HAL
open science

Convergence results on greedy algorithms for high-dimensional eigenvalue problems

Virginie Ehrlacher

To cite this version:

Virginie Ehrlacher. Convergence results on greedy algorithms for high-dimensional eigenvalue problems. ESAIM: Proceedings and Surveys, 2014, 45, pp.148 - 157. 10.1051/proc/201445015 . hal-01435976

HAL Id: hal-01435976

https://hal.science/hal-01435976v1

Submitted on 16 Jan 2017

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.

ESAIM: PROCEEDINGS AND SURVEYS, September 2014, Vol. 45, p. 148-157
J.-S. Dhersin, Editor

CONVERGENCE RESULTS ON GREEDY ALGORITHMS FOR HIGH-DIMENSIONAL EIGENVALUE PROBLEMS *

Virginie Ehrlacher [1]

Abstract. In this paper, we present two greedy algorithms for the computation of the lowest eigenvalue (and an associated eigenvector) of a high-dimensional eigenvalue problem, which have been introduced and analyzed recently in a joint work with Eric Cancès and Tony Lelièvre [1]. The performance of our algorithms is illustrated on toy numerical test cases, and compared with that of another greedy algorithm for eigenvalue problems introduced by Ammar and Chinesta [13].

Résumé. Dans ce document, nous présentons deux algorithmes gloutons for le calcul de la plus petite valeur propre (et d'un vecteur propre associé) d'un problème aux valeurs propres en grande dimension, qui ont été récemment introduits et analysés dans un travail commun avec Eric Cancès et Tony Lelièvre [1]. Le comportement numérique de ces algorithmes est illustré sur de petits cas tests, et comparé à celui d'un autre algorithme glouton proposé antérieurement par Ammar et Chinesta [13].

Introduction

High dimensional problems are encountered in many application fields, among which electronic structure calculations, molecular dynamics, uncertainty quantification, multiscale homogenization, and mathematical finance. The numerical simulation of these problems, which requires specific approaches due to the so-called *curse of dimensionality* [2], has fostered the development of a wide variety of new numerical methods and algorithms, such as sparse grids [3, 4], reduced bases [5], sparse tensor products [6], and adaptive polynomial approximations [7].

In this article, we focus on an approach introduced by Ladevèze [8], Chinesta [9], Nouy [10] and coauthors in different contexts, relying on the use of *greedy algorithms* [11]. This class of methods is also called *Progressive Generalized Decomposition* [12] in the literature.

The idea of these methods consist in approximating a function $u(x_1, \cdots, x_d)$ depending on a possibly very large number of variables $x_1, \cdots, x_d$ as a sum of so-called *tensor product* functions,

$$u(x_1, \cdots, x_d) \approx \sum_{k \geq 1} r_k^{(1)}(x_1) \cdots r_k^{(d)}(x_d),$$

where each of these tensor product functions appearing in the above sum is computed in an iterative way as *the best next tensor product* in a sense which depends on the nature of the problem *u is solution of* .

---

* *This work has been done while the author was a long-term visitor at IPAM (UCLA).*

[1] Université Paris Est, CERMICS, Projet MICMAC, Ecole des Ponts ParisTech - INRIA, 6 & 8 avenue Blaise Pascal, 77455 Marne-la-Vallée Cedex 2, France; e-mail: `ehrlachv@cermics.enpc.fr`

© EDP Sciences, SMAI 2014

Article published online by EDP Sciences and available at http://www.esaim-proc.org or http://dx.doi.org/10.1051/proc/201445015

Greedy algorithms have been extensively studied in the framework of convex unconstrained minimization problems, in other words when u is the solution of a problem of the form

$$u = \operatorname*{argmin}_{v \in V} \mathcal{E}(v), \tag{1}$$

where V is a Hilbert space of functions depending on the d variables $x_1, \cdots, x_d$, and $\mathcal{E}$ is a convex energy functional [14–16]. However, the analysis of such algorithms for other kinds of problems is less advanced [12]. We refer to [17] for a review of the mathematical issues arising in the application of greedy algorithms to non-symmetric linear problems. To our knowledge, the literature on greedy algorithms for eigenvalue problems is very limited. Penalized formulations of constrained minimization problems enable one to recover the structure of unconstrained minimization problems and to benefit from the sound theoretical framework existing in this case [14, 22]. The only reference we are aware of about greedy algorithms for eigenvalue problems without the use of a penalized formulation is an article by Ammar and Chinesta [13], in which the authors propose a greedy algorithm to compute the lowest eigenstate of a bounded from below self-adjoint operator, and apply it to electronic structure calculation. No analysis for this algorithm is given though.

In this paper, we present two new greedy algorithms for the computation of the lowest eigenstate of high-dimensional eigenvalue problems which were introduced and analyzed recently in [1]. We present the main theoretical results proved in the latter paper and illustrate the numerical behaviour of the algorithms on some toy numerical test cases. The outline of the paper is the following: we first introduce some notation and assumptions in Section 1; the description of the greedy algorithms and the associated theoretical convergence results are detailed in Section 2. We refer the reader to [1] for the detailed proofs of these results and an exhaustive description of the implementation of these algorithms in practice. The last section 3 contains some toy numerical tests for illustration.

1. Preliminaries

1.1. Notation and main assumptions

Let us consider two Hilbert spaces V and H, endowed respectively with the scalar products $\langle \cdot, \cdot \rangle_V$ and $\langle \cdot, \cdot \rangle$, such that, unless it is otherwise stated,

(HV) the embedding $V \hookrightarrow H$ is dense and compact.

The associated norms are denoted respectively by $\| \cdot \|_V$ and $\| \cdot \|$. Let us recall that it follows from (HV) that the weak convergence in V implies the strong convergence in H.

Let $a : V \times V \to \mathbb{R}$ be a symmetric continuous bilinear form on $V \times V$ such that

(HA) $\exists \gamma, \nu > 0$, such that $\forall v \in V,\ a(v, v) \geq \gamma \|v\|_V^2 - \nu \|v\|^2$.

The bilinear form $\langle \cdot, \cdot \rangle_a$, defined by

$$\forall v, w \in V, \quad \langle v, w \rangle_a := a(v, w) + \nu \langle v, w \rangle, \tag{2}$$

is a scalar product on V, whose associated norm, denoted by $\| \cdot \|_a$, is equivalent to the norm $\| \cdot \|_V$. Besides, we can also assume without loss of generality that the constant ν is chosen so that for all $v \in V$, $\|v\|_a \geq \|v\|$.

It is well-known (see e.g. [18]) that, under the above assumptions (namely (HA) and (HV)), there exists a sequence $(\psi_p, \mu_p)_{p \in \mathbb{N}^*}$ of solutions to the elliptic eigenvalue problem

$$\begin{cases} \text{find } (\psi, \mu) \in V \times \mathbb{R} \text{ such that } \|\psi\| = 1 \text{ and} \\ \forall v \in V,\ a(\psi, v) = \mu \langle \psi, v \rangle \end{cases} \tag{3}$$

such that $(\mu_p)_{p \in \mathbb{N}^*}$ forms a non-decreasing sequence of real numbers going to infinity and $(\psi_p)_{p \in \mathbb{N}^*}$ is an orthonormal basis of H. We focus here on the computation of μ_1, the lowest eigenvalue of $a(\cdot, \cdot)$, and of an associated H-normalized eigenvector. Let us note that, from (HA), for all $p \in \mathbb{N}^*$, $\mu_p + \nu > 0$.

In the case when the embedding $V \hookrightarrow H$ is dense but not compact (i.e. when (HV) does not hold), the spectrum of the unique self-adjoint operator A on H with form domain V associated with the quadratic form $a(\cdot,\cdot)$ can be purely continuous; in this case, (3) has no solution. However, if A has at least one discrete eigenvalue located below the minimum of its essential spectrum, convergence results for the second algorithm we propose can be established. This is the object of Proposition 2.1.

Definition 1.1. *A set $\Sigma \subset V$ is called a* dictionary *of V if Σ satisfies the following three conditions:*

(HΣ1): *Σ is a non-empty cone, i.e. $0 \in \Sigma$ and for all $(z,t) \in \Sigma \times \mathbb{R}$, $tz \in \Sigma$;*
(HΣ2): *Σ is weakly closed in V;*
(HΣ3): $\mathrm{Span}(\Sigma)$ *is dense in V.*

In practical applications for high-dimensional eigenvalue problems, the set Σ is typically an appropriate set of tensor formats used to perform the greedy algorithms presented in Section 2.1. We also denote by

$$\Sigma^* := \Sigma \setminus \{0\}. \tag{4}$$

1.2. Prototypical example

Let us present a prototypical example of the high-dimensional eigenvalue problems we have in mind, along with possible dictionaries.

Let $\mathcal{X}_1, \ldots, \mathcal{X}_d$ be bounded regular domains of $\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_d}$ respectively. Let $V = H_0^1(\mathcal{X}_1 \times \cdots \times \mathcal{X}_d)$ and $H = L^2(\mathcal{X}_1 \times \cdots \times \mathcal{X}_d)$. It follows from the Rellich-Kondrachov theorem that these spaces satisfy assumption (HV). Let $b : \mathcal{X}_1 \times \cdots \times \mathcal{X}_d \to \mathbb{R}$ be a measurable real-valued function such that

$$\exists \beta, B > 0, \text{ such that } \beta \leq b(x_1, \ldots, x_d) \leq B, \text{ for a.e. } (x_1, \ldots, x_d) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_d.$$

Besides, let $W \in L^q(\mathcal{X}_1 \times \cdots \times \mathcal{X}_d)$ with $q = 2$ if $m \leq 3$, and $q > m/2$ for $m \geq 4$ where $m := m_1 + \cdots + m_d$. A prototypical example of a continuous symmetric bilinear form $a : V \times V \to \mathbb{R}$ satisfying (HA) is

$$\forall v, w \in V, \quad a(v,w) := \int_{\mathcal{X}_1 \times \cdots \times \mathcal{X}_d} \left(b \nabla v \cdot \nabla w + W v w \right). \tag{5}$$

In this particular case, the eigenvalue problem (3) also reads

$$\begin{cases} \text{find } (\psi, \mu) \in H_0^1(\mathcal{X}_1 \times \cdots \times \mathcal{X}_d) \times \mathbb{R} \text{ such that } \|\psi\|_{L^2(\mathcal{X}_1 \times \cdots \times \mathcal{X}_d)} = 1 \text{ and} \\ -\mathrm{div}\,(b \nabla \psi) + W \psi = \mu \psi \text{ in } \mathcal{D}'(\mathcal{X}_1 \times \cdots \times \mathcal{X}_d). \end{cases}$$

For all $1 \leq j \leq d$, we denote by $V_j := H_0^1(\mathcal{X}_j)$. Some examples of dictionaries Σ based on different tensor formats satisfying (HΣ1), (HΣ2) and (HΣ3) are the set of rank-1 tensor-product functions

$$\Sigma^{\otimes} := \left\{ r^{(1)} \otimes \cdots \otimes r^{(d)} \mid \forall 1 \leq j \leq d,\ r^{(j)} \in V_j \right\}, \tag{6}$$

as well as other tensor formats [6, 19], for instance the sets of rank-R Tucker, rank-R Tensor Train, or rank-R Tensor Chain functions, with $R \in \mathbb{N}^*$.

2. Greedy algorithms for eigenvalue problems

In the rest of the paper, we present two different greedy algorithms to compute an eigenpair associated to the lowest eigenvalue of the elliptic eigenvalue problem (3).

The first one relies on the minimization of the Rayleigh quotient of $a(\cdot,\cdot)$ and is introduced in Section 2.1.1. The second one, presented in Section 2.1.2, is based on the use of a residual for problem (3). We recall the algorithm introduced in [13] in Section 2.1.3. Orthogonal versions of these algorithms are defined in Section 2.1.4.

Section 2.2 contains our main convergence results. The choice of a good initial guess for all these algorithms is discussed in [1] and is not detailed here for the sake of brevity. For a detailed proof of the results stated in this section, we refer the reader to [1].

For all $v \in V$, we denote by

$$\mathcal{J}(v) := \begin{cases} \frac{a(v,v)}{\|v\|^2} & \text{if } v \neq 0, \\ +\infty & \text{if } v = 0, \end{cases}$$

the Rayleigh quotient associated to (3), and

$$\lambda_\Sigma := \inf_{z\in\Sigma} \mathcal{J}(z) = \inf_{z\in\Sigma^*} \frac{a(z,z)}{\|z\|^2}.$$

Note that, since $\Sigma \subset V$, $\lambda_\Sigma \geq \mu_1 = \inf_{v\in V} \mathcal{J}(v)$.

2.1. Description of the algorithms

2.1.1. *Pure Rayleigh Greedy Algorithm*

The following algorithm, called hereafter the *Pure Rayleigh Greedy Algorithm* (PRaGA) algorithm, is inspired from the Pure Greegy Algorithm for convex minimization problems (see [14, 15] for instance).

Pure Rayleigh Greedy Algorithm (PRaGA):

- **Initialization:** choose an initial guess $u_0 \in V$ such that $\|u_0\| = 1$ and such that $\lambda_0 := a(u_0, u_0) < \lambda_\Sigma$;
- **Iterate on** $n \geq 1$**:** find $z_n \in \Sigma$ such that

$$z_n \in \operatorname*{argmin}_{z\in\Sigma} \mathcal{J}(u_{n-1} + z), \tag{7}$$

and set $u_n := \frac{u_{n-1}+z_n}{\|u_{n-1}+z_n\|}$ and $\lambda_n := a(u_n, u_n)$.

The choice of an initial guess $u_0 \in V$ satisfying $\|u_0\| = 1$ and $a(u_0, u_0) \leq \lambda_\Sigma$ is discussed [1].

The following lemma holds, stating that the iterations of the PRaGA are well-defined:

Lemma 2.1. *Let V and H be separable Hilbert spaces satisfying (HV), Σ a dictionary of V and $a : V \times V \to \mathbb{R}$ a symmetric continuous bilinear form satisfying (HA). Then, all the iterations of the PRaGA algorithm are well-defined in the sense that for all $n \in \mathbb{N}^*$, there exists at least one solution to the minimization problem (7). Besides, the sequence $(\lambda_n)_{n\in\mathbb{N}^*}$ is non-increasing.*

2.1.2. *Pure Residual Greedy Algorithm*

The *Pure Residual Greedy Algorithm* (PReGA) we propose is based on the use of a residual for problem (3).

Pure Residual Greedy Algorithm (PReGA):

- **Initialization:** choose an initial guess $u_0 \in V$ such that $\|u_0\| = 1$ and let $\lambda_0 := a(u_0, u_0)$;
- **Iterate on** $n \geq 1$**:** find $z_n \in \Sigma$ such that

$$z_n \in \operatorname*{argmin}_{z\in\Sigma} \frac{1}{2}\|u_{n-1} + z\|_a^2 - (\lambda_{n-1} + \nu)\langle u_{n-1}, z\rangle, \tag{8}$$

and set $u_n := \frac{u_{n-1}+z_n}{\|u_{n-1}+z_n\|}$ and $\lambda_n := a(u_n, u_n)$.

The denomination *Residual* can be justified as follows: it is easy to check that for all $n \in \mathbb{N}^*$, the minimization problem (8) is equivalent to the minimization problem

$$\text{find } z_n \in \Sigma \text{ such that } z_n \in \operatorname*{argmin}_{z\in\Sigma} \frac{1}{2}\|R_{n-1} - z\|_a^2, \tag{9}$$

where $R_{n-1} \in V$ is the Riesz representant in V of the linear form $l_{n-1} : v \in V \mapsto \lambda_{n-1}\langle u_{n-1}, v\rangle - a(u_{n-1}, v)$. In other words, R_{n-1} is the unique element in V such that

$$\forall v \in V, \quad \langle R_{n-1}, v\rangle_a = \lambda_{n-1}\langle u_{n-1}, v\rangle - a(u_{n-1}, v).$$

The linear form l_{n-1} can indeed be seen as a residual for (3) since $l_{n-1} = 0$ if and only if λ_{n-1} is an eigenvalue of $a(\cdot,\cdot)$ and u_{n-1} an associated H-normalized eigenvector.

Let us point out that, in order to carry out the PReGA in practice, one needs to know the value of a constant ν ensuring (HA), whereas this is not needed for the PRaGA, neither for the algorithm (PEGA) introduced in [13] and considered in the next section.

Lemma 2.2. *Let V and H be separable Hilbert spaces such that the embedding $V \hookrightarrow H$ is dense, Σ a dictionary of V and $a : V \times V \to \mathbb{R}$ a symmetric continuous bilinear form satisfying (HA). Then, all the iterations of the PReGA algorithm are well-defined in the sense that for all $n \in \mathbb{N}^*$, there exists at least one solution to the minimization problem (8).*

2.1.3. *Pure Explicit Greedy Algorithm*

The above two algorithms are new, at least to our knowledge. In this section, we describe the algorithm already proposed in [13], which we call in the rest of the paper the *Pure Explicit Greedy Algorithm* (PEGA).

Unlike the above two algorithms, the PEGA is not defined for general dictionaries Σ satisfying (HΣ1), (HΣ2) and (HΣ3). We need to assume in addition that Σ is an embedded manifold in V. In this case, for all $z \in \Sigma$, we denote by $T_\Sigma(z)$ the tangent subspace to Σ at the point z in V.

Let us point out that, if Σ is an embedded manifold in V, for all $n \in \mathbb{N}^*$, the Euler equations associated to the minimization problems (7) and (8) respectively read:

$$\forall \delta z \in T_\Sigma(z_n), \quad a\left(u_{n-1} + z_n, \delta z\right) = \lambda_n \langle u_{n-1} + z_n, \delta z\rangle, \tag{10}$$

and

$$\forall \delta z \in T_\Sigma(z_n), \quad a\left(u_{n-1} + z_n, \delta z\right) + \nu\langle z_n, \delta z\rangle = \lambda_{n-1}\langle u_{n-1}, \delta z\rangle. \tag{11}$$

The PEGA consists in solving at each iteration $n \in \mathbb{N}^*$ of the greedy algorithm the following equation, which is of a similar form as the Euler equations (10) and (11) above,

$$\forall \delta z \in T_\Sigma(z_n), \quad a\left(u_{n-1} + z_n, \delta z\right) = \lambda_{n-1}\langle u_{n-1} + z_n, \delta z\rangle. \tag{12}$$

More precisely, the PEGA algorithm reads:

Pure Explicit Greedy Algorithm (PEGA):

- **Initialization:** choose an initial guess $u_0 \in V$ such that $\|u_0\| = 1$ and let $\lambda_0 := a(u_0, u_0)$;
- **Iterate for $n \geq 1$:** find $z_n \in \Sigma$ such that

$$\forall \delta z \in T_\Sigma(z_n), \quad a\left(u_{n-1} + z_n, \delta z\right) - \lambda_{n-1}\langle u_{n-1} + z_n, \delta z\rangle = 0, \tag{13}$$

  and set $u_n := \frac{u_{n-1}+z_n}{\|u_{n-1}+z_n\|}$ and $\lambda_n := a(u_n, u_n)$.

Notice that (13) is very similar to (10) except that λ_{n-1} is used instead of λ_n. It can be seen as an *explicit* version of the PRaGA, hence the name *Pure Explicit Greedy Algorithm* .

Note that it is not clear whether there always exists a solution z_n to (13), since (13) does not derive from a minimization problem, unlike the other two algorithms. We were unable to prove convergence results for the PEGA.

2.1.4. *Orthogonal algorithms*

We introduce here slightly modified versions of the PRaGA, PReGA and PEGA, inspired from the *Orthogonal Greedy Algorithm* for convex minimization problems (see for instance [15, 16]).

Orthogonal (Rayleigh, Residual or Explicit) Greedy Algorithm (ORaGA, OReGA and OEGA):

- **Initialization:** choose an initial guess $u_0 \in V$ such that $\|u_0\| = 1$ and let $\lambda_0 := a(u_0, u_0)$. For the ORaGA, we need to assume that $\lambda_0 := a(u_0, u_0) < \lambda_\Sigma$.
- **Iterate on $n \geq 1$:**
  - for the ORaGA: find $z_n \in \Sigma$ satisfying (7);
  - for the OReGA: find $z_n \in \Sigma$ satisfying (8);
  - for the OEGA: find $z_n \in \Sigma$ satisfying (13);

  find $\left(c_0^{(n)}, \ldots, c_n^{(n)}\right) \in \mathbb{R}^{n+1}$ such that

$$\left(c_0^{(n)}, \ldots, c_n^{(n)}\right) \in \underset{(c_0,\ldots,c_n)\in\mathbb{R}^{n+1}}{\operatorname{argmin}} \mathcal{J}\left(c_0 u_0 + c_1 z_1 + \cdots + c_n z_n\right), \tag{14}$$

  and set $u_n := \frac{c_0^{(n)} u_0 + c_1^{(n)} z_1 + \cdots + c_n^{(n)} z_n}{\|c_0^{(n)} u_0 + c_1^{(n)} z_1 + \cdots + c_n^{(n)} z_n\|}$; if $\langle u_{n-1}, u_n\rangle \leq 0$, set $u_n := -u_n$; set $\lambda_n := a(u_n, u_n)$.

Let us point out that the original algorithm proposed in [13] is the OEGA. Besides, for the three algorithms and all $n \in \mathbb{N}^*$, there always exists at least one solution to the minimization problems (14).

For the sake of brevity, we do not detail here how the problems (7), (8) and (13) are solved in practice and refer the reader to [1] for further details in the case when the dictionary Σ is the set of rank-1 tensor product functions.

The orthogonal versions of the greedy algorithms can be easily implemented from the pure versions: at any iteration $n \in \mathbb{N}^*$, only an additional step is performed, which consists in choosing an approximate eigenvector u_n as a linear combination of the elements $u_0, z_1, \ldots, z_n$ minimizing the Rayleigh quotient associated to the bilinear form $a(\cdot,\cdot)$. Since u_n is called to be the approximation of an eigenvector associated to the lowest eigenvalue of $a(\cdot,\cdot)$, which is a minimizer of the Rayleigh quotient on the Hilbert space V, this additional step is a natural extension of the Orthogonal Greedy Algorithm for the minimization of convex energy functionals [15].

2.2. Convergence results

2.2.1. *The infinite-dimensional case*

Theorem 2.1. *Let V and H be separable Hilbert spaces satisfying (HV), Σ a dictionary of V and $a : V \times V \to \mathbb{R}$ a symmetric continuous bilinear form satisfying (HA). The following properties hold for the PRaGA, ORaGA, PReGA and OReGA:*

(1) *All the iterations of the algorithms are well-defined.*
(2) *The sequence $(\lambda_n)_{n\in\mathbb{N}}$ is non-increasing and converges towards a limit λ which is an eigenvalue of $a(\cdot,\cdot)$ for the scalar product $\langle\cdot,\cdot\rangle$.*
(3) *The sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in V and any subsequence of $(u_n)_{n\in\mathbb{N}}$ which weakly converges in V also strongly converges in V towards an H-normalized eigenvector associated with λ. This implies in particular that*
$$d_a(u_n, F_\lambda) := \inf_{w\in F_\lambda} \|w - u_n\|_a \underset{n\to\infty}{\longrightarrow} 0,$$
where F_λ denotes the set of the H-normalized eigenvectors of $a(\cdot,\cdot)$ associated with λ.
(4) *If λ is a simple eigenvalue, then there exists an H-normalized eigenvector w_λ associated with λ such that the whole sequence $(u_n)_{n\in\mathbb{N}}$ converges to w_λ strongly in V.*

It may happen that $\lambda > \mu_1$, if the initial guess u_0 is not properly chosen. This point is discussed in more details in [1]. If λ is degenerate, it is not clear whether the whole sequence $(u_n)_{n\in\mathbb{N}}$ converges. We will see in Section 2.2.2 that in finite dimension, at least for the pure versions of these algorithms, the whole sequence

$(u_n)_{n\in\mathbb{N}}$ always converges towards an element $w_\lambda \in V$ which is an eigenvector of $a(\cdot,\cdot)$ associated with λ. However, λ may still be strictly greater than μ_1, even in this case. However, if the initial guess u_0 is chosen such that λ_0 is strictly lower any eigenvalue of $a(\cdot,\cdot)$ except μ_1, then $(\lambda_n)_{n\in\mathbb{N}}$ necessarily converges to μ_1.

In addition, for the PReGA and the OReGA, we can prove similar convergence results without assuming that the Hilbert space V is compactly embedded in H, provided that the self-adjoint operator A associated with the quadratic form $a(\cdot,\cdot)$ has at least one eigenvalue below the minimum of its essential spectrum.

Proposition 2.1. *Let V and H be separable Hilbert spaces such that the embedding $V \hookrightarrow H$ is dense (but not necessarily compact), Σ a dictionary of V, $a : V\times V \to \mathbb{R}$ a symmetric continuous bilinear form satisfying (HA), and A the self-adjoint operator on H associated to $a(\cdot,\cdot)$. Let us assume also that $\min\sigma(A) < \min\sigma_{\text{ess}}(A)$, where $\sigma(A)$ and $\sigma_{\text{ess}}(A)$ respectively denote the spectrum and the essential spectrum of A, and that the initial guess u_0 satisfies $\min\sigma(A) \le \lambda_0 := a(u_0,u_0) < \min\sigma_{\text{ess}}(A)$. Then, the following properties hold for the PReGA and the OReGA:*

(1) *All the iterations of the algorithms are well-defined.*
(2) *The sequence $(\lambda_n)_{n\in\mathbb{N}}$ is non-increasing and converges towards a limit λ which is an eigenvalue of $a(\cdot,\cdot)$ for the scalar product $\langle\cdot,\cdot\rangle$ such that $\lambda < \min\sigma_{\text{ess}}(A)$.*
(3) *The sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in V and any subsequence of $(u_n)_{n\in\mathbb{N}}$ which weakly converges in V also strongly converges in V towards an H-normalized eigenvector associated with λ. This implies in particular that*
$$d_a(u_n, F_\lambda) := \inf_{w\in F_\lambda} \|w-u_n\|_a \underset{n\to\infty}{\longrightarrow} 0,$$
where F_λ denotes the set of H-normalized eigenvectors of $a(\cdot,\cdot)$ associated with λ.
(4) *If λ is a simple eigenvalue, then there exists an H-normalized eigenvector w_λ associated with λ such that the whole sequence $(u_n)_{n\in\mathbb{N}}$ converges to w_λ strongly in V.*

2.2.2. *The finite-dimensional case*

From now on, for any differentiable function $f : V \to \mathbb{R}$, and all $v_0 \in V$, we denote by $f'(v_0)$ the derivative of the function f at the point $v_0 \in V$. More precisely, $f'(v_0) \in V'$ is the unique continuous linear form on V such that for all $v \in V$,

$$f(v) = f(v_0) + \langle f'(v_0), v-v_0\rangle_{V',V} + r(v), \text{ with } \lim_{\|v\|_a\to 0} \frac{r(v)}{\|v\|_a} = 0.$$

Besides, we define the injective norm on V' associated to Σ as follows:

$$\forall l \in V',\ \|l\|_* = \sup_{z\in\Sigma^*} \frac{\langle l, z\rangle_{V',V}}{\|z\|_a}.$$

In the rest of this section, we assume that V, hence H (since the embedding $V \hookrightarrow H$ is dense), are finite dimensional vector spaces. The convergence results below heavily rely on the Łojasiewicz inequality [20] and the ideas presented in [21] for the proof of convergence of gradient-based algorithms for the Hartree-Fock equations.

The Łojasiewicz inequality [20] reads as follows:

Lemma 2.3. *Let Ω be an open subset of the finite-dimensional Euclidean space V, and f an analytic real-valued function defined on Ω. Then, for each $v_0 \in \Omega$, there is a neighborhood $U \subset \Omega$ of v_0 and two constants $K \in \mathbb{R}_+$ and $\theta \in (0, 1/2]$ such that for all $v \in U$,*

$$|f(v) - f(v_0)|^{1-\theta} \le K\|f'(v)\|_*. \tag{15}$$

This inequality can be understood in this way: it can be easily proved in the case when v_0 is not a critical point of f. When v_0 is a non-degenerate critical point, i.e. when the Hessian of f at v_0 is invertible, then it

is easy to see that θ can be chosen to be equal to $\frac{1}{2}$ by using a simple Taylor expansion. Moreover, when v_0 is a degenerate critical point of f, the analyticity assumption ensures that there exists $N \in \mathbb{N}^*$ such that the Nth-order derivatives cannot vanish simultaneously, and the exponent θ can be chosen to be equal to $\frac{1}{N}$.

In our context, the following lemma can be proved [1].

Lemma 2.4. *Let V and H be finite-dimensional Euclidean spaces, $\Omega := \{v \in V,\ 1/2 < \|v\| < 3/2\}$, λ be an eigenvalue of the bilinear form $a(\cdot,\cdot)$ and F_λ be the set of the H-normalized eigenvectors of $a(\cdot,\cdot)$ associated to λ. Then, $\mathcal{J} : \Omega \to \mathbb{R}$ is analytic, and there exists $K \in \mathbb{R}_+$, $\theta \in (0, 1/2]$ and $\varepsilon > 0$ such that*

$$\text{for all } v \in \Omega \text{ such that } d(v, F_\lambda) := \inf_{w \in F_\lambda} \|v - w\| \leq \varepsilon, \quad |\mathcal{J}(v) - \lambda|^{1-\theta} \leq K \|\mathcal{J}'(v)\|_*. \tag{16}$$

Using this lemma, the following convergence rates in finite dimension can be obtained:

Theorem 2.2. *Let V and H be finite dimensional Euclidian spaces and $a : V \times V \to \mathbb{R}$ be a symmetric bilinear form. The following properties hold for both PRaGA and PReGA:*

(1) *the whole sequence $(u_n)_{n\in\mathbb{N}}$ strongly converges in V to some $w_\lambda \in F_\lambda$;*
(2) *the convergence rates are as follows, depending on the value of the parameter θ in (16):*
   - *if $\theta = 1/2$, there exists $C \in \mathbb{R}_+$ and $0 < \sigma < 1$ such that for all $n \in \mathbb{N}$,*

$$\|u_n - w_\lambda\|_a \leq C\sigma^n; \tag{17}$$

   - *if $\theta \in (0, 1/2)$, there exists $C \in \mathbb{R}_+$ such that for all $n \in \mathbb{N}^*$,*

$$\|u_n - w_\lambda\|_a \leq C n^{-\frac{\theta}{1-2\theta}}. \tag{18}$$

3. Numerical results

We present here some numerical results obtained with these algorithms (PRaGA, PReGA, PEGA and their orthogonal versions) on toy examples involving only two Hilbert spaces ($d = 2$). We refer the reader to [13] for numerical examples involving a larger number of variables. These basic numerical tests performed with small-dimensional matrices lead us to think that the greedy algorithms presented above converge in general towards the lowest eigenvalue of the bilinear form under consideration, except in pathological situations which are not likely to be encountered in practice.

In this simple example, we take $V = H = \mathbb{R}^{N_x \times N_y}$, $V_x = \mathbb{R}^{N_x}$ and $V_y = \mathbb{R}^{N_y}$ for some $N_x, N_y \in \mathbb{N}^*$ (here typically $N_x = N_y = 51$). Let $D^{1x}, D^{2x} \in \mathbb{R}^{N_x \times N_x}$ and $D^{1y}, D^{2y} \in \mathbb{R}^{N_y \times N_y}$ be (randomly chosen) symmetric definite positive matrices. We aim at computing the lowest eigenstate of the symmetric bilinear form

$$a(U, V) = \mathrm{Tr}\left(U^T (D^{1x} V D^{1y} + D^{2x} V D^{2y})\right),$$

or, in other words, of the symmetric fourth order tensor A defined by

$$\forall 1 \leq i, k \leq N_x,\ 1 \leq j, l \leq N_y, \quad A_{ij,kl} = D^{1x}_{ik} D^{1y}_{jl} + D^{2x}_{ik} D^{2y}_{jl}.$$

Let us denote by μ_1 the lowest eigenvalue of the tensor A, by I the identity operator, and by $P_{\mu_1} \in \mathcal{L}(\mathbb{R}^{N_x \times N_y})$ the orthogonal projector onto the eigenspace of A associated with μ_1. Figure 1 shows the decay of the error on the eigenvalues $\log_{10}(|\mu_1 - \lambda_n|)$ and of the error on the eigenvectors $\log_{10}(\|(I - P_{\mu_1})U_n\|_F)$, where $\|\cdot\|_F$ denotes the Frobenius norm of $\mathbb{R}^{N_x \times N_y}$, as a function of n for the three algorithms and their orthogonal versions.

These tests were performed with several matrices $D^{1x}, D^{1y}, D^{2x}, D^{2y}$, either drawn randomly or chosen such that the eigenspace associated with the lowest eigenvalue is of dimension greater than 1. In any case, the three greedy algorithms converge towards a particular eigenstate associated with the lowest eigenvalue of the tensor A.

Besides, the rate of convergence always seems to be exponential with respect to n. The error on the eigenvalues decays twice as fast as the error on the eigenvectors, as usual when dealing with the approximation of linear eigenvalue problems.

We observe that the PRaGA and PEGA have similar convergence properties with respect to the number of iterations n. The behaviour of the PReGA strongly depends on the value ν chosen in (HA): the larger ν, the slower the convergence of the PReGA. To ensure the efficiency of this method, it is important to choose the numerical parameter $\nu \in \mathbb{R}$ appearing in (2) as small as possible so that (HA) remains true. If the value of ν is well-chosen, the PReGA may converge as fast as the PRaGA or the PEGA. In the example presented in Figure 1 where ν is chosen to be 0 and $\mu_1 \approx 116$, we can clearly see that the rate of convergence of the PReGA is poorer than the rates of the PRaGA and PEGA.

We also observe that the use of the ORaGA, OReGA and OEGA, instead of the pure versions of the algorithms, improves the convergence rate with respect to the number of iterations $n \in \mathbb{N}^*$. However, as n increases, the cost of the n-dimensional optimization problems (14) becomes more and more significant.

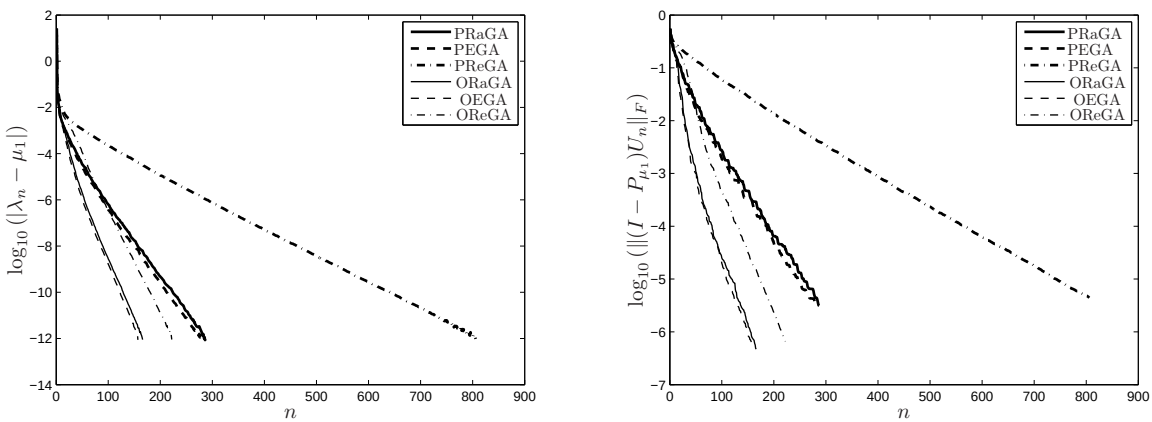

FIGURE 1. Decay of the error of the three algorithms and their orthogonal versions: eigenvalues (left) and eigenvectors (right).

REFERENCES

[1] E. Cancès, V. Ehrlacher and T. Lelièvre, *Greedy algorithms for high-dimensional eigenvalue problems*, http://arxiv.org/abs/1304.2631 (2013).

[2] R.E. Bellman, *Dynamic Programming*, Princeton University Press, 1957.

[3] H. Bungartz and M. Griebel, *Sparse grids*, Acta Numerica 13 (2004) 147–269.

[4] T. von Petersdorff and C. Schwab, *Numerical solution of parabolic equations in high dimensions*, M2AN Mathematical Modelling and Numerical Analysis 38 (2004) 93–127.

[5] 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*, ESAIM: Mathematical Modelling and Numerical Analysis 46 (2012) 595–603.

[6] W. Hackbusch, *Tensor spaces and numerical tensor calculus*, Springer, 2012.

[7] A. Chkifa, A. Cohen, R. DeVore and C. Schwab, *Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs*, ESAIM: Mathematical Modelling and Numerical Analysis 47 (2013) 253–280.

[8] P. Ladevèze, *Nonlinear computational structural mechanics: new approaches and non-incremental methods of calculation*, Springer (Berlin), 1999.

[9] A. Ammar, B. Mokdad, F. Chinesta and R. Keunings, *A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids*, Journal of Non-Newtonian Fluid Mechanics 139 (2006) 153–176.

[10] A. Nouy, *Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations* Archives of Computational Methods in Engineering 16 (2009) 251–285.

[11] V.N. Temlyakov, *Greedy Approximation*, Acta Numerica 17 (2008) 235–409.
[12] F. Chinesta, P. Ladevèze and E. Cueto, *A short review on model order reduction based on Proper Generalized Decomposition*, Archives of Computational Methods in Engineering 18 (2011) 395–404.
[13] A. Ammar and F. Chinesta, *Circumventing the curse of dimensionality in the solution of highly multidimensional models encountered in Quantum Mechanics using meshfree finite sums decompositions*, Lecture notes in Computational Science and Engineering 65 (2008) 1–17.
[14] E. Cancès, V. Ehrlacher and T. Lelièvre, *Convergence of a greedy algorithm for high-dimensional convex problems*, Mathematical Models and Methods in Applied Sciences 21 (2011) 2433–2467.
[15] A. Nouy and A. Falco, *Proper Generalized Decomposition for nonlinear convex problems in tensor Banach spaces*, Numerische Mathematik 121 (2012) 503–530.
[16] C. Le Bris, T. Lelièvre and Y. Maday, *Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations*, Constructive Approximation 30 (2009) 621–651.
[17] E. Cancès, V. Ehrlacher and T. Lelièvre, *Greedy algorithms for high-dimensional non-symmetric linear problems*, arXiv:1210.6688 (2012).
[18] M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV: Analysis of Operators*, Academic Press, 1978.
[19] V. Khoromskai, B.N. Khoromskij and R. Schneider, *QTT representation of the Hartree and Exchange operators in electronic structure calculations*, Computational Methods in Applied Mathematics 11 (2011) 327–341.
[20] S. Lojasiewicz, *Ensembles semi-analytiques*, Institut des Hautes Etudes Scientifiques, 1965.
[21] A. Levitt, *Convergence of gradient-based algorithms for the Hartree-Fock equations*, ESAIM: Mathematical Modelling and Numerical Analysis (M2AN) 46 (2012) 1321–1336.
[22] A. Falco and A. Nouy, *Constrained Tensor Product Approximations based on Penalized Best Approximations*, http://hal.archives-ouvertes.fr/hal-00577942/ (2011).