2n)^{-\alpha+\zeta+\beta} < \tau_{2n} &\leq \sqrt{2n}\frac{1}{\prod_{i=1}^{n}\gamma_{n+i}^{1/n}}\tau_n^{\delta}d_m^{1-\delta} \\ &\leq \sqrt{2n}C_\zeta(2n)^\zeta C_1^\delta n^{\delta(-\alpha+\zeta+\beta)}C_0^{1-\delta}(\delta n)^{-\alpha(1-\delta)}, \end{aligned} \tag{3.20}$$

where we have used the fact that $\tau_n \leq C_1 n^{-\alpha+\zeta+\beta}$ and $d_m \leq C_0 m^{-\alpha}$. It follows that

$$C_1^{1-\delta} < 2^{\alpha-\beta+\frac{1}{2}}C_\zeta C_0^{1-\delta}\delta^{-\alpha(1-\delta)}n^{\frac{1}{2}+\delta(\zeta+\beta)-\beta}$$

and therefore

$$C_1 < 2^{\frac{\alpha-\beta+\frac{1}{2}}{1-\delta}}C_\zeta^{\frac{1}{1-\delta}}C_0\delta^{-\alpha}n^{\frac{\zeta+\beta}{1-\delta}\left(\delta-\frac{\beta-\frac{1}{2}}{\zeta+\beta}\right)}.$$

Since, for $n \geq \lfloor 2(\zeta+\beta)\rfloor + 1 > 2(\zeta+\beta)$, we have

$$\begin{aligned} \delta &< \frac{\beta-\frac{1}{2}}{\zeta+\beta} + \frac{1}{n} \\ &< \frac{\beta}{\zeta+\beta}, \end{aligned} \tag{3.21}$$

then,

$$\frac{1}{1-\delta} < \frac{\zeta+\beta}{\zeta}. \tag{3.22}$$

Hence,

$$\frac{\zeta+\beta}{1-\delta}\left(\delta-\frac{\beta-\frac{1}{2}}{\zeta+\beta}\right) < \left(\frac{\zeta+\beta}{1-\delta}\right)\frac{1}{n} < \frac{(\zeta+\beta)^2}{\zeta}\frac{1}{n}, \tag{3.23}$$

where we have used inequalities (3.21) and (3.22). By using (3.23), it follows that

$$n^{\frac{\zeta+\beta}{1-\delta}\left(\delta-\frac{\beta-\frac{1}{2}}{\zeta+\beta}\right)} < n^{\frac{(\zeta+\beta)^2}{\zeta}\frac{1}{n}} < 2^{\frac{(\zeta+\beta)^2}{\zeta}}. \tag{3.24}$$

This yields:

$$C_1 < 2^{\frac{\alpha-\beta+\frac{1}{2}}{1-\delta}}C_\zeta^{\frac{1}{1-\delta}}C_0\delta^{-\alpha}2^{\frac{(\zeta+\beta)^2}{\zeta}} < 2^{\left(\frac{\zeta+\beta}{\zeta}\right)(\alpha+\zeta+\frac{1}{2})}C_\zeta^{\frac{1}{1-\delta}}C_0\delta^{-\alpha}. \tag{3.25}$$

Furthermore, for $-\alpha+\zeta+\beta < 0$ (which is the meaningful case), and using the fact that $\beta > \dfrac{1}{2}$, we have:

$$2^{\frac{\zeta+\beta}{\zeta}\left(\alpha+\zeta+\frac{1}{2}\right)} < 2^{\frac{\alpha}{\zeta}(\alpha+\zeta+\beta)} < 2^{\frac{2\alpha^2}{\zeta}} \tag{3.26}$$

and

$$C_\zeta^{\frac{1}{1-\delta}} < \max\left(1; C_\zeta^{\frac{\zeta+\beta}{\zeta}}\right) \tag{3.27}$$

Also, from (3.19), we have

$$\delta^{-\alpha} < \left(\frac{\zeta+\beta}{\beta-\frac{1}{2}}\right)^{\alpha} \tag{3.28}$$

By inserting inequalities (3.26), (3.27) and (3.28) in relation (3.25), the desired contradiction follows:

$$C_1 < C_0 2^{\frac{2\alpha^2}{\zeta}} \left(\frac{\zeta+\beta}{\beta-\frac{1}{2}}\right)^{\alpha} \max\left(1; C_\zeta^{\frac{\zeta+\beta}{\zeta}}\right).$$

Likewise, if $M = 2n+1$, hence is odd, the actually $M \geq 2\lfloor 2(\zeta+\beta)\rfloor + 3$, implying that $n \geq \lfloor 2(\zeta+\beta)\rfloor + 1$:

$$C_1 2^{-\alpha+\zeta+\beta}(2n)^{-\alpha+\zeta+\beta} < C_1(2n+1)^{-\alpha+\zeta+\beta} < \tau_{2n+1} \leq \tau_{2n}. \tag{3.29}$$

But, since from equation (3.20) we have

$$\tau_{2n} \leq \sqrt{2n} C_\zeta (2n)^\zeta C_1^\delta n^{\delta(-\alpha+\zeta+\beta)} C_0^{1-\delta}(\delta n)^{-\alpha(1-\delta)}, \tag{3.30}$$

then, following the same argument as above, we get:

$$C_1 < C_0 2^{\left(\frac{\zeta+\beta}{\zeta}\right)\left(\frac{1}{2}+2\alpha-\beta\right)} \left(\frac{\zeta+\beta}{\beta-\frac{1}{2}}\right)^{\alpha} \max\left(1; C_\zeta^{\frac{\zeta+\beta}{\zeta}}\right) \tag{3.31}$$

$$< C_0 2^{\frac{2\alpha^2}{\zeta}} \left(\frac{\zeta+\beta}{\beta-\frac{1}{2}}\right)^{\alpha} \max\left(1; C_\zeta^{\frac{\zeta+\beta}{\zeta}}\right), \tag{3.32}$$

where we have used the fact that $\beta > 1/2$ in the last inequality. □

3.3. Convergence rates of the interpolation error. Lemmas 3.8, 3.9 are the keys to derive the decay rates of the interpolation error of the GEIM greedy algorithm for any behavior of the sequence (γ_n). This is the purpose of the following theorem:

Theorem 3.12 (Convergence rates for GEIM in a Banach space).

i) Assume that $d_n \leq C_0 n^{-\alpha}$ for any $n \geq 1$, then the interpolation error of the GEIM greedy selection process satisfies for any $\varphi \in F$ the inequality $\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1+\Lambda_n)C_0\beta_n n^{-\alpha}$, where the parameter β_n is defined as in lemma 3.8.

ii) Assume that $d_n \leq C_0 e^{-c_1 n^\alpha}$ for $n \geq 1$ and $C_0 \geq 1$, then the interpolation error of the GEIM greedy selection process satisfies for any $\varphi \in F$ the inequality $\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1+\Lambda_n)C_0\beta_n e^{-c_2 n^\alpha}$, where β_n and c_2 are defined as in lemma 3.9.

Proof. It can be inferred from equation (2.4) that, $\forall \varphi \in F$, $\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1+\Lambda_n)\|\varphi - P_n(\varphi)\|_{\mathcal{X}} \leq (1+\Lambda_n)\tau_n$ according to the definition of τ_n. We conclude the proof by bounding τ_n thanks to lemmas 3.8, 3.9. □

From corollary 3.10 and lemma 3.11, we can also derive convergence rates in the case where (Λ_n) is a monotonically increasing sequence. This is summarized in

Corollary 3.13. *If (Λ_n) is a monotonically increasing sequence, then the sequence (γ_n) in the GEIM procedure is monotonically decreasing. The following decay rates in the generalized interpolation error can be inferred:*

i) If $d_n \leq C_0 n^{-\alpha}$ for any $n \geq 1$, then the interpolation error of the GEIM greedy selection process can be bounded as

$$\forall \varphi \in F, \quad \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq \begin{cases} 2C_0\left(1+\frac{1}{\eta}\right)(1+\Lambda_1), & \text{if } n = 1. \\ C_0 2^{3\alpha+1}\ell_2 \frac{(1+\Lambda_n)^3}{\eta^2} n^{-\alpha}, & \text{if } n \geq 2. \end{cases}$$

If we write n as $n = 4\ell + k$ (with $\ell \in \{0, 1, \dots\}$ and $k \in \{0, 1, 2, 3\}$), then $\ell_2 = 2\left(\ell + \lceil \frac{k}{4} \rceil\right)$.

ii) If $d_n \leq C_0 e^{-c_1 n^\alpha}$ for $n \geq 1$ and $C_0 \geq 1$, then the interpolation error of the GEIM greedy selection process can be bounded as

$$\forall \varphi \in F, \quad \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq \begin{cases} 2C_0 \left(1 + \dfrac{1}{\eta}\right)(1 + \Lambda_1), & \text{if } n = 1, \\ C_0 \sqrt{2} \dfrac{(1 + \Lambda_n)^2}{\eta} \sqrt{n} e^{-c_2 n^\alpha}, & \text{if } n \geq 2, \end{cases}$$

where $c_2 = c_1 2^{-2\alpha - 1}$.

iii) If $d_n \leq C_0 n^{-\alpha}$ and that $\gamma_n^{-1} \leq C_\zeta n^\zeta$ for any $n \geq 1$, then the interpolation error of the GEIM greedy selection process satisfies for any $\beta > 1/2$,

$$\forall \varphi \in F, \quad \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq \eta C_\zeta C_1 n^{-\alpha + 2\zeta + \beta},$$

where the parameter C_1 is defined as in lemma 3.11.

Proof. i) and ii) easily follow from corollary 3.10 and iii) is derived by using lemma 3.11. ∎

Remark 3.14. *The evolution of the Lebesgue constant Λ_n as a function of n is a subject of great interest. From the theoretical point of view, crude estimates exist and provide an exponential upper bound. This is however far from being what has been obtained in practical applications where, for a reasonable enough choice of the dictionary Σ, the sequence (Λ_n) presents, in the worst case scenario, a linearly increasing behavior (see [13] for a discussion about this issue and also [1], [8], [15] in the case of the traditional EIM). Assuming this type of behavior for (Λ_n), from corollary 3.13-iii, it follows that a polynomial decrease of the Kolmogorov n-width of order $\mathcal{O}(n^{-3})$ should be enough to ensure the convergence of the interpolation error of GEIM.*

4. Convergence rates of GEIM in a Hilbert space.

4.1. Preliminary notations and properties. In this section, $\mathcal{X}$ is a Hilbert space equipped with its induced norm $\|f\|_{\mathcal{X}} = (f, f)_{\mathcal{X}}$, where $(.,.)_{\mathcal{X}}$ is the scalar product in $\mathcal{X}$.

In the same spirit as in the case of a Banach space, we define the sequence (τ_n) as in formula 3.2 but now, for any $f \in F$, $P_n(f)$ corresponds to the unique element of X_n that is the orthogonal projection of f onto X_n. Note that lemma 3.1 still holds in the Hilbert setting. We address the task of deriving convergence rates for the interpolation of GEIM by applying the same technique of section 3, i.e. by first deriving convergence rates on (τ_n) (see section 4.2). The obtained results will be compared to the ones presented in [6] in corollary 3.3 for the case $\gamma_n = \gamma$ and that are recalled here:

Lemma 4.1 (Corollary 3.3 – (ii) of [6] – Polynomial decay rates for (τ_n) when $\gamma_n = \gamma$).

If $d_n \leq C_0 n^{-\alpha}$ for $n = 1, 2, \dots$, then $\tau_n \leq C_1 n^{-\alpha}$, $n = 1, 2, \dots$, with $C_1 = 2^{5\alpha+1} \gamma^{-2} C_0$.

Lemma 4.2 (Corollary 3.3 – (iii) of [6] – Exponential decay rates for (τ_n) when $\gamma_n = \gamma$).

If $d_n \leq C_0 e^{-c_1 n^\alpha}$ for $n = 1, 2, \dots$, then $\tau_n < \sqrt{2C_0} \gamma^{-1} e^{-c_2 n^\alpha}$, $n = 1, 2, \dots$, where $c_2 = 2^{-1-2\alpha} c_1$.

4.2. Convergence rates for (τ_n). In order to extend lemmas 4.1 and 4.2 to the more general case where γ depends on the dimension n, the following preliminary theorem is required:

THEOREM 4.3. *For any $N \geq 0$, consider a weak Greedy algorithm with the property of lemma 3.1 and constant γ_N. We have the following inequalities between τ_N and d_N: for any $K \geq 1$, $1 \leq m < K$*

$$\prod_{i=1}^{K} \tau_{N+i}^2 \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^2} \left(\frac{K}{m}\right)^m \left(\frac{K}{K-m}\right)^{K-m} \tau_{N+1}^{2m} d_m^{2(K-m)}.$$

Proof. See appendix B. □

This theorem yields corollaries 4.4 and 4.5, that are the analogue for the Hilbert setting of corollaries 3.6 and 3.7.

COROLLARY 4.4. *For $N \geq 1$, we have*

$$\tau_n \leq \sqrt{2} \frac{1}{\prod_{i=1}^{n} \gamma_i^{1/n}} \min_{1 \leq m < n} d_m^{\frac{n-m}{n}}. \tag{4.1}$$

In particular,

$$\tau_{2n} \leq \sqrt{2} \frac{1}{\prod_{i=1}^{2n} \gamma_i^{\frac{1}{2n}}} \sqrt{d_n}. \tag{4.2}$$

COROLLARY 4.5. *For $N \geq 0$, $K \geq 1$ and $1 \leq m < K$:*

$$\tau_{N+K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2} \tau_{N+1}^{m/K} d_m^{1-m/K}. \tag{4.3}$$

Proof. [Proofs of corollaries 4.4 and 4.4] The proofs of these two results follow very similar guidelines as corolaries 3.6 and 3.7. The only difference is that here the staring point is theorem 4.3 instead of 3.5. □

The absence of the factor $\sqrt{n}$ in these corollaries will be the key to derive improved results in Hilbert spaces.

Using theorem 4.3, convergence rates in the sequence (τ_n) when (d_n) has a polynomial or an exponential decay can be inferred and lead to lemmas 4.6 and 4.7. In these results, no assumption on the behavior of (γ_n) has been made:

LEMMA 4.6 (Polynomial decay of (d_n)). *For any $n \geq 1$, let $n = 4\ell + k$ (where $\ell \in \{0, 1, \dots\}$ and $k \in \{0, 1, 2, 3\}$). Assume that there exists a constant $C_0 > 0$ such that $\forall n \geq 1$, $d_n \leq C_0 n^{-\alpha}$, then*

$$\tau_n \leq C_0 \beta_n n^{-\alpha}, \tag{4.4}$$

where

$$\beta_n = \beta_{4\ell+k} := \begin{cases} 2\left(1 + \frac{1}{\eta}\right) & \text{if } n = 1 \\ \sqrt{2\beta_{\ell_1}} \frac{1}{\prod_{i=1}^{\ell_2} \gamma_{\ell_1 - \lceil \frac{k}{4} \rceil + i}^{\frac{1}{\ell_2}}} (2\sqrt{2})^{\alpha} & \text{if } n \geq 2 \end{cases}$$

and $\ell_1 = 2\ell + \lfloor \frac{2k}{3} \rfloor$, $\ell_2 = 2\left(\ell + \lceil \frac{k}{4} \rceil\right)$.

Proof. The proof is similar to the one proposed in lemma 3.8: the case $n = 1$ directly follows from lemma 3.4 and in the case $n \geq 2$, we write $n = N + K$ with $N \geq 0$ and $K \geq 2$. Corollary 4.5 yields:

$$\tau_{N+K} \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{1/K}} \sqrt{2} \tau_{N+1}^{m/K} d_m^{1-m/K}.$$

By using that $d_m \leq C_0 m^{-\alpha}$ and the recurrence hypothesis $\tau_{N+1} \leq \beta_{N+1}(N+1)^{-\alpha}$, we get:

$$\tau_{N+K} \leq C_0 \sqrt{2} \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^{\frac{1}{K}}} \beta_{N+1}^{\frac{m}{K}} \xi(N,K,m)^{\alpha} (N+K)^{-\alpha},$$

where $\xi(N,K,m) = \dfrac{N+K}{m} \left(\dfrac{m}{N+1} \right)^{\frac{m}{K}}$ for any $1 \leq m < K$ and any given index $n = N + K \geq 2$, where $N \geq 0$ and $K \geq 2$. It suffices now to decompose any $n \geq 2$ as $n = 4\ell + k$ with $\ell \in \{0, 1, \dots\}$ and $k \in \{0, 1, 2, 3\}$ and use the same choices of N, K and m described in the proof of lemma 3.8 to derive the result. □

Lemma 4.7 (Exponential decay in (d_n)). *Assume that there exists a constant* $C_0 \geq 1$ *such that* $\forall n \geq 1$, $d_n \leq C_0 e^{-c_1 n^{\alpha}}$, *then*

$$\tau_n \leq C_0 \beta_n e^{-c_2 n^{\alpha}},$$

where $c_2 := c_1 2^{-2\alpha - 1}$ *and*

$$\beta_n := \begin{cases} 2\left(1 + \dfrac{1}{\eta}\right), & \text{if } n = 1 \\ \sqrt{2} \dfrac{1}{\prod_{i=1}^{2\lfloor \frac{n}{2} \rfloor} \gamma_i^{\frac{1}{2\lfloor \frac{n}{2} \rfloor}}}, & \text{if } n \geq 2. \end{cases}$$

Proof. The proof is the same as lemma 3.9 but uses corollary 4.4 instead of corollary 3.6. □

As in the Banach cases, it is important to study the convergence rates in the particular case where (Λ_n) is monotonically increasing. The following corollary accounts for it.

Corollary 4.8. *In the case where* (γ_n) *is a monotonically decreasing sequence, the following bounds can be derived for* τ_n*:*

i) If $d_n \leq C_0 n^{-\alpha}$ *for any* $n \geq 1$, *then* $\tau_n \leq C_0 \tilde{\beta}_n n^{-\alpha}$, *with*

$$\tilde{\beta}_n := \begin{cases} 2\left(1 + \dfrac{1}{\eta}\right), & \text{if } n = 1 \\ 2^{3\alpha+1} \left(\dfrac{1}{\gamma_n}\right)^2, & \text{if } n \geq 2. \end{cases}$$

ii) If $d_n \leq C_0 e^{-c_1 n^\alpha}$ *for* $n \geq 1$ *and* $C_0 \geq 1$*, then* $\tau_n \leq C_0 \tilde{\beta}_n e^{-c_2 n^{-\alpha}}$*, with*

$$\tilde{\beta}_n := \begin{cases} 2\left(1+\dfrac{1}{\eta}\right), & \text{if } n=1 \\ \sqrt{2}\dfrac{1}{\gamma_n}, & \text{if } n \geq 2. \end{cases}$$

Proof. The proof is derived by following the same guidelines as the proof of corollary 3.10. □

Remark 4.9. *As a direct consequence of corollary 4.8, if* γ_n *is constant, we recover slighly better results as the ones in [6] for* $n \geq 2$ *(see lemmas 4.1 and 4.2 above).*

4.3. Convergence rates of the interpolation error. Lemmas 4.6 and 4.7 are the keys to derive the decay rates of the interpolation error of the GEIM Greedy algorithm. This is the purpose of the following theorem that is based on the inequality

$$\forall \varphi \in F,\ \|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1+\Lambda_n)\|\varphi - P_n(\varphi)\|_{\mathcal{X}} \leq (1+\Lambda_n)\tau_n,$$

together with lemmas 4.6 and 4.7 :

Theorem 4.10 (Convergence rates for GEIM).

1. *Assume that* $d_n \leq C_0 n^{-\alpha}$ *for any* $n \geq 1$*, then the interpolation error of the GEIM Greedy selection process satisfies for any* $\varphi \in F$ *the inequality* $\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1+\Lambda_n) C_0 \beta_n n^{-\alpha}$*, where the parameter* β_n *is defined as in lemma 4.6.*
2. *Assume that* $d_n \leq C_0 e^{-c_1 n^\alpha}$ *for* $n \geq 1$ *and* $C_0 \geq 1$*, then the interpolation error of the GEIM Greedy selection process satisfies for any* $\varphi \in F$ *the inequality* $\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq (1+\Lambda_n) C_0 \beta_n e^{-c_2 n^\alpha}$*, where* β_n *and* c_2 *are defined as in lemma 4.7.*

Then, similarly as in the previous section, we derive

Corollary 4.11. *If* (Λ_n) *is a monotonically increasing sequence, using corollary 4.8, the following decay rates in the generalized interpolation error can be derived:*

- *For any* $\varphi \in F$*, if* $d_n \leq C_0 n^{-\alpha}$ *for any* $n \geq 1$*, then the interpolation error of the GEIM Greedy selection process can be bounded as*

$$\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq \begin{cases} 2C_0\left(1+\dfrac{1}{\eta}\right)(1+\Lambda_1), & \text{if } n=1. \\ C_0 2^{3\alpha+1}\dfrac{(1+\Lambda_n)^3}{\eta^2} n^{-\alpha}, & \text{if } n \geq 2. \end{cases}$$

- *For any* $\varphi \in F$*, if* $d_n \leq C_0 e^{-c_1 n^\alpha}$ *for* $n \geq 1$ *and* $C_0 \geq 1$*, then the interpolation error of the GEIM Greedy selection process can be bounded as*

$$\|\varphi - \mathcal{J}_n[\varphi]\|_{\mathcal{X}} \leq \begin{cases} 2C_0\left(1+\dfrac{1}{\eta}\right)(1+\Lambda_1), & \text{if } n=1, \\ C_0\sqrt{2}\dfrac{(1+\Lambda_n)^2}{\eta} e^{-c_2 n^\alpha}, & \text{if } n \geq 2, \end{cases}$$

where $c_2 = 2^{-2\alpha-1}$.

5. Conclusion. Under the hypothesis of polynomial or exponential decay of the Kolmogorov n-width $d_n(F,\mathcal{X})$, it has been proven that the convergence rates of the interpolation error in GEIM are nearly-optimal and that the lack of optimality comes from the Lebesgue constant of the method that, depending on the case, impacts of the convergence by adding terms of order $\mathcal{O}(\Lambda_n^2)$ or $\mathcal{O}(\Lambda_n^3)$.

Given the fact that, for reasonable enough dictionaries Σ, it has been observed in practical applications that (Λ_n) is linear in the worst case scenario (see [1], [8], [15], [13]), our results prove that a decay of order $\mathcal{O}(n^{-3})$ in $d_n(F,\mathcal{X})$ should be enough to ensure the convergence of the interpolation errors of GEIM.

Appendix A. Proof of Theorem 3.5. We begin by recalling a preliminary lemma for matrices that is proven in [6].

LEMMA A.1. *Let $G=(g_{i,j})$ be a $K\times K$ lower triangular matrix with rows $\boldsymbol{g_1},\dots,\boldsymbol{g_K}$, W be any m dimensional subspace of $\mathbb{R}^K$, and P be the orthogonal projection of $\mathbb{R}^K$ onto W. Then,*

$$\prod_{i=1}^{K} g_{i,i}^2 \leq \left\{ \frac{1}{m}\sum_{i=1}^{K}\|P\boldsymbol{g_i}\|_{\ell_2}^2 \right\}^m \left\{ \frac{1}{K-m}\sum_{i=1}^{K}\|\boldsymbol{g_i}-P\boldsymbol{g_i}\|_{\ell_2}^2 \right\}^{K-m} \tag{A.1}$$

where $\|.\|_{\ell_2}$ is the euclidean norm of a vector in $\mathbb{R}^K$.

For the proof of theorem 3.5, we consider a lower triangular matrix $A=(a_{i,j})_{i,j=1}^{\infty}$ defined in the following way. For each $j=1,\dots,$ we let $\lambda_j\in\mathcal{L}(\mathcal{X})$ be the linear functional of norm one that satisfies:

$$(i)\ \lambda_j(X_j)=0, \qquad (ii)\ \lambda_j(\varphi_j)=\operatorname{dist}(\varphi_j,X_j), \tag{A.2}$$

where $X_j=\operatorname{span}\{\varphi_0,\dots,\varphi_{j-1}\}$, $j=1,2,\dots,$ is the interpolating space given by the greedy algorithm of GEIM. The existence of such a functional is a consequence of the Hahn-Banach theorem. We let A be the matrix with entries

$$a_{i,j}=\lambda_j(\varphi_i).$$

LEMMA A.2. *The matrix A has the following properties:*
P1: *The diagonal elements of A satisfy $\gamma_n\tau_n \leq a_{n,n}\leq \tau_n$*
P2: *For every $j<i$, one has: $|a_{i,j}|\leq \operatorname{dist}(\varphi_i,X_j)\leq\tau_j$.*
P3: *For every $j>i$, $a_{i,j}=0$.*
Proof.

P1: We have

$$a_{j,j}=\lambda_j(\varphi_j)=\operatorname{dist}(\varphi_j,X_j)=\|\varphi_j-P_j(\varphi_j)\|_{\mathcal{X}}\leq\max_{\varphi\in F}\|\varphi-P_j(\varphi)\|_{\mathcal{X}}=\tau_j.$$

Lemma 3.1 directly yields the second part of the inequality: $a_{j,j}\geq\gamma_j\tau_j$.

P2: For any $j<i$ and any $g\in X_j$, we have

$$|a_{i,j}|=|\lambda_j(\varphi_i)|=|\lambda_j(\varphi_i-g)|\leq\|\lambda_j\|_{\mathcal{L}(\mathcal{X})}\|\varphi_j-g\|_{\mathcal{X}},$$

where we have used the fact that $\lambda_j(g)=0$ because $g\in X_j$. Therefore, since $\|\lambda_j\|_{\mathcal{L}(\mathcal{X})}=1$,

$$|a_{i,j}|\leq\|\varphi_j-g\|_{\mathcal{X}},\quad\forall\, g\in X_j,$$

hence

$$|a_{i,j}|\leq\|\varphi_i-P_j(\varphi_i)\|_{\mathcal{X}}\leq\tau_j.$$

P3: Clearly, for $j > i$, $a_{i,j} = \lambda_j(\varphi_i) = 0$ because $\varphi_i \in X_j$ in this case.
$\square$

We can now prove theorem 3.5, i.e.:

THEOREM A.3. *For any $N \geq 0$, consider a weak Greedy algorithm with the property of lemma 3.1 and constant γ_N. We have the following inequalities between τ_N and d_N: for any $K \geq 1,\ 1 \leq m < K$*

$$\prod_{i=1}^{K} \tau_{N+i}^2 \leq \frac{1}{\prod_{i=1}^{K} \gamma_{N+i}^2} 2^K K^{K-m} \left(\sum_{i=1}^{K} \tau_{N+i}^2 \right)^m d_m^{2(K-m)} \tag{A.3}$$

Proof. We consider the $K \times K$ matrix G which is formed by the rows and columns of A with indices from $\{N+1, \ldots, N+K\}$. Let Y_m be the Kolmogorov subspace of $\mathcal{X}$ for which $\mathrm{dist}(F, Y_m) = d_m(F, \mathcal{X})$. For each i, there is an element $h_i \in Y_m$ such that

$$\|\varphi_i - h_i\|_{\mathcal{X}} = \mathrm{dist}(\varphi_i, Y_m) \leq d_m(F, \mathcal{X})$$

and therefore

$$|\lambda_j(\varphi_i) - \lambda_j(h_i)| = |\lambda_j(\varphi_i - h_i)| \leq \|\lambda_j\|_{\mathcal{L}(\mathcal{X})} \|\varphi_i - h_i\|_{\mathcal{X}} \leq d_m(F, \mathcal{X}). \tag{A.4}$$

We now consider the vectors $(\lambda_{N+1}(h), \ldots, \lambda_{N+K}(h))$, $h \in X_m$. They span a space $W \subset \mathbb{R}^K$ of dimension $\leq m$. We assume that $\dim(W) = m$ (a slight notational adjustment has to be made if $\dim(W) < m$). It follows from (A.4) that each row $\boldsymbol{g_i}$ of G can be approximated by a vector from W in the ℓ_∞ norm to accuracy d_m, and therefore in the ℓ_2 norm to accuracy $\sqrt{K} d_m$. Let P be the orthogonal projection of $\mathbb{R}^K$ onto W. Hence, we have

$$\|\boldsymbol{g_i} - P\boldsymbol{g_i}\|_{\ell_2} \leq \sqrt{K} d_m, \quad i = 1, \ldots, K. \tag{A.5}$$

It also follows from property P2 that

$$\|P\boldsymbol{g_i}\|_{\ell_2} \leq \|\boldsymbol{g_i}\|_{\ell_2} \leq \left\{ \sum_{j=1}^{i} \tau_{N+j}^2 \right\}^{1/2},$$

and therefore

$$\sum_{i=1}^{K} \|P\boldsymbol{g_i}\|_{\ell_2}^2 \leq \sum_{i=1}^{K} \sum_{j=1}^{i} \tau_{N+j}^2 \leq K \sum_{i=1}^{K} \tau_{N+i}^2. \tag{A.6}$$

Next, we apply lemma A.1 for this G and W and use property P1 and estimates (A.5) and (A.6) to derive

$$\begin{aligned}
\prod_{i=1}^{K} \gamma_{N+i}^2 \tau_{N+i}^2 &\leq \left\{ \frac{K}{m} \sum_{i=1}^{K} \tau_{N+i}^2 \right\}^m \left\{ \frac{K^2}{K-m} d_m^2 \right\}^{K-m} \\
&= K^{K-m} \left(\frac{K}{m} \right)^m \left(\frac{K}{K-m} \right)^{K-m} \left\{ \sum_{i=1}^{K} \tau_{N+i}^2 \right\}^m d_m^{2(K-m)} \\
&\leq 2^K K^{K-m} \left\{ \sum_{i=1}^{K} \tau_{N+i}^2 \right\}^m d_m^{2(K-m)},
\end{aligned}$$

where we have used the fact that $x^{-x}(1-x^{x-1}) \leq 2$ for $0<x<1$. This completes the proof.
□

Appendix B. Proof of Theorem 4.3.

In this section, $\mathcal{X}$ is a Hilbert space. We will denote by $(\varphi_n^*)_{n\geq 0}$ the orthonormal system obtained from $(\varphi_n)_{n\geq 0}$ by Gram-Schmidt orthonormalisation. It follows that the orthogonal projector P_n from $\mathcal{X}$ onto X_n is given by

$$P_n\varphi = \sum_{i=0}^{n-1} (\varphi, \varphi_i^*)_{\mathcal{X}} \varphi_i^*, \quad n = 1, 2, \ldots$$

and, in particular,

$$\varphi_n = P_{n+1}\varphi_n = \sum_{j=0}^{n} a_{n,j}\varphi_j^*, \quad a_{n,j} = (\varphi_n, \varphi_j^*)_{\mathcal{X}},\ j \leq n.$$

There is no loss of generality in assuming that the infinite dimensional Hilbert space $\mathcal{X}$ is $\ell_2(\mathbb{N})$ and that $\varphi_j^* = e_j$, where e_j is the vector with a one in the coordinate indexed by j and is zero in all other coordinates, i.e. $(e_j)_i = \delta_{j,i}$.

In a similar manner as in the Banach space case, we associate with the greedy procedure of GEIM the lower triangular matrix:

$$A := (a_{i,j})_{i,j=0}^{\infty}, \quad a_{i,j} := 1,\ j > i.$$

This matrix incorporates all the information about the weak greedy algorithm on F. The following two properties characterize any lower triangular matrix A generated by such a greedy algorithm.

Lemma B.1. *The matrix A has the following two properties:*

P1: *The diagonal elements of A satisfy $\gamma_n\tau_n \leq |a_{n,n}| \leq \tau_n$.*

P2: *For every $m \geq n$, one has $\sum\limits_{j=n}^{m} a_{m,j}^2 \leq \tau_n^2$.*

Proof.

P1: For any $n \geq 1$, since $\varphi_n - P_n\varphi_n = a_{n,n}\varphi_n^*$, it follows that For any $n \geq 1$, $|a_{n,n}| = \|\varphi_n - P_n\varphi_n\| \leq \tau_n$. The fact that $|a_{n,n}| \geq \gamma_n\tau_n$ directly follows from lemma 3.1.

P2: For $m \geq n$, $\sum\limits_{j=n}^{m} a_{m,j}^2 = \|\varphi_m - P_n\varphi_m\|^2 \leq \max\limits_{\varphi\in F}\|\varphi - P_n\varphi\|^2 = \tau_n^2$.

□

We can now prove theorem 4.3, i.e.

Theorem B.2. *For any $N \geq 0$, consider a weak Greedy algorithm with the property of lemma 3.1 and constant γ_N. We have the following inequalities between τ_N and d_N: for any $K \geq 1,\ 1 \leq m < K$*

$$\prod_{i=1}^{K} \tau_{N+i}^2 \leq \frac{1}{\prod\limits_{i=1}^{K} \gamma_{N+i}^2} \left(\frac{K}{m}\right)^m \left(\frac{K}{K-m}\right)^{K-m} \tau_{N+1}^{2m} d_m^{2(K-m)}.$$

Proof. We consider the $K \times K$ matrix $G = (g_{i,j})$ which is formed by the rows and columns of A with indices from $\{N+1, \dots, N+K\}$. Each row $\boldsymbol{g_i}$ is the restriction of φ_{N+i} to the coordinates $N+1, \dots, N+K$. Let Y_m be the Kolmogorov subspace of $\mathcal{X}$ for which $\mathrm{dist}(F, Y_m) = d_m(F, \mathcal{X})$. Then, $\mathrm{dist}(\varphi_{N+i}, Y_m) \leq d_m,\ i = 1, \dots, K$. Let $\tilde{W}$ be the linear subspace which is the restriction of Y_m to the coordinates $N+1, \dots, N+K$. In general, $\dim(\tilde{W}) \leq m$. Let W be an m dimensional space, $W \subset \mathrm{span}\{e_{N+1}, \dots, e_{N+K}\}$, such that $\tilde{W} \subset W$ and P and $\tilde{P}$ are the projections in $\mathbb{R}^K$ onto W and $\tilde{W}$, respectively. Clearly,

$$\|P\boldsymbol{g_i}\|_{\ell_2} \leq \|\boldsymbol{g_i}\|_{\ell_2} \leq \tau_{N+1}, \quad i = 1, \dots, K, \tag{B.1}$$

where we have used property P2 in the last inequality. Note that

$$\|\boldsymbol{g_i} - P\boldsymbol{g_i}\|_{\ell_2} \leq \|\boldsymbol{g_i} - \tilde{P}\boldsymbol{g_i}\|_{\ell_2} = \mathrm{dist}\left(\boldsymbol{g_i}, \tilde{W}\right) \leq \mathrm{dist}\left(\varphi_{N+i}, Y_m\right) \leq d_m, \quad i = 1, \dots, K. \tag{B.2}$$

It follows from property P1 that

$$\prod_{i=1}^{K} |a_{N+i,N+i}|^2 \geq \prod_{i=1}^{K} \gamma_{N+i}^2 \tau_{N+i}^2. \tag{B.3}$$

We now apply lemma A.1 for this G and W, and use estimates (B.1), (B.2) and (B.3) to derive the result.
$\square$

REFERENCES

[1] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera, *An empirical interpolation method: Application to efficient reduced-basis discretization of partial differential equations.*, C. R. Acad. Sci. Paris, Série I., 339 (2004), pp. 667–672.

[2] P. Binev, A. Cohen, W. Dahmen, R.A. DeVore, G. Petrova, and P. Wojtaszczyk, *Convergence rates for greedy algorithms in reduced basis methods.*, SIAM J. Math. Anal., 43 (2011), pp. 1457–1472.

[3] A. Buffa, Y. Maday, A.T. Patera, C. Prud'Homme, and G. Turinici, *A priori convergence of the greedy algorithm for the parametrized reduced basis method*, ESAIM: Mathematical Modelling and Numerical Analysis, 46 (3) (2012), pp. 595–603.

[4] A. Cohen and R. DeVore, *Kolmogorov widths under holomorphoc mappings*, Submitted, (2014).

[5] P. J. Davis, *Interpolation and Approximation*, Blaisdell Publishing Company, 1963.

[6] R. A. DeVore, G. Petrova, and P. Wojtaszczyk, *Greedy algorithms for reduced bases in Banach spaces*, Constructive Approximation, (2012), pp. 1–12.

[7] M.S. Floater and K. Hormann, *Barycentric rational interpolation with no poles and high rates of approximation*, Numer. Math., 107 (2007), pp. 315–331.

[8] M.A. Grepl, Y. Maday, N.C. Nguyen, and A.T. Patera, *Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations.*, ESAIM, Math. Model. Numer. Anal., 41(3) (2007), pp. 575–605.

[9] J. Kleijnen and W. van Beers, *Robustness of kriging when interpolating in random simulation with heterogeneous variances: Some experiments*, European Journal of Operational Research, 165 (2005), pp. 826 – 834.

[10] A. Kolmogoroff, *Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse*, Annals of Mathematics, 37 (1936), pp. 107–110.

[11] H. Liu and S. Maghsoodloo, *Simulation optimization based on taylor kriging and evolutionary algorithm*, Applied Soft Computing, 11 (2011), pp. 3451 – 3462.

[12] Y. Maday and O. Mula, *A generalized empirical interpolation method: Application of reduced basis techniques to data assimilation*, in Analysis and Numerics of Partial Differential Equations, Franco Brezzi, Piero Colli Franzone, Ugo Gianazza, and Gianni Gilardi, eds., vol. 4 of Springer INdAM Series, Springer Milan, 2013, pp. 221–235.

[13] Y. Maday, O. Mula, A.T. Patera, and M. Yano, *The generalized Empirical Interpolation Method: stability theory on Hilbert spaces with an application to the Stokes equation*, Submitted, – (2014), pp. –.
[14] Y. Maday, O. Mula, and 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.
[15] Y. Maday, N.C. Nguyen, A.T. Patera, and G.S.H. Pau, *A general multipurpose interpolation procedure: the magic points*, Comm. Pure Appl. Anal., 8(1) (2009), pp. 383–404.
[16] Y. Maday, A.T. Patera, J.D. Penn, and M. Yano, *A Parametrized-Background Data-Weak approach to variational data assimilation: formulation, analysis and application to acoustics*, Int. J. Num. Meth. Eng., (2014). Accepted.
[17] I. H. Sloan, *Nonpolynomial interpolation*, J. Approx. Theory, 39 (1983), pp. 97 – 117.