



**HAL**  
open science

# Regularity and hp discontinuous Galerkin finite element approximation of linear elliptic eigenvalue problems with singular potentials

Yvon Maday, Carlo Marcati

## ► To cite this version:

Yvon Maday, Carlo Marcati. Regularity and hp discontinuous Galerkin finite element approximation of linear elliptic eigenvalue problems with singular potentials. 2018. hal-01900286

**HAL Id: hal-01900286**

**<https://hal.sorbonne-universite.fr/hal-01900286>**

Preprint submitted on 21 Oct 2018

**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.

# REGULARITY AND $hp$ DISCONTINUOUS GALERKIN FINITE ELEMENT APPROXIMATION OF LINEAR ELLIPTIC EIGENVALUE PROBLEMS WITH SINGULAR POTENTIALS

YVON MADAY<sup>†,\*</sup> AND CARLO MARCATI<sup>†</sup>

**ABSTRACT.** We study the regularity in weighted Sobolev spaces of Schrödinger-type eigenvalue problems, and we analyse their approximation via a discontinuous Galerkin (dG)  $hp$  finite element method. In particular, we show that, for a class of singular potentials, the eigenfunctions of the operator belong to analytic-type non homogeneous weighted Sobolev spaces. Using this result, we prove that the an isotropically graded  $hp$  dG method is spectrally accurate, and that the numerical approximation converges with exponential rate to the exact solution. Numerical tests in two and three dimensions confirm the theoretical results and provide an insight into the behaviour of the method for varying discretisation parameters.

## 1. INTRODUCTION

Many problems in physics and chemistry are modeled through elliptic eigenvalue problems with singular potential. This is the case, for example, of the electronic Schrödinger equation, where the attraction between the nuclei and the electrons is proportional to the inverse of their distance. In this paper we propose and analyze the application of an isotropically graded  $hp$  discontinuous Galerkin (dG) finite element method for the approximation of the solution to linear elliptic eigenvalue problems. The central idea is that, for a wide class of singular potentials, the exact eigenfunctions are highly regular in weighted Sobolev spaces – i.e., they are smooth in Sobolev spaces when multiplied by weights that are null at the singularities. The weighted Sobolev spaces considered were introduced in the analysis of elliptic problems in domains with non-smooth boundary [Kon67]; when applied to elliptic problems in domains with corners and edges, the graded  $hp$  refinement gives rise to exponentially convergent methods [GB86d, GB86e, SSW13b, SSW13a].

Our goal is firstly, therefore, to show that the solution to the eigenvalue problems has sufficient regularity to be approximated with exponential convergence by the discontinuous  $hp$  space. Then, this can be used to prove that the solution provided by the  $hp$  dG finite element method converges with this exponential rate.

In Section 2 we start by briefly introducing the functional setting of homogeneous and non homogeneous weighted Sobolev spaces, and by stating our eigenvalue problem. We do so in a quite general way, which includes both singularities on the boundary and in the interior of the domain. Note however that in three dimensions we do not consider anisotropic approximation along the edge, thus the singularities only arise in practice from potentials.

---

<sup>†</sup> SORBONNE UNIVERSITÉ, UNIVERSITÉ PARIS-DIDEROT SPC, CNRS, LABORATOIRE JACQUES-LOUIS LIONS, LJLL, F-75005 PARIS,

\* INSTITUT UNIVERSITAIRE DE FRANCE

*E-mail addresses:* yvon.maday@upmc.fr, carlo.marcatti@upmc.fr.

2010 *Mathematics Subject Classification.* 35J10, 65N25, 65N30.

*Key words and phrases.*  $hp$  graded finite element method, discontinuous Galerkin, elliptic eigenvalue problem, Schrödinger equation, weighted Sobolev spaces, elliptic regularity.

In Section 3 we consider the issues related to the regularity of solutions to linear elliptic problem with singular points. We are mainly interested in singular points as a consequence of singular potentials, but we place ourselves in the more general case of a conical domain. The analysis therefore applies also to corner domains in two and three dimensions, a situation that has been widely studied, see, among the others, [CDN12, ES97, KMR97, MR10].

Let us consider a conical domain, i.e., a bounded domain  $\Omega \subset \mathbb{R}^d$  such that, after localization of the singularity at the origin, in polar coordinates,  $\Omega \cap S_{d-1} = (0, \zeta) \times U$ , with  $\zeta > 0$ ,  $U \subset \mathbb{S}_{d-1}$ , and  $S_{d-1}$  is the  $d - 1$  dimensional sphere. While most of the literature is concerned with the analysis in homogeneous weighted Sobolev spaces, denoted here as  $\mathcal{K}_\gamma^s(\Omega)$ , here we focus on inhomogeneous spaces, denoted as  $\mathcal{J}_\gamma^s(\Omega)$ . The latter spaces have been studied mainly as the domain of solutions to elliptic problems in corner domains with Neumann boundary conditions. The similarity arises from the fact that problems with singular potential and Neumann boundary problems in domains with conical points share solutions that a priori, have nonzero imposed value at the singular points.

The reason why a regularity result in non homogeneous weighted spaces is more relevant than its homogeneous counterpart lies in the fact that, by taking wider spaces — in general,  $\mathcal{K}_\gamma^{s,p}(\Omega) \subset \mathcal{J}_\gamma^{s,p}(\Omega)$  — we can obtain an estimate with a bigger weight  $\gamma$ . This is relevant since it can, in some situations, give insight into the boundedness of a function.

From the point of view of the Mellin transformation, working in non homogeneous spaces consists in isolating some singularities of the Mellin transform of the solution, bounding the rest of the function using the theory of homogeneous spaces, and finally bounding the terms in the expansion of the solution corresponding to the singularities via embeddings in higher order non homogeneous spaces. To illustrate this, consider the Mellin symbol  $\mathfrak{L}$  related to a Laplacian operator in a conical domain given by  $\mathfrak{L}(\omega, \partial_\omega, \lambda) = -((\lambda + d - 2)\lambda + \Delta_U)$ , with  $-\Delta_U$  representing the Laplace-Beltrami operator on  $U \subset \mathbb{S}_{d-1}$ , in the case where  $\Delta_U$  has a null eigenvalue (corresponding to spherically symmetric functions). The symbol has a single (resp. double) zero for  $\lambda = 0$  in three (resp. two) dimensions. In three dimension, this zero corresponds to a constant in the asymptotic expansion of the solution near the singularity; in two dimensions, we would have a constant and logarithmic term  $\log(|x|)$ , but this would not be in  $H^1(\Omega)$ . In the asymptotic expansion of the solution near the singularity, we will therefore find a constant, followed by a term due to the potential or the geometry of the domain. The former case depends on the asymptotic expansion of the potential near the singularity, while the latter depends on the eigenvalues of  $\Delta_U$ . In the following sections, we will suppose that the term following the constant in the expansion goes as  $|x|^\varepsilon$  for an  $0 < \varepsilon < 1$ . As an example of a potential that would generate such a behavior, consider  $V(x) = |x|^{-2+\varepsilon}$ . A geometry causing an expansion containing  $|x|^\varepsilon$  would instead be one such that  $\varepsilon(\varepsilon + d - 2) \in \sigma(-\Delta_U, B_{\partial U})$ , i.e., there exists a function  $\hat{u} : \mathbb{R}^+ \rightarrow H^s(U)$ ,  $s \geq 2$  such that

$$(\mathfrak{L}\hat{u})(\lambda) = -((\lambda + d - 2)\lambda - (\varepsilon + d - 2)\varepsilon) \hat{u}(\lambda).$$

As it can easily be seen,  $\varepsilon$  is indeed a zero of the symbol above. More practically, this happens if we consider a two dimensional wedge with aperture  $\pi/\varepsilon$  (in this two dimensional wedge case we have therefore also  $\varepsilon \geq 1/2$ ), as it will be outlined later in Section 2.2.

Returning to weighted Sobolev spaces, in light of the analysis of the operator given above, we can consider a simple case by neglecting the higher order terms, and consider a function  $v(x) = v(0) + |x|^\varepsilon$ , with  $v(0) \neq 0$  as a model of our solution. As long as  $|x| \ll 1$ , those terms are indeed the predominant ones. The norm  $\|r^{-d/p}v\|_{L^p(\Omega)}$  is clearly unbounded, thus  $v \notin \mathcal{K}_\gamma^{s,p}(\Omega)$  for any  $s \in \mathbb{N}$  and any  $\gamma \geq d/p$ . It is easy to see, though, that the statement  $v \in \mathcal{K}_\gamma^{s,p}(\Omega)$  for any  $s \in \mathbb{N}$  and  $\gamma < d/p$  does not tell the whole story, since  $v - v(0) \in \mathcal{K}_\gamma^{s,p}(\Omega)$  also for larger values of

$\gamma \in (d/p, d/p + \varepsilon)$ . The non homogeneous weighted spaces give therefore a framework where functions such as  $v$  can be treated more naturally than in homogeneous spaces.

We define the spaces treated above in more detail and outline the relationships between the homogeneous and non homogeneous ones in the following Section 2.1. Then, in Section 2.2 we specify the class of operators we treat here. The main regularity result for those operators is then given in Section 3. Specifically, we give an elliptic regularity result in non homogeneous weighted Sobolev spaces for operators with singular potential, and we follow with an observation on how this can be used as a basis to obtain “analytic regularity” in weighted spaces – see Corollary 5.

In Section 4 we introduce the  $hp$  discontinuous Galerkin method we use to approximate the solution to the eigenproblems considered and we prove our convergence results.

Historically, dG methods have been originally introduced for the approximation of first order steady equations in [RH73, CR73] in the context of neutron transport equations and of the Stokes equation. For second order elliptic problems, the development of discontinuous Galerkin methods is based on the ideas in [Nit72], with interior penalty methods being introduced in [Whe78] and developed in [Arn82]. A wide range of different methods have been proposed throughout the years, including, among others, the local discontinuous Galerkin (LDG) method [CS98], and the already mentioned class of interior penalty (IP) methods, in its symmetric (SIP), nonsymmetric (NIP) and incomplete (IIP) versions. See [Riv08, HW08, DE12] for an overview of discontinuous Galerkin methods.

The  $hp$  version of finite element (FE) methods, introduced in [GB86a, GB86b, GB86c] in one dimension and in [GB86d, GB86e] in more dimensions, combines adaptivity in space in low regularity regions with adaptivity in polynomial degree in high regularity regions. When applied to elliptic problems with point singularities, the numerical solutions obtained with the  $hp$  FE method can converge with exponential rate, provided that the exact solutions belong to the spaces  $\mathcal{K}_\gamma^\varpi(\Omega)$  or  $\mathcal{J}_\gamma^\varpi(\Omega)$  defined in Section 2.1. We also signal the recent research on  $hp$  methods in polygonal and polyhedral domains, see, among others, [CDS05, SW10, SSW13b, SSW13a, SSW16].

We focus on the symmetric version of the interior penalty method, since the original problem is itself symmetric and preserving symmetry improves both the numerical stability and the theoretical convergence rate of the method. In particular, as shown in Theorem 3, the eigenvalues can be shown to converge at a rate twice that of the eigenfunctions. Our convergence analysis follows closely what has been shown in [ABP06], with some minor differences related to the specificity of the regularity of the eigenspaces and to the approximation of the  $hp$  space.

We conclude, in Section 5, with some numerical tests in two and three dimensions, in which we confirm our theoretical results and investigate the role of the sources of numerical error that we did not consider in the analysis. Those tests can also be used, in practice, to devise specifically crafted  $hp$  spaces for more complex problems.

## 2. NOTATION AND STATEMENT OF THE PROBLEM

Let us consider a bounded domain  $\Omega \subset \mathbb{R}^d$ ,  $d = 2, 3$  which will be specified later and let  $\mathcal{C}$  be a set of isolated points in  $\Omega$ ; for the sake of simplicity we consider the case of a single point  $\mathcal{C} = \{c\}$ ; the theory can be trivially extended to the case of a finite number of points. We then denote by  $r = r(x)$  the distance  $|x - c|$  where  $|\cdot|$  is the euclidean norm of  $\mathbb{R}^d$  (it would be a smooth function representing the distance from the nearest point in  $\mathcal{C}$  if there were more than one).

We denote by  $C \in \mathbb{R}^+$  a generic constant independent of the discretization, and write  $A \lesssim B$  (resp.  $A \gtrsim B$ ) if  $A \leq CB$  (resp.  $A \geq CB$ ) and  $A \simeq B$  if both  $A \lesssim B$  and  $A \gtrsim B$  hold.

**2.1. Weighted Sobolev spaces.** For  $k \in \mathbb{N}$  and  $1 \leq p < \infty$  we introduce on a set  $D \subset \Omega$  the homogeneous weighted norm

$$(1) \quad \|u\|_{\mathcal{K}_\gamma^{k,p}(D)}^p = \sum_{j=0}^k \sum_{|\alpha|=j} \|r^{j-\gamma} \partial^\alpha u\|_{L^p(D)}^p$$

with seminorm

$$|u|_{\mathcal{K}_\gamma^{k,p}(D)}^p = \sum_{|\alpha|=k} \|r^{k-\gamma} \partial^\alpha u\|_{L^p(D)}^p$$

and denote the space  $\mathcal{K}_\gamma^{k,p}(\Omega)$  as the space of all functions with bounded  $\mathcal{K}_\gamma^{k,p}(\Omega)$  norm. The case where  $p = \infty$  follows from the usual modification in Sobolev spaces. We also introduce the inhomogeneous norm

$$(2) \quad \|u\|_{\mathcal{J}_\gamma^{k,p}(\Omega)}^p = \sum_{j=0}^k \sum_{|\alpha|=j} \|r^{\max(-\gamma+|\alpha|,\rho)} \partial^\alpha u\|_{L^p(\Omega)}^p,$$

for  $\gamma - d/p < k$  and  $\rho \in (-d/p, -\gamma + k]$ , if  $1 \leq p < \infty$ ,  $\rho \in [0, -\gamma + k]$  if  $p = \infty$ . We write  $\mathcal{K}_\gamma^k(\Omega) = \mathcal{K}_\gamma^{k,2}(\Omega)$  and  $\mathcal{J}_\gamma^k = \mathcal{J}_\gamma^{k,2}$ . We also remark that, for  $k \geq 1$  and  $\gamma - d/2 < 0$ ,  $\mathcal{K}_\gamma^k(\Omega) = \mathcal{J}_\gamma^k(\Omega)$ . Furthermore, if  $v \in \mathcal{J}_\gamma^k(\Omega)$  for  $k \geq 1$  and  $0 < \gamma - d/2 < 1$  (condition under which  $|v(\mathbf{c})| \lesssim \|v\|_{\mathcal{J}_\gamma^k(\Omega)}$ ),

$$(3) \quad \|v - v(\mathbf{c})\|_{\mathcal{K}_\gamma^k(\Omega)} + |v(\mathbf{c})| \simeq \|v\|_{\mathcal{J}_\gamma^k(\Omega)}.$$

On the boundary, for integer  $k \geq 1$ , we introduce the space  $\mathcal{K}_{\gamma-1/p}^{k-1/p,p}(\partial\Omega)$  (resp.  $\mathcal{J}_{\gamma-1/p}^{k-1/p,p}(\partial\Omega)$ ) of traces of functions from  $\mathcal{K}_\gamma^{k,p}(\Omega)$  (resp.  $\mathcal{J}_\gamma^{k,p}(\Omega)$ ) with norm

$$\|u\|_{\mathcal{K}_{\gamma-1/p}^{k-1/p,p}(\partial\Omega)} = \inf\{\|v\|_{\mathcal{K}_\gamma^{k,p}(\Omega)}, v|_{\partial\Omega} = u\},$$

and

$$\|u\|_{\mathcal{J}_{\gamma-1/p}^{k-1/p,p}(\partial\Omega)} = \inf\{\|v\|_{\mathcal{J}_\gamma^{k,p}(\Omega)}, v|_{\partial\Omega} = u\}.$$

Note that on portions of the boundary not touching the singularity  $\mathbf{c}$ , the weighted trace spaces coincide with classical Sobolev trace spaces.

Finally, we introduce the spaces

$$\mathcal{J}_\gamma^{\infty,p}(\Omega) = \left\{ v \in \mathcal{J}_\gamma^{\infty,p}(\Omega) : \exists A, C \in \mathbb{R} \text{ s.t. } \forall k \in \mathbb{N}, \|v\|_{\mathcal{J}_\gamma^{k,p}(\Omega)} \leq CA^k k! \right\},$$

and

$$\mathcal{K}_\gamma^{\infty,p}(\Omega) = \left\{ v \in \mathcal{K}_\gamma^{\infty,p}(\Omega) : \exists A, C \in \mathbb{R} \text{ s.t. } \forall k \in \mathbb{N}, |v|_{\mathcal{K}_\gamma^{k,p}(\Omega)} \leq CA^k k! \right\}.$$

In the following, for an  $S \subset \Omega$ , we denote by  $(\cdot, \cdot)_S$  the  $L^2(S)$  scalar product and by  $\|\cdot\|_S$  the  $L^2(S)$  norm.

**2.2. Statement of the problem.** Let us now assume that in a neighborhood of  $\mathbf{c}$ , the domain  $\Omega \subset \mathbb{R}^d$  is conical, i.e., there exists a ball  $B_\zeta(\mathbf{c})$  centered in  $\mathbf{c}$  with radius  $\zeta > 0$  such that, going from a cartesian to a polar representation,

$$\Omega \cap B_\zeta(\mathbf{c}) = (0, \zeta) \times U$$

where  $U \subset \mathbb{S}_{d-1}$  the  $d-1$  dimensional sphere, and  $\partial U$  is smooth.

In this domain we set the problem of finding  $\lambda \in \mathbb{R}$  and  $u \in H^1(\Omega)$  such that  $\|u\| = 1$  and

$$(4) \quad \begin{aligned} L(x, \partial_x)u &= -\Delta u + V(x)u = \lambda u && \text{in } \Omega \\ B(x, \partial_x)u &= 0 && \text{on } \partial\Omega \end{aligned}$$

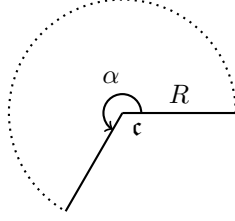


FIGURE 1. Two dimensional wedge

where  $B(x, \partial_x)$  is a boundary operator with analytic coefficients of order  $m \leq 1$  covering  $L(x, \partial_x)$ , i.e., such as the problem defined by  $(L, B)$  is elliptic. Furthermore,  $V : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is a potential such that  $V \in \mathcal{K}_{\varepsilon-2}^{\varpi, \infty}(\Omega) \cap L^p(\Omega)$  for some  $0 < \varepsilon \leq 1$  and  $p > d/2$ , and  $V$  is bounded from below by a positive constant. We will omit the dependence of  $L$  and  $B$  on  $x$  and  $\partial_x$  when it will not be strictly necessary. Finally, we denote by  $a(\cdot, \cdot)$  be the bilinear form associated to  $L$ , i.e.

$$a(u, v) = (\nabla u, \nabla v) + (Vu, v).$$

We recall the definition of the Mellin transformation

$$\hat{u}(\lambda) = (\mathcal{M}_{r \rightarrow \lambda}) u = \int_{\mathbb{R}_+} u(r, \omega) r^{-\lambda-1} dr$$

where  $(r, \omega) \in \mathbb{R}_+ \times \mathbb{S}_{d-1}$  are spherical coordinates. We denote the leading part of the operator  $L$  in (4) by  $L^0 = -\Delta$  and introduce the Mellin symbol  $\mathfrak{L}(\omega, \partial_\omega, \lambda)$  of the leading part  $L^0$ , such as

$$(5) \quad r^{-2} \mathfrak{L}(\omega, \partial_\omega, r \partial_r) = L^0(x, \partial_x),$$

i.e.,  $\mathfrak{L}(\omega, \partial_\omega, \lambda) = -((\lambda + d - 2)\lambda + \Delta_U)$ .

We also suppose, for ease of notation, that the smallest nonzero eigenvalue of the Laplace-Beltrami operator  $-\Delta_U$  on  $U$  with boundary operator  $B_{\partial U}$  on  $\partial U$  is bigger than  $\varepsilon(\varepsilon + d - 2)$ , i.e.,

$$(6) \quad \min \{ \mu > 0 : \mu \in \sigma(-\Delta_U, B_{\partial U}) \} \geq \varepsilon(\varepsilon + d - 2).$$

This condition, combined with  $V \in \mathcal{K}_{\varepsilon-2}^{\varpi, \infty}(\Omega)$  guarantees that the positive pole with smallest real part of the Mellin transform of the solution  $\hat{u}(\lambda)$  lies in the half-space  $\{\text{Re}(\lambda) \geq \varepsilon\}$ , see [KMR97, Chapter 6]. As an example of condition (6), consider a two dimensional domain that coincides near the origin with the wedge with angle of aperture  $\alpha \in (0, 2\pi)$

$$\{0 < r < R, \vartheta \in (0, \alpha)\},$$

where  $r$  and  $\vartheta$  are polar coordinates, as in Figure 1. On the boundary we impose either homogeneous Dirichlet or homogeneous Neumann boundary conditions. Then, (6) is equivalent to

$$\alpha \leq \frac{\pi}{\varepsilon}.$$

### 3. REGULARITY OF THE SOLUTION

The first lemma concerns the regularity of the solution of (4): we specialize here the results of [KMR97]. We also introduce the set  $I_d$  as

$$I_d = \begin{cases} (-1, \varepsilon) \setminus \{0\} & \text{if } d = 3 \\ [0, \varepsilon) & \text{if } d = 2. \end{cases}$$

In what follows, we analyze the set of weighted spaces where the operator  $(L, B)$  is an isomorphism. We place ourselves in the Hilbertian setting ( $p = 2$ ). In general, we avoid considering

the cases where  $\gamma - d/2 \in \mathbb{N}$ , since for those  $\gamma$  the operator is not Fredholm, the exception being  $\gamma = 1$  when  $d = 2$ , since  $\mathcal{J}_1^1(\Omega) = H^1(\Omega)$ . Given  $\gamma \in I_d$ , the image of the operator  $(L, B)$  applied to  $\mathcal{J}_\gamma^k(\Omega)$  is given by  $\mathcal{J}_{\gamma-2}^{k-2}(\Omega) \times \mathcal{J}_{\gamma-m-1/2}^{k-m-1/2}(\partial\Omega)$ . In the following lemma we will show that the operator is an isomorphism between those spaces.

The idea of the proof is then to start from in homogeneous weighted spaces and then to extend the results to the non homogeneous ones, by function decomposition.

**Lemma 1.** *The operator  $(L, B)$  is an isomorphism between the spaces*

$$(7) \quad \mathcal{J}_\gamma^k(\Omega) \rightarrow \mathcal{J}_{\gamma-2}^{k-2}(\Omega) \times \mathcal{J}_{\gamma-m-1/2}^{k-m-1/2}(\partial\Omega)$$

for  $\gamma - d/2 \in I_d$ ,  $k \geq 2$ .

*Proof.* Let  $\mathcal{F}_\gamma = (L_\gamma, B_\gamma) : \mathcal{K}_\gamma^2(\Omega) \rightarrow \mathcal{K}_{\gamma-2}^0(\Omega) \times \mathcal{K}_{\gamma-m-1/2}^{3/2-m}(\partial\Omega)$ . The operator  $\mathcal{F}_\gamma$  is Fredholm for all  $\gamma - d/2 \notin \mathbb{N}$  [KMR97]; its index is defined as

$$\text{ind } \mathcal{F}_\gamma = \dim(\ker \mathcal{F}_\gamma) - \dim(\ker \mathcal{F}_\gamma^*).$$

In the case  $d = 3$  the index is given by

$$\text{ind } \mathcal{F}_\gamma = \begin{cases} 0 & \text{if } \gamma - 3/2 \in (-1, 0) \\ -1 & \text{if } \gamma - 3/2 \in (0, \varepsilon). \end{cases}$$

When  $d = 2$ , instead,

$$\text{ind } \mathcal{F}_\gamma = \begin{cases} 1 & \text{if } \gamma - 1 \in (-1, 0) \\ -1 & \text{if } \gamma - 1 \in (0, \varepsilon). \end{cases}$$

Let us first consider the case  $\gamma - 3/2 \in (-1, 0)$  and  $d = 3$ . The operator  $\mathcal{F}_\gamma$  is coercive on  $H^1(\Omega) = \mathcal{J}_1^1(\Omega) = \mathcal{K}_1^1(\Omega)$ . It is then an isomorphism between the spaces  $\mathcal{K}_1^2(\Omega)$  and  $\mathcal{K}_{-1}^0(\Omega) \times \mathcal{K}_{1/2-m}^{3/2-m}(\partial\Omega)$ . Therefore,  $\mathcal{F}_\gamma$  is an isomorphism between the spaces (7) for all  $-1 < \gamma - 3/2 < 0$ , see [KMR97, Corollary 6.3.3].

In the case where  $\gamma = 1$  and  $d = 2$ , the uniqueness of the solution in  $H^1(\Omega)$  implies that the operator is an isomorphims between the spaces (7).

Let us now consider the case  $\gamma - d/2 \in (0, \varepsilon)$  and go back to the generic case  $d = 2, 3$ . We introduce  $\beta$  such that  $\beta - d/2 \in (-1, 0)$  and consider a solution  $u \in \mathcal{K}_\beta^s(\Omega) \cap H^1(\Omega)$ ,  $s \geq 2$ , to

$$\begin{aligned} L_\beta u &= f & \text{in } \Omega \\ B_\beta u &= g & \text{on } \partial\Omega \end{aligned}$$

for  $(f, g) \in \mathcal{J}_{\gamma-2}^{s-2}(\Omega) \times \mathcal{J}_{\gamma-m-1/2}^{s-m-1/2}(\partial\Omega)$ . The Mellin transform  $\mathfrak{L}(\lambda)$  of the principal part of  $L$  has a single zero at  $\lambda = 0$  if  $d = 3$  and a double zero if  $d = 2$ . We can decompose  $u$  as

$$u = w + u(\mathbf{c})$$

where  $w \in \mathcal{K}_\gamma^s(\Omega)$ . This is straightforward for  $d = 3$ ; for  $d = 2$  there could be a term proportional to  $\log(r)$ , but this term would not belong to  $H^1(\Omega)$ . Then,  $w$  is solution to

$$\begin{aligned} L_\gamma w &= f - Vu(\mathbf{c}) & \text{in } \Omega \\ B_\gamma w &= g - B_\gamma u(\mathbf{c}) & \text{on } \partial\Omega \end{aligned}$$

In this case  $\text{ind } \mathcal{F}_\gamma = -1$  but the right hand side in the above equation belongs to the image of  $\mathcal{F}_\gamma$ , by definition. Furthermore,  $f - Vu(\mathbf{c}) \in \mathcal{K}_{\gamma-2}^{s-2}(\Omega)$  and  $g - B_\gamma u(\mathbf{c}) \in \mathcal{K}_{\gamma-1/2-m}^{s-1/2-m}(\partial\Omega)$ . Therefore,

$$\|w\|_{\mathcal{K}_\gamma^s(\Omega)} \leq C \left( \|f\|_{\mathcal{K}_{\gamma-2}^{s-2}(\Omega)} + \|g - Bu(\mathbf{c})\|_{\mathcal{K}_{\gamma-1/2-m}^{s-1/2-m}(\partial\Omega)} + |u(\mathbf{c})| \right)$$

We now conclude as in [KMR97, Theorem 7.1.1]: since for any  $\delta > 0$ , there exists a  $C_\delta$  such that

$$|u(\mathbf{c})| \leq \delta \|u\|_{\mathcal{J}_\gamma^2(\Omega)} + C_\delta \|u\|_{\mathcal{J}_{\gamma-1}^1(\Omega)},$$

we can write, for  $s \geq 2$ ,

$$\begin{aligned} \|u\|_{\mathcal{J}_\gamma^s(\Omega)} &\leq C \left( \|w\|_{\mathcal{K}_\gamma^s(\Omega)} + |u(\mathbf{c})| \right) \\ &\leq C \left( \delta \|u\|_{\mathcal{J}_\gamma^2(\Omega)} + \|f\|_{\mathcal{K}_{\gamma-2}^{s-2}(\Omega)} + C_\delta \|u\|_{\mathcal{J}_{\gamma-1}^1(\Omega)} + \|g\|_{\mathcal{J}_{\gamma-m-1/2}^{s-m-1/2}(\partial\Omega)} \right). \end{aligned}$$

Since  $\gamma - 1 \leq d/2$ , by the arguments of the first part of the proof

$$\begin{aligned} \|u\|_{\mathcal{J}_{\gamma-1}^1(\Omega)} &\leq \|u\|_{\mathcal{J}_{\gamma-1}^2(\Omega)} \leq C \left( \|f\|_{\mathcal{J}_{\gamma-3}^0(\Omega)} + \|g\|_{\mathcal{J}_{\gamma-3/2-m}^{3/2-m}(\Omega)} \right) \\ &\leq C \left( \|f\|_{\mathcal{K}_{\gamma-2}^{s-2}(\Omega)} + \|g\|_{\mathcal{J}_{\gamma-1/2-m}^{s-1/2-m}(\partial\Omega)} \right) \end{aligned}$$

for all  $s \geq 2$ . The choice of a sufficiently small  $\delta$  then concludes the proof.  $\square$

In the following lemma we extend the estimates for corner domains developed in [CDN12, Theorem 3.7] to the case of an operator with singular potential in three dimensions. The proof follows directly from the one in the cited reference and is therefore omitted. Let  $\gamma \in I_d$  and let  $Lg \in \mathcal{K}_{\gamma-2}^\infty(\Omega)$ . We consider a dyadic decomposition of  $\Omega$  given by

$$\Omega_n = \{x \in \Omega : 2^{-n-1} < \|x\|_{\ell^\infty} < 2^{-n}\}, \quad n \geq 1$$

and denote  $\Omega'_n$  as the interior of  $\bar{\Omega}_{n-1} \cup \bar{\Omega}_n \cup \bar{\Omega}_{n+1}$ .

**Lemma 2.** *For any  $n \geq 2$  and  $s \geq 2$ , the estimate*

$$(8) \quad |g|_{\mathcal{K}_\gamma^{s,2}(\Omega_n)} \leq C^s s! \left( \sum_{j=1}^{s-2} \frac{1}{j!} |Lg|_{\mathcal{K}_{\gamma-2}^{j,2}(\Omega'_n)} + \|g\|_{\mathcal{K}_\gamma^{1,2}(\Omega'_n)} \right)$$

holds, with  $C$  independent of  $s$  and  $n$ .

We now prove an embedding result that bounds  $L^\infty(\Omega)$  norms in weighted spaces with norms of higher derivatives for  $p = 2$ . This is simply the weighted version of the classical embedding of  $H^s(\Omega)$  into  $L^\infty(\Omega)$  for  $s > d/2$ , and the proof follows almost directly via dyadic decomposition.

**Lemma 3.** *For any  $\gamma - d/2 \notin \mathbb{N}$  and for any  $t > s + d/2$  there exists  $C > 0$  such that for any  $u \in \mathcal{J}_\gamma^t(\Omega)$ ,*

$$(9) \quad \|u\|_{\mathcal{J}_{\gamma-d/2}^{s,\infty}(\Omega)} \leq C \|u\|_{\mathcal{J}_\gamma^t(\Omega)}.$$

*Proof.* To prove (9) we use the fact that  $\mathcal{J}_\gamma^t(\Omega) = \mathcal{K}_\gamma^t(\Omega) \oplus \mathbb{Q}_{[\gamma-d/2]}(\Omega)$  and decompose  $u = v + w$  such that

$$v \in \mathcal{K}_\gamma^{s,2}(\Omega) \quad \text{and} \quad w \in \mathbb{Q}_{[\gamma-d/2]}(\Omega).$$

Furthermore, we have

$$\|u\|_{\mathcal{J}_\gamma^t(\Omega)} \simeq \|v\|_{\mathcal{K}_\gamma^t(\Omega)} + \|w\|_{\mathbb{Q}_{[\gamma-d/2]}(\Omega)}$$

for any chosen norm  $\|\cdot\|_{\mathbb{Q}_{[\gamma-d/2]}(\Omega)}$ , thanks to the equivalency of norms in finite dimensional spaces, see [CDN10] and [KMR97, Theorem 7.1.1]. Then, by the triangle inequality and the definition of the norms in the weighted spaces,

$$\begin{aligned} \|u\|_{\mathcal{J}_{\gamma-d/2}^{s,\infty}(\Omega)} &\leq \|v\|_{\mathcal{J}_{\gamma-d/2}^{s,\infty}(\Omega)} + \|w\|_{\mathcal{J}_{\gamma-d/2}^{s,\infty}(\Omega)} \\ &\leq C \|v\|_{\mathcal{K}_{\gamma-d/2}^{s,\infty}(\Omega)} + \|w\|_{\mathcal{J}_{\gamma-d/2}^{s,\infty}(\Omega)}, \end{aligned}$$



and we consider separately the two terms at the right hand side. Consider the annuli

$$\Gamma_j = \{x \in \Omega : 2^{-j} < \|x\|_{\ell^\infty} < 2^{-j+1}\}, j \in \mathbb{N}$$

and let  $\hat{\Gamma} = \Gamma_0$ . Then, scaling and using a Sobolev inequality,

$$\begin{aligned} \|v\|_{\mathcal{K}_{\gamma-d/2}^{s,\infty}(\Gamma_j)} &\simeq 2^{j(\gamma-d/2)} \|\hat{v}\|_{W^{s,\infty}(\hat{\Gamma})} \\ &\lesssim 2^{j(\gamma-d/2)} \|\hat{v}\|_{H^t(\hat{\Gamma})} \\ &\lesssim 2^{j(\gamma-d/2)} \|\hat{v}\|_{\mathcal{K}_\gamma^t(\hat{\Gamma})} \\ &\simeq \|v\|_{\mathcal{K}_\gamma^t(\Gamma_j)} \\ &\lesssim \|u\|_{\mathcal{J}_\gamma^t(\Omega)}, \end{aligned}$$

where the quantities with a hat are rescaled on  $\hat{\Gamma}$ . Therefore

$$\|v\|_{\mathcal{K}_{\gamma-d/2}^{s,\infty}(\Omega)} = \sup_j \|v\|_{\mathcal{K}_{\gamma-d/2}^{s,\infty}(\Gamma_j)} \leq C \|u\|_{\mathcal{J}_\gamma^t(\Omega)}.$$

Since  $w$  lies in the finite dimensional space of polynomials of degree  $\lfloor \gamma - d/2 \rfloor$ , we can conclude with (9), where the constant  $C$  can depend on the domain  $\Omega$ , on the dimension  $d$  and on  $\gamma$ , but does not depend on  $s$  and  $t$ .  $\square$

The weighted analytic estimates then follow for  $p = \infty$ . Lemma 3 directly implies the following statement.

**Corollary 4.** *Let  $\gamma - d/2 \notin \mathbb{N}$ . If  $u \in \mathcal{J}_\gamma^{\varpi,2}(\Omega)$ , then  $u \in \mathcal{J}_{\gamma-d/2}^{\varpi,\infty}(\Omega)$ .*

It is now evident that, using Lemmas 1 and 2, we can prove that when the right hand side and the potential of (4) obey analytic growth estimates on the weighted norms of the derivatives, the solution  $u$  is in the same regularity class. This is the content of the following corollary.

**Corollary 5.** *If  $u$  is solution to (4) with  $V : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that  $V \in \mathcal{K}_{\varepsilon-2}^{\varpi,\infty}(\Omega)$  for some  $0 < \varepsilon \leq 1$ , then  $u \in \mathcal{J}_\gamma^{\varpi,2}(\Omega)$  for any  $\gamma < d/2 + \varepsilon$ .*

*Proof.* By the initial regularity of  $u \in H^1(\Omega)$  and Lemma 1 we have that  $u \in \mathcal{J}_\gamma^2(\Omega)$  for  $\gamma \in I_d$ . We can decompose  $u = (u - u(\mathbf{c})) + u(\mathbf{c})$  and apply (8) to  $g = u - u(\mathbf{c})$ . Then,  $V \in \mathcal{K}_{-2+\varepsilon}^{\varpi,\infty}(\Omega)$ , and  $|u(\mathbf{c})| \leq C$  by Lemma 3, hence  $(L - \lambda)g = \lambda u(\mathbf{c}) - V u(\mathbf{c}) \in \mathcal{K}_{\gamma-2}^{\varpi}(\Omega)$ . In addition  $\|u - u(\mathbf{c})\|_{\mathcal{K}_\gamma^1(\Omega)} \lesssim \|u\|_{\mathcal{J}_\gamma^1(\Omega)} + |u(\mathbf{c})|$ . Summing the left and right hand sides of (8) over all  $\Omega_k$  gives the existence of  $C, A \in \mathbb{R}^+$  such that

$$|u|_{\mathcal{K}_\gamma^s(\Omega)} \leq CA^s s!,$$

for all  $s \geq 2$  and  $\gamma \in I_d$ , thus  $u \in \mathcal{J}_\gamma^{\varpi}(\Omega)$ .  $\square$

#### 4. NUMERICAL APPROXIMATION

In this section, we consider the approximation of the linear elliptic eigenvalue problem (4) obtained through an isotropically graded discontinuous Galerkin  $hp$  method.

The contents of the section are largely based on [ABP06], where the convergence of the discontinuous Galerkin method is proven for linear elliptic eigenvalue problems. The result obtained in that paper is an extension to discontinuous Galerkin methods of the theory developed almost three decades earlier, see [DNR78a, DNR78b]. A thorough presentation of the approximation of eigenvalue problems is also given in [CL91, Chapter II].

The only differences with the analysis in [ABP06] are due to the presence of a potential and to the specificity of approximation in isotropically refined  $hp$  finite element spaces.

In the following, we introduce the analysis developed in the aforementioned papers, specializing it to problems with singular points and interior penalty discontinuous Galerkin methods.

We also introduce the interior penalty discontinuous Galerkin methods that will be taken into consideration. We will be dealing with a symmetric operator and a coercive linear form, thus the spectrum is composed of real isolated eigenvalues of ascent one. The analysis can still be partially extended to non-symmetric problems, but it has to be taken into account that the operators are not self-adjoint. We conclude the section by introducing the “solution operators”  $T$ , for the continuous problem, and  $T_\delta$ , for the discrete approximation.  $T$  and  $T_\delta$  are continuous and invertible operators, with the same eigenspaces as the ones of the original problems and with reciprocal eigenvalues. The analysis will center around those operators, and the final results can easily be applied back to the original problems. Finally, we will need a way to measure a “distance” between eigenspaces: this is the role of  $\delta(\cdot, \cdot)$  and  $\hat{\delta}(\cdot, \cdot)$  defined in (15).

The interest of the analysis of the approximation of an eigenvalue problem lies not only in the convergence of the numerical eigenpairs to the exact ones, but also in the non pollution and completeness of eigenfunctions and eigenvalues. Basically, a good approximation of an eigenproblem should not introduce any spurious numerical eigenvalue or eigenvector (non-pollution) and should approximate all eigenpairs (completeness). In Theorem 1 we show that the spectrum is not polluted, while in Theorem 2 the completeness of the approximation is shown (more precisely, Theorem 2 gives both completeness and convergence for finite dimensional eigenspaces, while simple completeness is a consequence of Lemma 10). Note that, in practice, some techniques may still introduce spurious eigenvalues in the approximation: consider for example the “strong” imposition of boundary conditions in a numerical code, where the matrix resulting from the approximation of the operator is modified in order to set the degrees of freedom at the boundary, see, e.g., the documentation of [ABD<sup>+</sup>17]. This is out of the scope of the present analysis; furthermore, the spurious eigenvalues can often be easily identified.

Finally, in Section 4.3, the focus is on the rate of convergence of the numerical eigenpairs. We consider finite dimensional exact eigenspaces and we introduce a projector from the exact to the numerical eigenspace, thus obtaining a somewhat algebraic problem, at least in the relationship between the eigenvalues and the (projected) operators  $\hat{T}$  and  $\hat{T}_\delta$  (the latter can be seen as tensors in the finite dimensional eigenspace). We obtain the expected quasi optimal estimates on the difference between exact and numerical eigenfunctions. The eigenvalue error, additionally, can be shown to converge with a higher rate of convergence — quadratically with respect to the eigenfunctions — if the method is adjoint consistent (symmetric, in our case).

We now introduce the discontinuous Galerkin interior penalty method.

**4.1. Interior penalty method.** Let  $\mathcal{T}$  be a mesh isotropically and geometrically graded around the points in  $\mathfrak{C}$ . We assume that the mesh is shape- and contact-regular and we indicate by  $\Omega_j$ ,  $j = 1, \dots, \ell$ , the set of elements and edges at the same level of refinement. We introduce on this mesh the  $hp$  space with refinement ratio  $\sigma$  and linear polynomial slope  $\mathfrak{s}$ , i.e., for an element  $K \in \mathcal{T}$  such that  $K \in \Omega_j$ ,

$$h_K \simeq h_j = \sigma^j \text{ and } p_K \simeq p_j = p_0 + \mathfrak{s}(\ell - j),$$

where  $h_K$  is the diameter of the element  $K$  and  $p_K$  is the polynomial order whose role will be specified in (10). We suppose that for any  $K \in \mathcal{T}$  there exists an affine transformation  $\Phi : K \rightarrow \hat{K}$  to the  $d$ -dimensional cube  $\hat{K}$  such that  $\Phi(K) = \hat{K}$ , and introduce the discrete space

$$(10) \quad X_\delta = \left\{ v_\delta \in L^2(\Omega) : (v|_K \circ \Phi^{-1}) \in \mathbb{Q}_{p_K}(\hat{K}) \forall K \in \mathcal{T} \right\},$$

where  $\mathbb{Q}_p$  is the space of polynomials of maximal degree  $p$  in any variable. Let then  $\mathcal{E}$  be the set of the edges (for  $d = 2$ ) or faces ( $d = 3$ ) of the elements in  $\mathcal{T}$  and

$$\begin{aligned} \mathbf{h}_e &= \min_{K \in \mathcal{T}: e \cap \partial K \neq \emptyset} h_K \\ \mathbf{p}_e &= \max_{K \in \mathcal{T}: e \cap \partial K \neq \emptyset} p_K. \end{aligned}$$

Note that edges and faces are open  $d-1$  dimensional sets. On an edge/face between two elements  $K_\sharp$  and  $K_b$ , i.e., on  $e \subset \partial K_\sharp \cap \partial K_b$ , the average  $\{\!\!\{ \cdot \}\!\!\}$  and jump  $\llbracket \cdot \rrbracket$  operators for a function  $w \in X(\delta)$  are defined by

$$\{\!\!\{ w \}\!\!\} = \frac{1}{2} \left( w|_{K_\sharp} + w|_{K_b} \right), \quad \llbracket w \rrbracket = w|_{K_\sharp} \mathbf{n}_\sharp + w|_{K_b} \mathbf{n}_b,$$

where  $\mathbf{n}_\sharp$  (resp.  $\mathbf{n}_b$ ) is the outward normal to the element  $K_\sharp$  (resp.  $K_b$ ). In the following, for an  $S \subset \Omega$ , we denote by  $(\cdot, \cdot)_S$  the  $L^2(S)$  scalar product and by  $\|\cdot\|_S$  the  $L^2(S)$  norm.

We indicate by  $a_\delta(\cdot, \cdot) : X_\delta \times X_\delta \rightarrow \mathbb{R}$  the interior penalty bilinear form, given by

$$\begin{aligned} (11) \quad a_\delta(u_\delta, v_\delta) &= (\nabla u_\delta, \nabla v_\delta)_\mathcal{T} - (\{\!\!\{ \nabla u_\delta \}\!\!\}, \llbracket v_\delta \rrbracket)_{\mathcal{E}_I} - \vartheta (\{\!\!\{ \nabla v_\delta \}\!\!\}, \llbracket u_\delta \rrbracket)_{\mathcal{E}} \\ &\quad + \sum_{e \in \mathcal{E}} \alpha_e \frac{\mathbf{p}_e^2}{\mathbf{h}_e} (\llbracket u_\delta \rrbracket, \llbracket v_\delta \rrbracket)_e + \int_\Omega V u_\delta v_\delta. \end{aligned}$$

Here,  $\mathcal{E}_I$  is the set of internal edges such that for all  $e \in \mathcal{E}_I$ ,  $e \cap \partial \Omega = \emptyset$ , and we have written

$$(\cdot, \cdot)_\mathcal{T} = \sum_{K \in \mathcal{T}} (\cdot, \cdot)_K \quad (\cdot, \cdot)_\mathcal{E} = \sum_{e \in \mathcal{E}} (\cdot, \cdot)_e.$$

The discrete eigenvalue problem then reads: find  $(\lambda_\delta, u_\delta) \in \mathbb{C} \times X_\delta$

$$(12) \quad a_\delta(u_\delta, v_\delta) = \lambda_\delta (u_\delta, v_\delta) \text{ for all } v_\delta \in X_\delta.$$

Choosing  $\vartheta = 1$  in (11) gives the symmetric interior penalty (SIP) method, while  $\vartheta = -1$  gives the non-symmetric interior penalty (NIP) method, and  $\vartheta = 0$  gives the incomplete interior penalty method (IIP). We remark that the choice  $\vartheta = 1$  is the only one that assures the symmetry of the method; the SIP method is *adjoint consistent*.

We write  $X = H^1(\Omega)$ ,  $X(\delta) = X + X_\delta$  and introduce the mesh dependent norms

$$\|v\|_{\text{DG}}^2 = \sum_{K \in \mathcal{T}} \|v\|_{H^1(K)}^2 + \sum_{e \in \mathcal{E}} \mathbf{p}_e^2 \mathbf{h}_e^{-1} \|\llbracket v \rrbracket\|_{L^2(e)}^2$$

and

$$\| \|v\|_{\text{DG}}^2 = \|v\|_{\text{DG}}^2 + \sum_{e \in \mathcal{E}} \mathbf{h}_e \mathbf{p}_e^{-2} \|\nabla v\|_{L^2(e)}^2.$$

Note that  $\|\cdot\|_{\text{DG}}$  is defined on  $X(\delta)$ , while  $\| \| \cdot \|_{\text{DG}}$  is defined only on the broken space

$$X(\delta) \cap H^{d/2}(\mathcal{T}) = \left\{ v \in X(\delta) : v \in H^{d/2}(K) \text{ for all } K \in \mathcal{T} \right\},$$

due to the presence of the boundary gradient term. We introduce the continuous solution operator

$$(13) \quad T : L^2(\Omega) \rightarrow X$$

such that

$$a(Tu, v) = (u, v), \text{ for all } v \in X$$

and its discrete counterpart, given by

$$(14) \quad T_\delta : L^2(\Omega) \rightarrow X_\delta$$

such that

$$a_\delta(T_\delta u, v) = (u, v), \text{ for all } v_\delta \in X_\delta.$$

The analysis of the relation between the spectra associated to the operator  $L$  in (4) and to the discrete bilinear form  $a_\delta$  can be reconducted to the analysis of the spectra of  $T$  and  $T_\delta$ . Since the bilinear form associated to  $L$  is continuous, coercive on  $X$ , and symmetric,

- (i) all the eigenvalues are real and strictly positive,
- (ii) the set of the eigenvalues of  $L$  is a countably infinite sequence diverging to  $\infty$ ,
- (iii) all the eigenspaces are finite dimensional,
- (iv) eigenfunctions associated with different eigenvalues are L2-orthogonal,
- (v) the eigenfunctions are complete  $L^2(\Omega)$  and in  $X$ .

In the following, the spectrum of  $T$  will be denoted by  $\sigma(T)$  and its resolvent set by  $\rho(T)$ . Similarly,  $\sigma(T_\delta)$  and  $\rho(T_\delta)$  will be respectively the spectrum and resolvent set of  $T_\delta$ . Let then

$$R_z(T) = (z - T)^{-1}$$

be the resolvent operator associated with  $T$ , and

$$R_z(T_\delta) = (z - T_\delta)^{-1}$$

be the resolvent operator associated with  $T_\delta$ , both defined for  $z \in \mathbb{C}$ . Finally, we introduce a measure of the gap between subspaces of  $X(\delta)$ : let  $Y$  and  $Z$  be close subspaces of  $X(\delta)$ ; then for an  $x \in X$

$$(15) \quad \begin{aligned} \delta(x, Y) &= \inf_{y \in Y} \|x - y\|_{\text{DG}}, & \delta(Y, Z) &= \sup_{y \in Y: \|y\|_{\text{DG}}=1} \delta(y, Z) \\ \hat{\delta}(Y, Z) &= \max(\delta(Y, Z), \delta(Z, Y)) \end{aligned}$$

## 4.2. Non pollution and completeness of the discrete spectrum and eigenspaces.

**4.2.1. Non pollution of the spectrum.** In this section we detail the technique used in [ABP06] to prove the non-pollution of the discrete spectrum. Note that, thus far,  $R_z(T_\delta)$  has only been defined formally. We will now show its existence and continuity, together with the existence and continuity of its inverse. This will imply the non pollution of the discrete spectrum and guarantee that, for a sufficient number of degrees of freedom, the discrete spectrum lies in the vicinity of the continuous one.

We start by introducing a lemma, whose proof we postpone to the end of the section.

**Lemma 6.** *Let  $z \in \rho(T)$  such that  $z \neq 0$  and  $u \in X(\delta)$ . Then, there exists  $C > 0$  such that*

$$\|(z - T)u\|_{\text{DG}} \geq C\|u\|_{\text{DG}}$$

where  $C$  depends on  $L$ , on  $\Omega$ , and on  $|z|$ .

By the triangle inequality, then,

$$(16) \quad \|(z - T_\delta)u\|_{\text{DG}} \geq \|(z - T)u\|_{\text{DG}} - \|(T - T_\delta)u\|_{\text{DG}}.$$

Now, the second term at the right hand side is the classical error of the method; by the coercivity and continuity of the discrete bilinear form, Lemma 1 and the approximation properties of the  $hp$  space, we have that

$$\|(T - T_\delta)u\|_{\text{DG}} \rightarrow 0 \text{ as } N \rightarrow \infty$$

where  $N$  is the dimension of  $X_\delta$ . Using Lemma 6 and the above estimate in (16), we obtain that, for a sufficient number of degrees of freedom,

$$(17) \quad \|(z - T_\delta)u\|_{\text{DG}} \geq C\|u\|_{\text{DG}}$$

for  $0 \neq z \in \rho(T)$ . For a fixed  $z$  and for a sufficient number of degrees of freedom (depending on  $z$ ), thus,  $z - T_\delta$  is invertible and  $R_z(T_\delta)$  is well defined. Furthermore, Lemma 6 implies that

$R_z(T)$  is well defined and bounded as an operator on the spaces  $X(\delta) \rightarrow X(\delta)$ . We have therefore shown that  $R_z(T_\delta)$  is bounded as a linear operator from  $X(\delta)$  to  $X(\delta)$ , and that the spectrum is not polluted; in the following we summarize this results. Denoting by  $\|\cdot\|_{\mathcal{L}(V,W)}$  the classical operator norm

$$\|F\|_{\mathcal{L}(V,W)} = \sup_{v \in V: \|v\|_V=1} \|Fv\|_W,$$

from (17) we conclude that

**Lemma 7.** *Let  $A \subset \rho(T)$  be a closed set. Then, for all  $z \in A$ , there exists a constant  $C$  such that*

$$\|R_z(T_\delta)\|_{\mathcal{L}(X(\delta),X(\delta))} \leq C.$$

The non-pollution of the spectrum follows directly, taking the complementary of the set  $A$  above.

**Theorem 1.** *Let  $B \supset \sigma(T)$  be an open set. Then, for a sufficient number of degrees of freedom,*

$$\sigma(T_\delta) \subset B.$$

We conclude the section with the proof of Lemma 6.

*Proof of Lemma 6.* Consider  $u \in X(\delta)$  and  $0 \neq z \in \rho(T)$ . Then, by the triangle inequality,

$$(18) \quad \|z\| \|u\|_{\text{DG}} \leq \|zTu\|_{\text{DG}} + \|(z-T)u\|_{\text{DG}}.$$

Let now  $v = zTu$ . Then, by the definition of  $T$ ,  $zu = Lv$  and

$$Lv - \frac{1}{z}v = (z-T)u$$

with the associated boundary conditions. Since  $z \in \rho(T)$ , the operator  $L - 1/z$  is invertible, and

$$\|zTu\|_{\text{DG}} = \|v\|_X \leq C\|(z-T)u\|_{L^2(\Omega)}.$$

The constant  $C$  clearly depends on  $z$ , on the operator  $L$ , and on  $\Omega$ . Inserting the above inequality into (18) one obtains the thesis.  $\square$

**4.2.2. Eigenspaces and completeness of the spectrum.** Consider a smooth closed curve  $\Gamma \subset \rho(T)$ . We introduce the spectral projectors

$$(19) \quad E = \frac{1}{2\pi i} \int_{\Gamma} R_z(T) dz \quad \text{and} \quad E_\delta = \frac{1}{2\pi i} \int_{\Gamma} R_z(T_\delta) dz$$

Clearly, both projectors depend on  $\Gamma$ , we omit that in our notation as is customary: suppose that  $\Gamma$  is fixed and that it encloses a single eigenvalue of  $T$ . The discrete projector  $E_\delta$  is, once again, well defined provided that the space  $X_\delta$  contains a sufficient number of degrees of freedom. Suppose that  $\Gamma$  contains an eigenvalue of  $T$ ; then,  $E$  is the projector on the eigenspace associated to the eigenvalue. The same holds for the discrete version.

We now wish to prove the convergence of the discrete projector to the continuous one, in the operator norm. We start by noting that

$$(z-T)^{-1} - (z-T_\delta)^{-1} = (z-T_\delta)^{-1}(T-T_\delta)(z-T)^{-1},$$

therefore,

$$\begin{aligned} \|R_z(T) - R_z(T_\delta)\|_{\mathcal{L}(L^2(\Omega),X(\delta))} &= \|R_z(T_\delta)(T-T_\delta)R_z(T)\|_{\mathcal{L}(L^2(\Omega),X(\delta))} \\ &\leq \|R_z(T_\delta)\|_{X(\delta),X(\delta)} \|(T-T_\delta)\|_{\mathcal{L}(L^2(\Omega),X(\delta))} \|R_z(T)\|_{L^2(\Omega),L^2(\Omega)}. \end{aligned}$$

Due to the boundedness of the continuous, see [ABP06], and discrete resolvent operators, see Lemma 7, we conclude that

$$(20) \quad \|E - E_\delta\|_{\mathcal{L}(L^2(\Omega),X(\delta))} \leq C\|(T-T_\delta)\|_{\mathcal{L}(L^2(\Omega),X(\delta))}.$$

**Lemma 8.** *Given the definition of  $E$  and  $E_\delta$  in (19), if  $X_\delta$  has a sufficient number of degrees of freedom, there holds*

$$\|E - E_\delta\|_{\mathcal{L}(L^2(\Omega), X(\delta))} \rightarrow 0.$$

Consider now the definitions given in (15). The convergence of the projectors allows for the proof of the convergence to zero of some “distances” between eigenspaces. The first almost direct result is in the following lemma.

**Lemma 9.** *Let  $\delta(\cdot, \cdot)$  be defined as in (15). Then,*

$$\delta(E_\delta(X_\delta), E(X)) \rightarrow 0$$

*Proof.* For any  $x_\delta \in E_\delta(X_\delta)$ ,  $E_\delta(x_\delta) = x_\delta$ . We remark that, due to the regularity result given in Lemma 1,  $E(L^2(\Omega)) = E(X)$ . Thus, for any  $x_\delta \in E_\delta(X_\delta)$  such that  $\|x_\delta\|_{\text{DG}} = 1$ ,

$$\begin{aligned} \inf_{x \in E(X)} \|x_\delta - x\|_{\text{DG}} &= \inf_{x \in E(L^2(\Omega))} \|x_\delta - x\|_{\text{DG}} \\ &= \inf_{y \in L^2(\Omega)} \|E_\delta x_\delta - Ey\|_{\text{DG}} \\ &\leq \|E_\delta - E\|_{\mathcal{L}(L^2(\Omega), X(\delta))}. \end{aligned}$$

Taking the supremum over all  $x_\delta \in E_\delta(X_\delta)$  one obtains the thesis.  $\square$

This is a proof of the non pollution of the eigenspaces: we have indeed shown that all numerical eigenfunction converge to an exact one. We continue by showing the completeness of the eigenspaces. This involves proving that any exact eigenfunction is approximated by a numerical one.

**Lemma 10.** *For any  $x \in E(X)$ ,*

$$\delta(x, E_\delta(X_\delta)) \rightarrow 0$$

*Proof.* Let  $x \in E(X)$  and  $x_\delta \in X_\delta$ . Then,

$$\|E_\delta x_\delta - x\|_{\text{DG}} \leq \|E\|_{\mathcal{L}(X(\delta), X(\delta))} \|x_\delta - x\|_{\text{DG}} + \|E - E_\delta\|_{\mathcal{L}(X(\delta), X(\delta))} \|x_\delta\|_{\text{DG}}.$$

Taking  $x_\delta$  as the projection of  $x$  in  $X_\delta$  and thanks to the convergence of  $E_\delta$  towards  $E$ , we obtain the thesis.  $\square$

We now restrict our focus to finite dimensional eigenspaces. Let then  $n = \dim(E(X))$  and  $n_\delta = \dim(E_\delta(X_\delta))$ : if  $n = \infty$ , then  $n_\delta \rightarrow \infty$ ; we consider the case where  $n$  is finite. If  $n$  is finite, the above lemma implies that

$$\delta(E(X), E_\delta(X_\delta)) \rightarrow 0.$$

**Remark 1.** *The eigenspace  $E(X)$  is invariant for  $T$ , hence if  $x \in E(X)$ , then  $R_z(T)x \in E(X)$ .*

Consider then an  $x \in E(X)$ : we have

$$(21) \quad \inf_{x_\delta \in X_\delta} \|E_\delta x_\delta - x\|_{\text{DG}} \leq \|E_\delta\|_{\mathcal{L}(X(\delta), X(\delta))} \inf_{x_\delta \in X_\delta} \|x_\delta - x\|_{\text{DG}} + \|(E - E_\delta)x\|_{\text{DG}}$$

Due to the approximation properties of  $X_\delta$  there exist  $C, b > 0$  such that

$$(22) \quad \inf_{x_\delta \in X_\delta} \|x - x_\delta\|_{\text{DG}} \leq Ce^{-bN^{1/(d+1)}},$$

with  $N = \dim(X_\delta)$ . In addition,

$$\begin{aligned} \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(R_z(T) - R_z(T_\delta))x\|_{\text{DG}} &= \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(R_z(T_\delta)(T - T_\delta)R_z(T))x\|_{\text{DG}} \\ &\leq C\|R_z(T_\delta)\|_{\mathcal{L}(X(\delta), X(\delta))} \\ &\quad \times \|(T - T_\delta)|_{E(X)}\|_{\mathcal{L}(L^2(\Omega), X(\delta))} \|R_z(T)\|_{\mathcal{L}(L^2(\Omega), L^2(\Omega))}, \end{aligned}$$

As above, the boundedness of the continuous, see [ABP06], and discrete resolvent operators, see Lemma 7, imply that

$$(23) \quad \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(E - E_\delta)x\|_{\text{DG}} \leq C \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)x\|_{\text{DG}}$$

Thanks to Remark 1, the right hand side of the above equation is the error of the numerical method for a problem with source term belonging to  $\mathcal{J}_\gamma^\infty(\Omega)$ : by Lemma 1, the approximation properties of the  $hp$  space, and the compactness of the unitary ball in the finite dimensional space  $E(X)$ , there exist  $C, b > 0$  such that

$$(24) \quad \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(E - E_\delta)x\|_{\text{DG}} \leq Ce^{-bN^{1/(d+1)}}.$$

Combining (21), (22) and (24), we have then the explicit rate

$$\delta(E(X), E_\delta(X_\delta)) \leq Ce^{-bN^{1/(d+1)}}.$$

We summarize this in the following statement.

**Theorem 2.** *If  $\dim(E(X)) < \infty$  and for a sufficient number of degrees of freedom, there exist  $C, b > 0$  such that*

$$\delta(E(X), E_\delta(X_\delta)) \leq Ce^{-bN^{1/(d+1)}}.$$

**4.3. Convergence of the eigenfunctions and eigenvalues.** In this section we consider the convergence of the numerical eigenfunctions and eigenvalues obtained through the  $hp$  approximation. As far as the eigenspaces are concerned, Lemma 9 proves that they are not polluted and Lemma 10 proves that they are complete. As a direct consequence of Theorem 2, furthermore, we have that for any  $x \in E(X)$  there exists  $x_\delta \in E_\delta(X_\delta)$  such that

$$\|x - x_\delta\|_{\text{DG}} \leq Ce^{-bN^{1/(d+1)}}.$$

We now consider the eigenvalues; we will do so in the case of a symmetric numerical scheme.

**4.3.1. Convergence of the eigenvalues.** We are mainly interested in the analysis of the convergence of the eigenvalues for the symmetric interior penalty method, obtained by choosing  $\vartheta = 1$  in (11). For the sake of generality, the first part of the section will, nonetheless, hold for non-symmetric methods and for a non symmetric operator, but we will indicate when the hypothesis of symmetry of the numerical method will become necessary. The final result obtained for the SIP method will be stronger than what can be obtained in the case of non-symmetric methods, since they lack the property of adjoint consistency.

We start by considering the operator  $\Lambda_\delta = E_\delta|_{E(X)} : E(X) \rightarrow E_\delta(X_\delta)$ . For a sufficient number of degrees of freedom, the operator is invertible. For any  $x \in E(X)$ ,

$$\|x\|_{\text{DG}} \leq \|(E - E_\delta)x\|_{\text{DG}} + \|E_\delta x\|_{\text{DG}}$$

and the convergence of  $E - E_\delta$  in the operator norm implies that for a sufficient number of degrees of freedom,  $\Lambda_\delta^{-1}$  is bounded. Let us then introduce the operators

$$(25) \quad \widehat{T} = T|_{E(X)} \quad \text{and} \quad \widehat{T}_\delta = \Lambda_\delta^{-1} T_\delta \Lambda_\delta,$$

both defined on the spaces  $E(X) \rightarrow E(X)$ . We consider the case where  $\Gamma$ , introduced in (19), contains a single eigenvalue  $\mu$  of  $T$ , with multiplicity  $n$ . Theorem 2 then implies (see [ABP06] for the details) that there exist  $\mu_{\delta j}$ ,  $j = 1, \dots, n$  that converge towards  $\mu$ . For every  $\mu_{\delta j}$  there exists, then, an  $x_j \in E(X)$  such that

$$\widehat{T}_\delta x_j = \mu_{\delta j} x_j.$$

Let now  $T'$  and  $T'_\delta$  be the adjoint operators to  $T$  and  $T_\delta$ , and let  $E'$  and  $E'_\delta$  be the associated spectral projectors. Furthermore, consider a  $y \in E'(X)$  such that  $(x, y) = 1$ : since for all  $x \in E(X)$ ,  $(T - \mu)x = 0$  (since all eigenvalues have ascent one), we have

$$\begin{aligned}\mu - \mu_{\delta j} &= \langle (\mu - \widehat{T}_\delta)x, y \rangle \\ &= \langle (T - \widehat{T}_\delta)x, y \rangle \\ &= \langle (T - \Lambda_\delta^{-1}T_\delta E_\delta)x, y \rangle\end{aligned}$$

Note now that  $\Lambda_\delta^{-1}E_\delta|_{E(X)} = I$  and that  $T_\delta$  and  $E_\delta$  commute on  $E(X)$ , thus

$$\begin{aligned}\mu - \mu_{\delta j} &= \langle (\Lambda_\delta^{-1}E_\delta)(T - T_\delta)x, y \rangle \\ &= \langle (T - T_\delta)x, y \rangle + \langle (\Lambda_\delta^{-1}E_\delta - I)(T - T_\delta)x, y \rangle.\end{aligned}$$

We remark that  $\ker((\Lambda_\delta^{-1}E_\delta - I)|_{E(X)}) = \ker(E_\delta)^\perp$ , hence

$$\Lambda_\delta^{-1}E_\delta - I : E(X) \rightarrow \text{im}(E_\delta)^\perp.$$

Using also the fact that  $E'y = y$ , the second term at the right hand side above can be written as

$$\langle (\Lambda_\delta^{-1}E_\delta - I)(T - T_\delta)x, y \rangle = \langle (\Lambda_\delta^{-1}E_\delta - I)(T - T_\delta)x, (E' - E'_\delta)y \rangle.$$

As already shown  $\Lambda_\delta^{-1}$  is bounded for a sufficient number of degrees of freedom, and so is  $E_\delta$ . Let us now choose  $\|x\| = \|y\| = 1$ : we have

$$|\langle (\Lambda_\delta^{-1}E_\delta - I)(T - T_\delta)x, y \rangle| \leq C \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)x\|_{\text{DG}} \sup_{\substack{y \in E'(X) \\ \|y\|=1}} \|(T' - T'_\delta)y\|_{\text{DG}}$$

where we have used (23) for the adjoint spectral projectors. We introduce two orthonormal bases  $\{\varphi_i\}_i$  and  $\{\varphi'_j\}_j$  for  $E(X)$  and  $E'(X)$  respectively. Since the spaces are finite dimensional, i.e.,  $n = \dim(E(X)) = \dim(E'(X)) < \infty$ , there exists a constant  $C > 0$  such that

$$\begin{aligned}\langle (T - T_\delta)x, y \rangle &\leq \sup_{\|x\|=\|y\|=1} |\langle (T - T_\delta)x, y \rangle| \\ &\leq C \sum_{i,j=1}^n |\langle (T - T_\delta)\varphi_i, \varphi'_j \rangle|,\end{aligned}$$

where  $C$  depends on  $n$ . We conclude that

$$(26) \quad |\mu - \mu_{\delta j}| \leq C \left( \sum_{i,j=1}^n \langle (T - T_\delta)\varphi_i, \varphi'_j \rangle + \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)x\|_{\text{DG}} \sup_{\substack{y \in E'(X) \\ \|y\|=1}} \|(T' - T'_\delta)y\|_{\text{DG}} \right),$$

**Remark 2.** The above estimate (26) holds since we have considered a case where all eigenvalues have ascent one. If this were not the case, one would find that

$$|\mu - \frac{1}{n} \sum_{j=1}^n \mu_{\delta j}| \leq C \left( \sum_{i,j=1}^n \langle (T - T_\delta)\varphi_i, \varphi'_j \rangle + \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)x\|_{\text{DG}} \sup_{\substack{y \in E'(X) \\ \|y\|=1}} \|(T' - T'_\delta)y\|_{\text{DG}} \right),$$

and

$$\max_{j=1, \dots, n} |\mu - \mu_{\delta j}| \leq C \left( \sum_{i,j=1}^n \langle (T - T_\delta)\varphi_i, \varphi'_j \rangle + \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)x\|_{\text{DG}} \sup_{\substack{y \in E'(X) \\ \|y\|=1}} \|(T' - T'_\delta)y\|_{\text{DG}} \right)^{1/\alpha},$$

$\alpha$  being the ascent of the eigenvalue  $\mu$ , see [DNR78b, CL91].



4.3.2. *Convergence of the eigenvalues for the SIP method.* We now restrict ourselves to the symmetric interior penalty method and consider the fact that our operator is self-adjoint: then,  $T' = T$ ,  $T_\delta' = T_\delta$ , and (26) reads

$$(27) \quad |\mu - \mu_{\delta j}| \leq C \left( \sum_{i,j=1}^n \langle (T - T_\delta)\varphi_i, \varphi_j \rangle + \sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)v\|_{\text{DG}}^2 \right).$$

At this stage, the goal is in bounding the first term at the right hand side of the inequality by something quadratic in nature, to show that it converges as fast as the second term. This is where the adjoint consistency of the SIP method is crucial. Let  $y \in E(X)$  with  $\|y\| = 1$  and let  $\psi \in X$  be the solution to the adjoint problem

$$(28) \quad \begin{aligned} L\psi &= y \text{ in } \Omega \\ B\psi &= 0 \text{ on } \partial\Omega. \end{aligned}$$

This implies  $Ty = \psi$ , hence

$$(29) \quad \inf_{v_\delta \in X_\delta} \|\psi - v_\delta\|_{\text{DG}} \leq C \sup_{\substack{x \in E(X) \\ \|x\|=1}} \inf_{v_\delta \in X_\delta} \|x - v_\delta\|_{\text{DG}}.$$

Consider then  $x, y \in E(X)$ , with  $\|x\| = \|y\| = 1$ :

$$\begin{aligned} \langle (T - T_\delta)x, y \rangle &= \langle (T - T_\delta)x, L\psi \rangle \\ &= a_\delta((T - T_\delta)x, \psi) \\ &= a_\delta((T - T_\delta)x, \psi - v_\delta). \end{aligned}$$

Finally, by the continuity of the bilinear form, the quasi optimality of the discontinuous Galerkin method, and using (29), we conclude that

$$\begin{aligned} |\langle (T - T_\delta)x, y \rangle| &\leq C \|(T - T_\delta)x\|_{\text{DG}} \|\psi - v_\delta\|_{\text{DG}} \\ &\leq C \sup_{\substack{x \in E(X) \\ \|x\|=1}} \inf_{v_\delta \in X_\delta} \|x - v_\delta\|_{\text{DG}}^2. \end{aligned}$$

Since clearly

$$\sup_{\substack{x \in E(X) \\ \|x\|=1}} \|(T - T_\delta)x\|_{\text{DG}}^2 \leq C \sup_{\substack{x \in E(X) \\ \|x\|=1}} \inf_{v_\delta \in X_\delta} \|x - v_\delta\|_{\text{DG}}^2,$$

from (27) we conclude that

$$\max_{j=1, \dots, n} |\mu - \mu_{\delta j}| \leq C \sup_{\substack{x \in E(X) \\ \|x\|=1}} \inf_{v_\delta \in X_\delta} \|x - v_\delta\|_{\text{DG}}^2.$$

Since for every eigenvalue  $\mu$  of  $T$ ,  $1/\mu$  is an eigenvalue of (4), we have proven the following theorem.

**Theorem 3.** *Let  $\lambda$  be an eigenvalue of problem (4) with associated eigenspace  $U = \text{span}(u_1, \dots, u_n)$ , with  $\|u_i\| = 1$  for  $i = 1, \dots, n$ . Then, there exist  $n$  eigenvalue-eigenfunction pairs  $\{(\lambda_{\delta j}, u_{\delta j})\}_j$  of the finite dimensional problem (12) such that for all  $j = 1, \dots, n$*

$$\begin{aligned} \min_{u \in U} \|u - u_{\delta j}\|_{\text{DG}} &\lesssim \sup_{u \in U} \inf_{v_\delta \in X_\delta} \|u - v_\delta\|_{\text{DG}} \\ |\lambda - \lambda_{\delta j}| &\lesssim \sup_{u \in U} \inf_{v_\delta \in X_\delta} \|u - v_\delta\|_{\text{DG}}. \end{aligned}$$

Furthermore, if the numerical solutions are obtained with the SIP method,

$$|\lambda - \lambda_{\delta j}| \lesssim \sup_{u \in U} \inf_{v_{\delta} \in X_{\delta}} \|u - v_{\delta}\|_{\text{DG}}^2.$$

Finally, there are no spurious numerical eigenvalues or eigenvectors.

Given the approximation properties of the hp method and considering that all eigenfunctions of (4) belong to the space  $\mathcal{J}_{\gamma}^{\infty}(\Omega)$  for a  $\gamma > d/2$ , we can also provide the following corollary.

**Corollary 11.** *Let  $\lambda, u, U, \lambda_{\delta j}$ , and  $u_{\delta j}$  be defined as in Theorem 3 and let  $N = \dim(X_{\delta})$ . Then, there exist  $C, b > 0$  such that for all  $u \in U$ , for all  $j = 1, \dots, n$*

$$\begin{aligned} \|u - u_{\delta j}\|_{\text{DG}} &\leq C e^{-bN^{1/(d+1)}} \\ |\lambda - \lambda_{\delta j}| &\leq C e^{-bN^{1/(d+1)}} \end{aligned}$$

Furthermore, if the numerical solutions are obtained with the SIP method,

$$|\lambda - \lambda_{\delta j}| \leq C e^{-2bN^{1/(d+1)}}.$$

## 5. NUMERICAL RESULTS

In this section, we perform some numerical experiments on the linear eigenvalue problem of finding  $(\lambda, u) \in \mathbb{R} \times H^1(\Omega)$  such that  $\|u\|_{L^2(\Omega)} = 1$  and

$$(30) \quad \begin{aligned} (-\Delta + V)u &= \lambda u \text{ in } \Omega \\ u &= 0 \text{ on } \partial\Omega. \end{aligned}$$

The domain  $\Omega$  is the  $d$ -dimensional cube with unitary edge  $(-1/2, 1/2)^d$ , and  $V$  is a potential with a singularity at the origin that will be specified in the single cases. Since no exact solution is available, every numerical solution is compared with the solution obtained at a higher degree of refinement than those presented.

In all cases, the mesh is isotropically and geometrically refined around the origin, with a geometric refinement ratio  $\sigma = 1/2$ . All elements are axiparallel  $d$ -dimensional cubes. This means that, introducing the refinement layers  $\Omega_j, j = 1, \dots, \ell$ , such that for all  $K \in \Omega_j$ ,

$$\inf_{x \in K} \|x\|_{\ell^{\infty}} = \sigma^{j+1} \quad j = 1, \dots, \ell - 1$$

we have

$$|K| = h_K^d = \sigma^{(j+1)d}.$$

Furthermore, the elements in  $\Omega_{\ell}$  have a vertex on the singularity. The polynomial slope  $\mathfrak{s}$ , defined as the parameter such that for all  $v_{\delta} \in X_{\delta}$ , if an element  $K \in \Omega_j$  then

$$v_{\delta}|_K \in \mathbb{Q}_{p_j}(K),$$

with

$$p_j = p_0 + \lfloor \mathfrak{s}(\ell - j) \rfloor$$

is instead variable between experiments, and it is one of the main parameters whose role in the approximation we investigate. The base polynomial degree is fixed at  $p_0 = 1$ .

All the simulations are obtained with C++ code based on the library `deal.II` [ABD<sup>+</sup>17]. Furthermore, we use PETSc [BAA<sup>+</sup>17] for the solution of algebraic linear systems, and SLEPc [HRV05] for the solution of the algebraic eigenvalue problem. The actual methods used will vary between the two and the three dimensional cases, and will be specified in the respective sections. The boundary conditions are imposed weakly, as is customary in the framework of discontinuous Galerkin methods, so no spurious eigenvalue is introduced, as shown in Section 4.

TABLE 1. Estimated coefficients. Potential:  $r^{-1/2}$ 

$\mathfrak{s}$	$b_{L^2}$	$b_{\text{DG}}$	$b_{L^\infty}$	$b_\lambda$
0.125	0.73	0.78	0.78	1.34
0.25	0.9	0.89	0.86	1.12
0.5	1.07	1	1	1.19

The results we will show in the following concern the estimation of the DG,  $L^2(\Omega)$  and  $L^\infty(\Omega)$  norms of the error, and of the difference between the computed and the “exact” eigenvalue. Furthermore, we will try to estimate the constants  $b_X$  such that

$$\|u - u_\delta\|_X \leq C_X \exp(-b_X N^{1/(d+1)}),$$

for  $X = \text{DG}, L^2(\Omega), L^\infty(\Omega)$ , and

$$|\lambda - \lambda_\delta| \leq C_\lambda \exp(-b_\lambda N^{1/(d+1)}).$$

Here,  $u_\delta \in X_\delta$  (resp.  $\lambda_\delta \in \mathbb{R}$ ) is the numerical eigenfunction (resp. eigenvalue) computed with  $\dim(X_\delta) = N$  and  $u$  (resp.  $\lambda$ ) is the exact one.

We start by illustrating the results obtained in the framework of a two dimensional approximation.

**5.1. Two dimensional case.** We solve problem (30) with  $d = 2$  on a mesh built as shown in Figure 2. An example of a numerically computed eigenfunction is shown in Figure 3a. We can see the combination of the effect of the laplacian with homogeneous Dirichlet boundary conditions and of the potential. The cusp introduced by the potential is partially hidden by the rest of the solution; in Figure 3b, where a close up of the solution over a line is represented, we can see it more clearly.

We consider three different potentials, given by  $V(x) = r^{-\alpha}$ , with  $\alpha \in \{1/2, 1, 3/2\}$ . Clearly, the bigger the exponent  $\alpha$ , the lower the regularity of the exact solution. In particular, from the point of view of classical Sobolev spaces, denoting  $u_\alpha$  as the solution of

$$\begin{aligned} (-\Delta + r^{-\alpha})u_\alpha &= \lambda_\alpha u_\alpha \text{ in } \Omega \\ u_\alpha &= 0 \text{ on } \partial\Omega, \end{aligned}$$

we have  $u_\alpha \in H^{3-\alpha-\xi}(\Omega)$ , for any  $\xi > 0$ . In particular, the problem with  $\alpha = 3/2$  roughly corresponds to a two dimensional elliptic problem in a domain with a crack, see [CD02]. When considering weighted Sobolev spaces, we have

$$(31) \quad u_\alpha \in \mathcal{J}_{3-\alpha-\xi}^\varpi(\Omega),$$

again for any  $\xi > 0$ .

From the algebraic point of view, the eigenpairs are computed using a Krylov-Schur method [Ste02]. Furthermore, a shift and invert spectral transformation is used to precondition and speed up computations. Due to the relatively small size of the problems we consider here, the linear system introduced by the shift and invert spectral transformation is solved via an LU decomposition. When considering the problem set in three dimensions, we will see how to deal with problems with more degrees of freedom, where memory availability becomes a concern.

**5.1.1. Analysis of the results.** The results on the error for the potential  $V(x) = r^{-1/2}$  are shown in Figure 4, and the estimated coefficients are given in Table 1. Similarly, when the potential is given by  $V(x) = r^{-1}$  the error curves are in Figure 5, with coefficients  $b_X$  in Table 2, and the case  $V(x) = r^{-3/2}$  is reported in Figure 6 and Table 3.

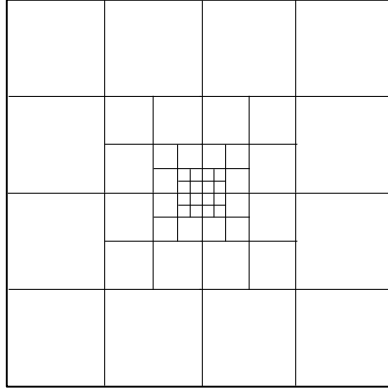
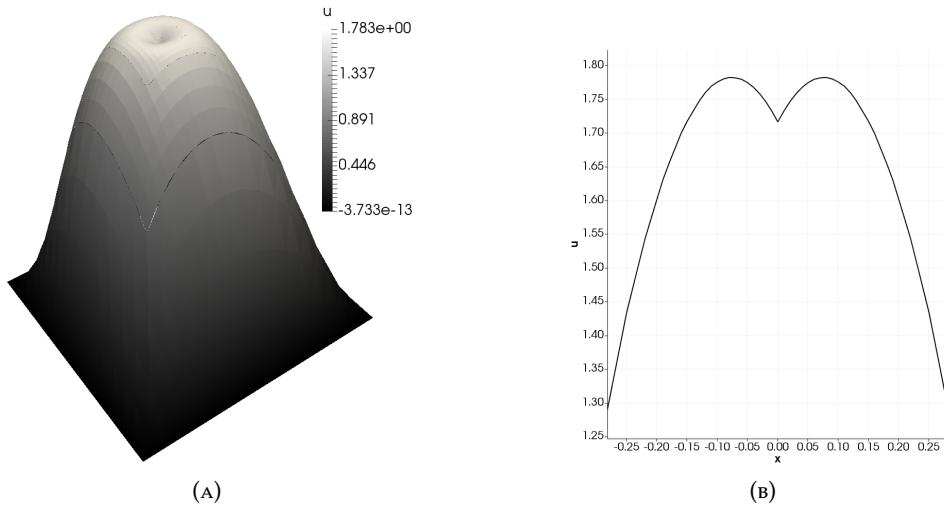

 FIGURE 2. Example of a two dimensional mesh, with  $\ell = 5$ 


FIGURE 3. Numerical solution to (30) with  $V(x) = r^{-1}$ . Figure a: representation vertically not to scale; the separation between some elements is an artifact of the visualization on grids with hanging nodes. Figure b: close up around the singularity of the function  $u(\cdot, 0)$ , i.e., of  $u$  on the line  $\{y = 0\}$ .

 TABLE 2. Estimated coefficients. Potential:  $r^{-1}$ 

$\mathfrak{s}$	$b_{L^2}$	$b_{\text{DG}}$	$b_{L^\infty}$	$b_\lambda$
0.125	0.76	0.75	0.71	1.37
0.25	0.87	0.85	0.84	1.12
0.5	0.72	0.72	0.63	0.64

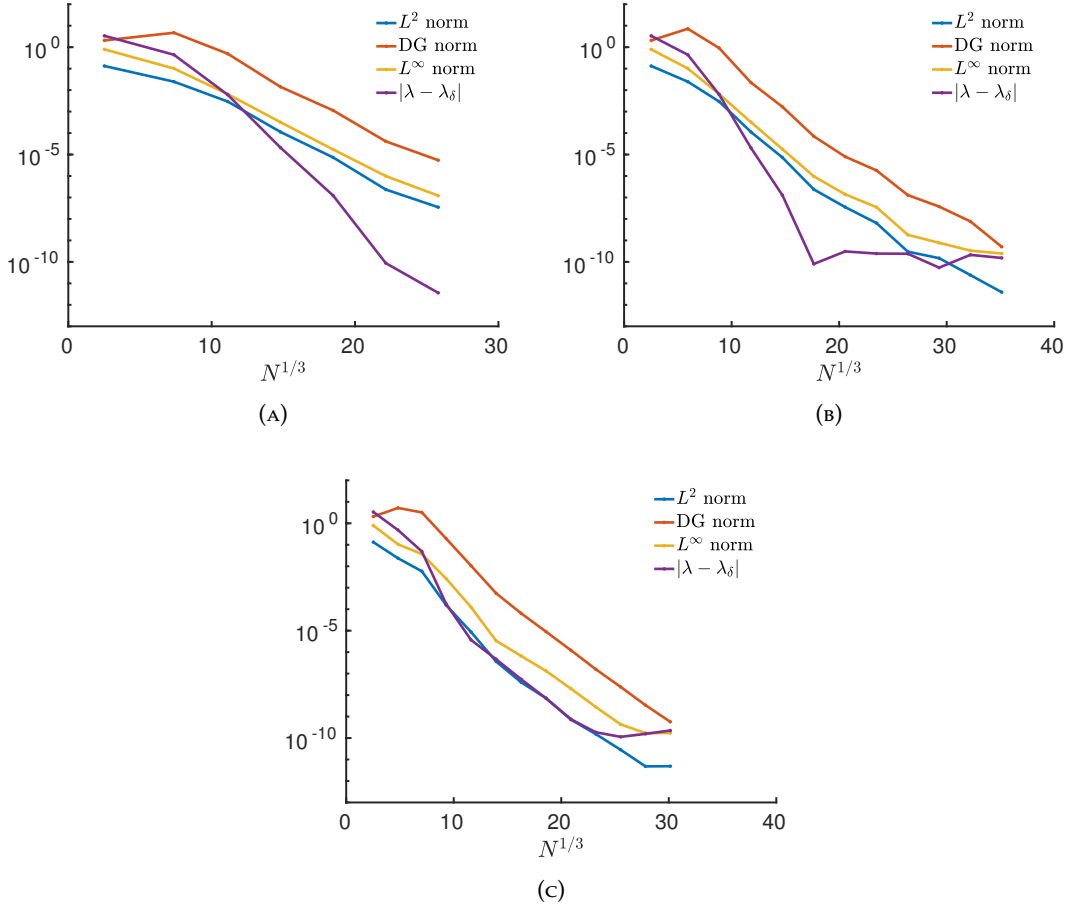


FIGURE 4. Errors for the numerical solution with potential  $V(x) = r^{-1/2}$ . Polynomial slope:  $\mathfrak{s} = 1/8$  in Figure a;  $\mathfrak{s} = 1/4$  in Figure b and  $\mathfrak{s} = 1/2$  in Figure c.

TABLE 3. Estimated coefficients. Potential:  $r^{-3/2}$

$\mathfrak{s}$	$b_{L^2}$	$b_{DG}$	$b_{L^\infty}$	$b_\lambda$
0.062	0.49	0.48	0.47	0.85
0.125	0.61	0.59	0.64	1.09
0.25	0.6	0.53	0.42	0.48

We can clearly see, that in many cases the error reaches at some point a plateau; we estimate the coefficients  $b_X$  by linear regression on the points before the plateau. This will be done for all subsequent potentials. Furthermore, as expected, the less regular the potential, the slower the convergence of the numerical solution.

Two phenomena are less expected from the point of view of the theory. The first one is the emergence of a plateau at relatively high values compared to the machine epsilon. Through the

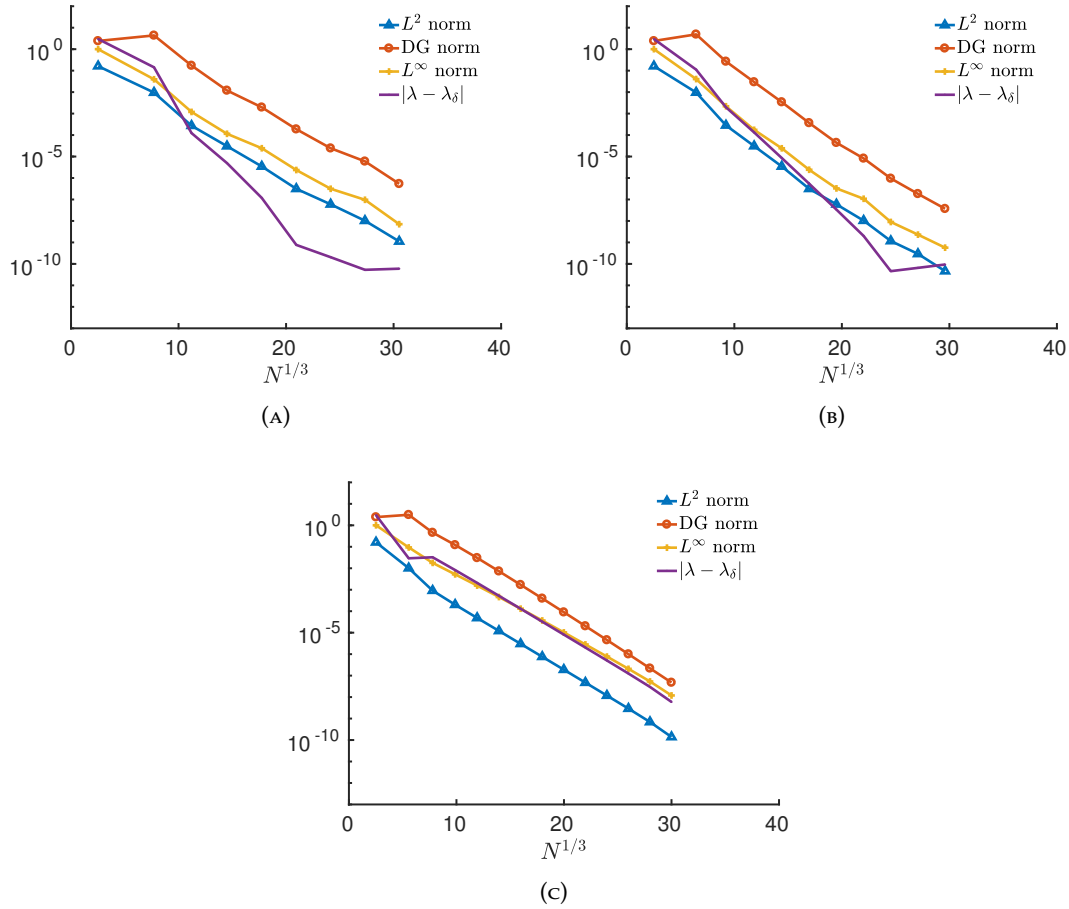


FIGURE 5. Errors for the numerical solution with potential  $V(x) = r^{-1}$ . Polynomial slope:  $s = 1/8$  in Figure a;  $s = 1/4$  in Figure b and  $s = 1/2$  in Figure c.

choice of different algebraic scheme, we can see that we get a lower plateau: this is an indication that the dominating error at the points where it is not converging to zero is the algebraic one. The fact that matrices arising from the  $hp$  method are ill conditioned explains the size of the algebraic error. In practical applications, the fact that a relative error of approximately  $10^{-12}$  can be reached should be sufficient.

The second “unexpected phenomenon” is evident when looking at Figures 4c, 5b, 6b, and 6c. We remark that, after an initial part where the eigenvalue converges faster than the other norms of the error, its rate of convergence then stabilizes to the same rate of the other norms. This can be shown [CCM10] to be dependent on the quadrature formula employed. When using a higher degree quadrature formula, the highest rate for the eigenvalue error is recovered, see Figure 7 and Table 4, obtained with a higher quadrature formula and compare them with Figure 5b and Table 2. As a side effect of a higher quadrature order, the plateau is raised.

In practice, one has to quite carefully balance computational cost, conditioning of the matrix, and speed of convergence. The usefulness of this numerical experiments lies therefore not only

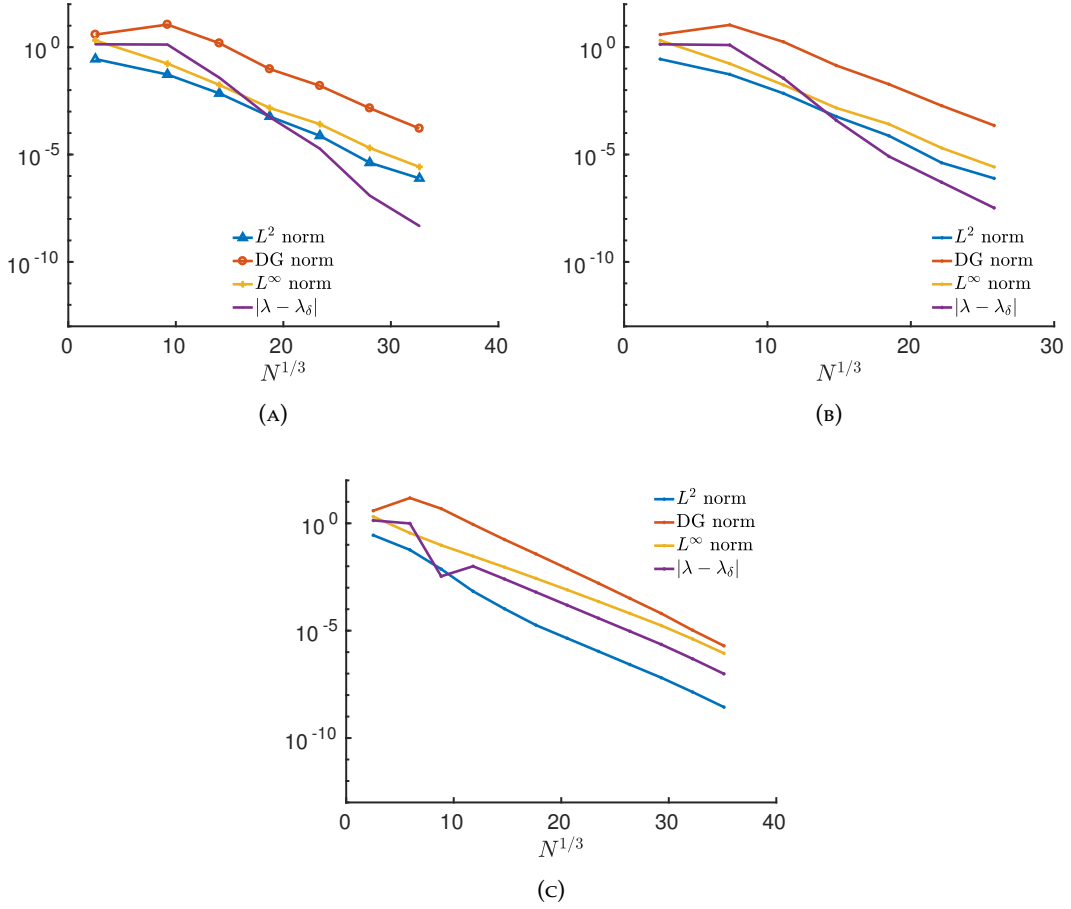


FIGURE 6. Errors for the numerical solution with potential  $V(x) = r^{-3/2}$ . Polynomial slope:  $\mathfrak{s} = 1/16$  in Figure a;  $\mathfrak{s} = 1/8$  in Figure b and  $\mathfrak{s} = 1/4$  in Figure c.

TABLE 4. Estimated coefficients. Potential:  $r^{-1}$ , high degree quadrature formula

$\mathfrak{s}$	$b_{L^2}$	$b_{DG}$	$b_{L^\infty}$	$b_\lambda$
0.25	0.89	0.81	0.85	1.5

in the fact that we verify our theoretical results and we see the impact of components of the error we did not account for in the theoretical analysis, but also in the fact that we see, practically, how the parameters affect the simulation for different exact solutions. Since by asymptotic analysis we can see, locally and *a priori*, how the solution of a problem behaves, this gives an indication on how to construct and locally *a priori* optimize the *hp* spaces.

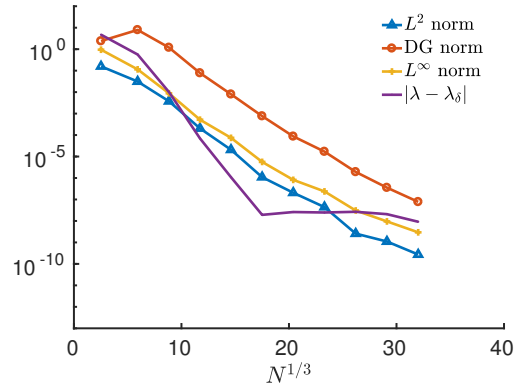


FIGURE 7. Errors of the numerical solution for  $V(x) = r^{-1}$  and a high degree quadrature formula. Polynomial slope  $\varepsilon = 0.25$ .

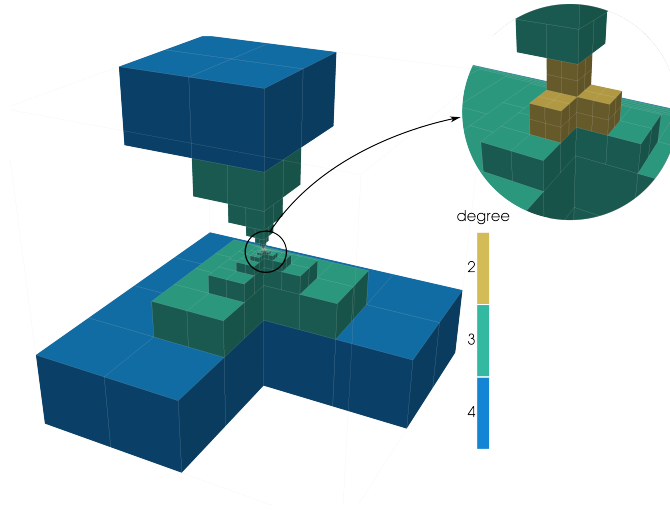


FIGURE 8. Example mesh for the three dimensional approximation

**5.2. Three dimensional case.** In the three dimensional case, we replicate the setting introduced in Section 5.1. In this case,  $\Omega = (-1/2, 1/2)^3$ . Note that the regularity of the solution of

$$\begin{aligned} (-\Delta + r^{-\alpha})u_\alpha &= \lambda_\alpha u_\alpha \text{ in } \Omega \\ u_\alpha &= 0 \text{ on } \partial\Omega, \end{aligned}$$

scales differently with respect to  $\alpha$ , if compared to the two dimensional case. Specifically, we have

$$u_\alpha \in H^{7/2-\alpha-\xi}(\Omega)$$

and

$$u_\alpha \in \mathcal{J}_{7/2-\alpha-\xi}^\infty(\Omega),$$

for any  $\xi > 0$ .



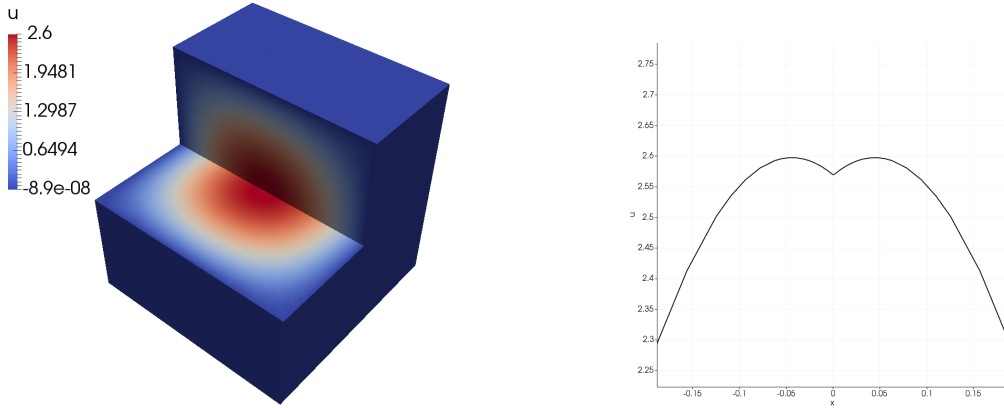


FIGURE 9. Numerical solution in the three dimensional case: solution in the cube, left, and close up near the origin of the restriction to the line  $\{y = z = 0\}$ , right

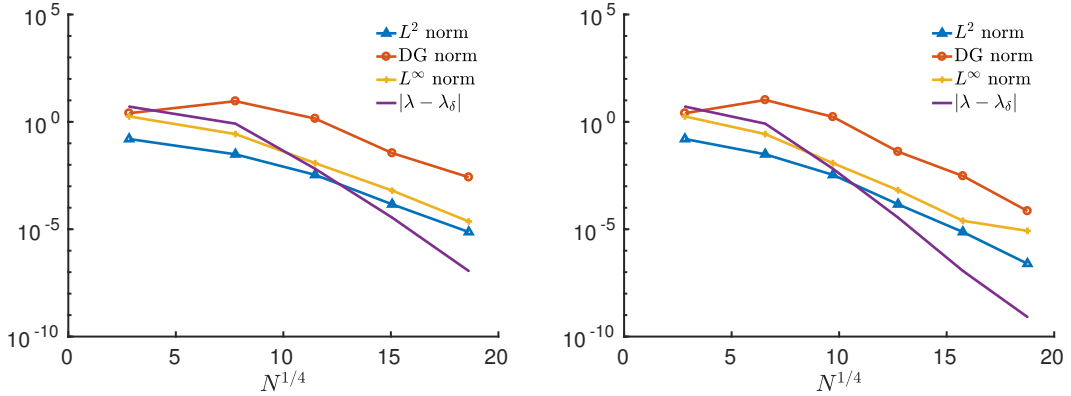


FIGURE 10. Errors of the numerical solution for  $V(x) = r^{-1/2}$ . Polynomial slope  $s = 1/8$ , left and  $s = 1/4$ , right.

The mesh is built in a tensor product way as in Section 5.1, with refinement ratio  $\sigma = 1/2$ . A representation of a mesh is given in Figure 8. The numerical solution for  $V(x) = r^{-1}$  is shown in Figure 9.

From the algebraic point of view, the assembled matrices are bigger in size and less sparse, thus a direct LU method is less feasible than in the previous case (up to completely unfeasible for the simulations with a high number of degrees of freedom). Hence, we turn to iterative methods, and try to employ an algebraic eigenvalue method that is not too sensible to the error introduced by the linear solver. Therefore, the search for the eigenvalues is done with a Jacobi-Davidson method [SV96]. Internally, we employ a biconjugate gradient stabilized method (BiCGS, [vdV92, SvdVF94]) as a linear solver, with simple Jacobi preconditioner. The tolerance for the linear solver is set at  $10^{-6}$ , while the tolerance of the Jacobi-Davidson method is set at  $10^{-8}$ .

5.2.1. *Analysis of the results.* Results for  $V(x) = r^{-1/2}$  are given in Figure 10 and Table 5, while the case  $V(x) = r^{-1}$  is analyzed in Figure 11 and Table 6 and the errors and estimates when

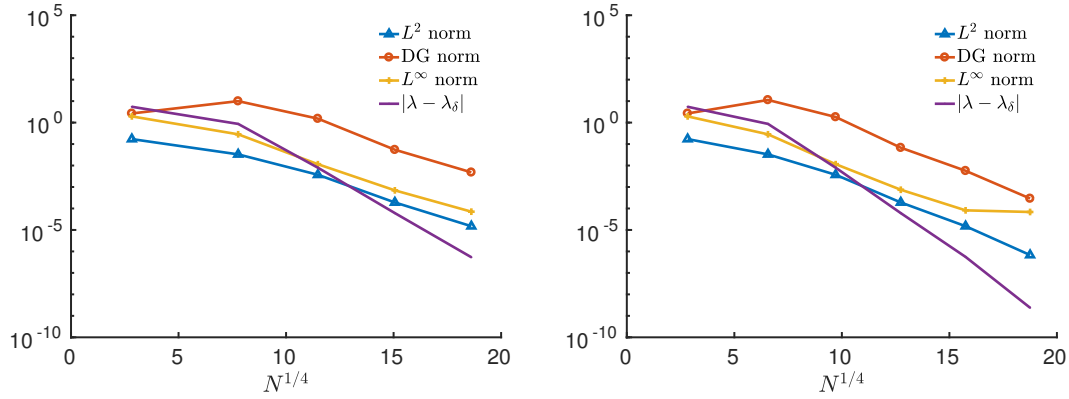


FIGURE 11. Errors of the numerical solution for  $V(x) = r^{-1}$ . Polynomial slope  $\mathfrak{s} = 1/8$ , left and  $\mathfrak{s} = 1/4$ , right.

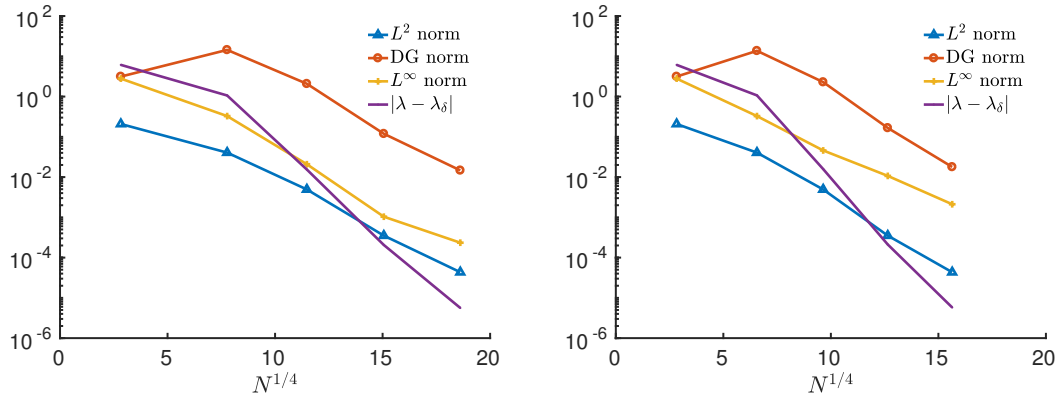


FIGURE 12. Errors of the numerical solution for  $V(x) = r^{-3/2}$ . Polynomial slope  $\mathfrak{s} = 1/8$ , left and  $\mathfrak{s} = 1/4$ , right.

TABLE 5. Estimated coefficients. Potential:  $r^{-1/2}$

$\mathfrak{s}$	$b_{L^2}$	$b_{DG}$	$b_{L^\infty}$	$b_\lambda$
0.125	0.78	0.78	0.86	1.46
0.25	0.97	0.99	0.89	1.72

TABLE 6. Estimated coefficients. Potential:  $r^{-1}$

$\mathfrak{s}$	$b_{L^2}$	$b_{DG}$	$b_{L^\infty}$	$b_\lambda$
0.125	0.72	0.73	0.77	1.32
0.25	0.89	0.88	0.71	1.61

TABLE 7. Estimated coefficients. Potential:  $r^{-3/2}$ 

$s$	$b_{L^2}$	$b_{DG}$	$b_{L^\infty}$	$b_\lambda$
0.125	0.64	0.65	0.69	1.13
0.25	0.77	0.75	0.55	1.35

$V(x) = r^{-3/2}$  are shown in Figure 12 and Table 7. The three dimensional approximation has far more degrees of freedom than the two dimensional one for a given level of refinement  $\ell$ , thus the results we show have lower levels of refinement than the two dimensional ones. This is partially balanced by the fact that the solutions are more regular, but the errors are still obviously higher than those of the two dimensional case, at the same number of degrees of freedom. In the three dimensional case, we do not see a great effect neither of the algebraic error nor of the quadrature formulas. The coefficients  $b_\lambda$  listed in Tables 5 to 7 are almost the double of the respective coefficients  $b_{DG}$ ; thus, if the effect of the quadrature error is present, it is nonetheless negligible compared to other sources of error for the quite comprehensive potentials and polynomial slopes considered in this experiments.

## REFERENCES

- [ABD<sup>+</sup>17] D. Arndt, W. Bangerth, D. Davydov, T. Heister, L. Heltai, M. Kronbichler, M. Maier, J.-P. Pelteret, B. Turcksin, and D. Wells, *The deal.II library, version 8.5*, Journal of Numerical Mathematics **25** (2017), no. 3, 137–145.
- [ABP06] P. F. Antonietti, A. Buffa, and I. Perugia, *Discontinuous Galerkin approximation of the Laplace eigenproblem*, Computer Methods in Applied Mechanics and Engineering **195** (2006), no. 25-28, 3483–3503.
- [Arn82] D. N. Arnold, *An Interior Penalty Finite Element Method with Discontinuous Elements*, SIAM Journal on Numerical Analysis **19** (1982), no. 4, 742–760.
- [BAA<sup>+</sup>17] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, K. Rupp, B. F. Smith, S. Zampini, H. Zhang, and H. Zhang, *PETSc Web page*, <http://www.mcs.anl.gov/petsc>, 2017.
- [CCM10] E. Cancès, R. Chakir, and Y. Maday, *Numerical Analysis of Nonlinear Eigenvalue Problems*, Journal of Scientific Computing **45** (2010), no. 1-3, 90–117.
- [CD02] M. Costabel and M. Dauge, *Crack Singularities for General Elliptic Systems*, Mathematische Nachrichten **235** (2002), no. 1, 29–49.
- [CDN10] M. Costabel, M. Dauge, and S. Nicaise, *Mellin Analysis of Weighted Sobolev Spaces with Nonhomogeneous Norms on Cones*, Around the Research of Vladimir Maz'ya I, Springer New York, 2010, pp. 105–136.
- [CDN12] ———, *Analytic Regularity for Linear Elliptic Systems in Polygons and Polyhedra*, Mathematical Models and Methods in Applied Sciences **22** (2012), no. 08, 1250015.
- [CDS05] M. Costabel, M. Dauge, and C. Schwab, *Exponential convergence of hp-FEM for Maxwell equations with weighted regularization in polygonal domains*, Mathematical Models and ... **15** (2005), no. 4, 575–622.
- [CL91] P. G. Ciarlet and J.-L. Lions, *Handbook of numerical analysis. Vol. II*, North-Holland, Amsterdam, 1991, Finite element methods. Part 1. MR 1115235
- [CR73] M. Crouzeix and P.-A. Raviart, *Conforming and nonconforming finite element methods for solving the stationary Stokes equations I*, Revue française d'automatique informatique recherche opérationnelle. Mathématique **7** (1973), no. R3, 33–75.
- [CS98] B. Cockburn and C.-W. Shu, *The Local Discontinuous Galerkin Method for Time-Dependent Convection-Diffusion Systems*, SIAM Journal on Numerical Analysis **35** (1998), no. 6, 2440–2463.
- [DE12] D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Mathématiques et Applications, vol. 69, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [DNR78a] J. Descloux, N. Nassif, and J. Rappaz, *On spectral approximation. I. The problem of convergence*, RAIRO Analyse Numérique **12** (1978), no. 2, 97–112, iii.
- [DNR78b] ———, *On spectral approximation. II. Error estimates for the Galerkin method*, RAIRO Analyse Numérique **12** (1978), no. 2, 113–119, iii.
- [ES97] Y. V. Egorov and B.-W. Schulze, *Pseudo-Differential Operators, Singularities, Applications*, Birkhäuser Basel, Basel, 1997.
- [GB86a] W. Gui and I. Babuška, *The h, p and h-p versions of the finite element method in 1 dimension. Part I. The Error Analysis of the p-Version*, Numerische Mathematik **612** (1986), 577–612.

- [GB86b] ———, *The  $h$ ,  $p$  and  $h$ - $p$  versions of the finite element method in 1 dimension. Part II. The Error analysis of the  $h$ - and  $h - p$  versions.*, *Numerische Mathematik* **49** (1986), no. 6, 613–657.
- [GB86c] ———, *The  $h$ ,  $p$  and  $h$ - $p$  versions of the finite element method in 1 dimension. Part III. The Adaptive  $h$ - $p$  Version*, *Numerische Mathematik* **683** (1986), 659–683.
- [GB86d] B. Guo and I. Babuška, *The  $h$ - $p$  version of the finite element method - Part 1: The basic approximation results*, *Computational Mechanics* **1** (1986), no. 1, 21–41.
- [GB86e] ———, *The  $h$ - $p$  version of the finite element method - Part 2: General results and applications*, *Computational Mechanics* **1** (1986), no. 3, 203–220.
- [HRV05] V. Hernandez, J. E. Roman, and V. Vidal, *SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems*, *ACM Transactions on Mathematical Software* **31** (2005), no. 3, 351–362.
- [HW08] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods*, *Texts in Applied Mathematics*, vol. 54, Springer New York, 2008.
- [KMR97] V. Kozlov, V. G. Maz'ya, and J. Rossmann, *Elliptic boundary value problems in domains with point singularities*, American Mathematical Society, 1997.
- [Kon67] V. A. Kondrat'ev, *Boundary value problems for elliptic equations in domains with conical or angular points*, *Trudy Moskovskogo Matematicheskogo Obščestva* **16** (1967), 209–292. MR 0226187
- [MR10] V. G. Maz'ya and J. Rossmann, *Elliptic Equations in Polyhedral Domains*, *Mathematical Surveys and Monographs*, vol. 162, American Mathematical Society, apr 2010.
- [Nit72] J. Nitsche, *On Dirichlet problems using subspaces with nearly zero boundary conditions*, *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, Elsevier, 1972, pp. 603–627.
- [RH73] W. H. Reed and T. Hill, *Triangular mesh methods for the neutron transport equation*, Tech. report, Los Alamos Scientific Lab., (USA), 1973.
- [Riv08] B. Rivière, *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*, Society for Industrial and Applied Mathematics, jan 2008.
- [SSW13a] D. Schötzau, C. Schwab, and T. P. Wihler,  *$hp$ -dGFEM for second order elliptic problems in polyhedra. II: Exponential convergence*, *SIAM Journal on Numerical Analysis* **51** (2013), no. 4, 2005–2035.
- [SSW13b] D. Schötzau, C. Schwab, and T. Wihler,  *$hp$ -dGFEM for Second-Order Elliptic Problems in Polyhedra I: Stability on Geometric Meshes*, *SIAM Journal on Numerical Analysis* **51** (2013), no. 3, 1610–1633.
- [SSW16] D. Schötzau, C. Schwab, and T. P. Wihler,  *$hp$ -dGFEM for second-order mixed elliptic problems in polyhedra*, *Mathematics of Computation* **85** (2016), no. 299, 1051–1083.
- [Ste02] G. W. Stewart, *A Krylov–Schur Algorithm for Large Eigenproblems*, *SIAM Journal on Matrix Analysis and Applications* **23** (2002), no. 3, 601–614.
- [SV96] G. L. Sleijpen and H. A. Van der Vorst, *A Jacobi–Davidson Iteration Method for Linear Eigenvalue Problems*, *SIAM Journal on Matrix Analysis and Applications* **17** (1996), no. 2, 401–425.
- [SvdVF94] G. L. G. Sleijpen, H. A. van der Vorst, and D. R. Fokkema, *BiCGstab(l) and other hybrid Bi-CG methods*, *Numerical Algorithms* **7** (1994), no. 1, 75–109.
- [SW10] B. Stamm and T. P. Wihler,  *$hp$ -optimal discontinuous Galerkin methods for linear elliptic problems*, *Mathematics of Computation* **79** (2010), 2117–2133.
- [vdV92] H. A. van der Vorst, *Bi-CGSTAB: A Fast and Smoothly Converging Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems*, *SIAM Journal on Scientific and Statistical Computing* **13** (1992), no. 2, 631–644.
- [Whe78] M. F. Wheeler, *An Elliptic Collocation-Finite Element Method with Interior Penalties*, *SIAM Journal on Numerical Analysis* **15** (1978), no. 1, 152–161.