Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

Analyticity and hp discontinuous Galerkin approximation of nonlinear Schrödinger eigenproblems

Y. Maday and C. Marcati

Research Report No. 2019-69
December 2019

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Funding ERC: 810367

ANALYTICITY AND hp DISCONTINUOUS GALERKIN APPROXIMATION OF NONLINEAR SCHRÖDINGER EIGENPROBLEMS

YVON MADAY† AND CARLO MARCATI⋆

ABSTRACT. We study a class of nonlinear eigenvalue problems of Schrödinger type, where the potential is singular on a set of points. Such problems are widely present in physics and chemistry, and their analysis is of both theoretical and practical interest. In particular, we study the regularity of the eigenfunctions of the operators considered, and we propose and analyze the approximation of the solution via an isotropically refined hp discontinuous Galerkin (dG) method.

We show that, for weighted analytic potentials and for up-to-quartic nonlinearities, the eigenfunctions belong to analytic-type non homogeneous weighted Sobolev spaces. We also prove quasi optimal *a priori* estimates on the error of the dG finite element method; when using an isotropically refined hp space the numerical solution is shown to converge with exponential rate towards the exact eigenfunction. In addition, we investigate the role of pointwise convergence in the doubling of the convergence rate for the eigenvalues with respect to the convergence rate of eigenfunctions. We conclude with a series of numerical tests to validate the theoretical results.

CONTENTS

† SORBONNE UNIVERSITÉ, CNRS, UNIVERSITÉ DE PARIS, LABORATOIRE JACQUES-LOUIS LIONS (LJLL), F-75005 PARIS, FRANCE AND INSTITUT UNIVERSITAIRE DE FRANCE

⋆ SEMINAR FOR APPLIED MATHEMATICS, D-MATH, ETH ZURICH, ZÜRICH, SWITZERLAND

E-mail addresses: yvon.maday@upmc.fr, carlo.marcati@sam.math.ethz.ch.

2010 *Mathematics Subject Classification.* 65N25 (Primary), 35J10, 35P30, 65N12, 65N30 (Secondary).

Key words and phrases. Nonlinear Schrödinger equation, analytic regularity, exponential convergence, finite element method.

Yvon Maday acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810367).

1. Introduction

This paper concerns the analysis of an elliptic nonlinear eigenvalue problem and its approximation with an hp discontinuous Galerkin finite element method. Specifically, we consider the problem of finding, in a domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, the smallest eigenvalue and associated eigenfunction (λ, u) such that $\|u\|_{L^2(\Omega)} = 1$ and

$$(-\Delta + V + f(u^2))u = \lambda u, \tag{1}$$

for a (singular) potential V and a nonlinearity f. Problems of this kind correspond to the Euler-Lagrange equations of energy minimization problems and are therefore widely present in physics and chemistry. Equations of the form (1) are also often referred to as nonlinear Schrödinger equations.

Our analysis is centered mainly on potentials that are singular at a set of isolated points; this includes the electric attraction generated by a Coulomb potential, i.e., $V(x) = 1/d(x, \mathfrak{c})$, where $d(\cdot, \cdot)$ is the euclidean distance between two points in $\mathbb{R}^d$, for some fixed point $\mathfrak{c} \in \Omega$, but applies more generally to any potential that, in the vicinity of the singular point, behaves as

$$V(x) \sim \frac{1}{d(x, \mathfrak{c})^\xi}, \tag{2}$$

for a $\xi < 2$. Clearly, V is not very regular in classical Sobolev spaces, thus we cannot expect the solution to be regular in those spaces either. Nonetheless, we can alternatively work in weighted Kondrat'ev-Babuška spaces, prove that the solution is sufficiently regular in these spaces, and thus design an appropriate hp discretization that converges exponentially to the exact solution, see [SSW13].

The nonlinear Schrödinger equation (1) and the weighted spaces are introduced in detail in Section 2. There, we also introduce our basic assumptions on the nonlinearity, which are similar to those introduced in [CCM10], and on the potential V. As the analysis progresses, we will introduce more restrictive hypotheses.

In Section 3, we then prove *a priori* convergence estimates on the eigenvalue and eigenfunction. Even though our focus is on hp methods, most of the proofs are more general. Suppose we consider a simpler h-type finite element method: the proof of Theorem 1 — i.e., convergence and quasi optimality of the numerical solution — holds, since we do not use any specific feature of hp refinement. The proof of convergence of the discontinuous Galerkin method for a nonlinear eigenvalue problem of the form (1) is a new result as far as we are aware. Previous results include the convergence of the discontinuous Galerkin method for linear eigenproblems [ABP06] and the convergence of conforming methods for the nonlinear problem [CCM10]. The main difference with the latter paper is that the discontinuous Galerkin method is not conforming, thus some relations between exact and numerical quantities, e.g., between the exact eigenvalue λ and the numerical one λ_δ, are less straightforward. In general, the convergence and quasi optimality of the numerical eigenvalue–eigenfunction pair proven in Theorem 1 should be readily extendable to any nonconforming symmetric method such that the thesis of Lemma 3, akin to coercivity and continuity of the numerical bilinear form, holds.

In the analysis of the approximation of linear eigenvalue problems by symmetric methods, one is often able to show that the rate of convergence of eigenvalues is double that of eigenfunction, see, e.g., [BO91]. In the latter part of Section 3, we show that if we introduce a conjecture on the local behavior of the solution to linear problems with localised right hand sides, we are able to recover a similar result.

In Section 4, we restrict the analysis to the case of polynomial, up-to-quartic nonlinearities. In this setting, the solutions of problem (1) are analytic in weighted Sobolev spaces: specifically, if the potential is of type (2) and $f(u^2) = |u|^\eta$, $\eta = 1, 2, 3$, then for $\gamma < d/2 + 2 - \xi$ there exist constants C_u and A_u such that for all $k \in \mathbb{N}$

$$\sum_{|\alpha|=k} \|d(x,\mathfrak{c})^{k-\gamma}\partial^\alpha u\|_{L^2(\Omega)} \leq C_u A_u^k k!,$$

where u is the ground state of (1) and $\alpha \in \mathbb{N}_0^d$ is a multi index. This estimate is proven in Theorem 2 and constitutes a result of independent interest. For previous weighted analytic regularity results for elliptic problems, we refer, among others, to [GS06, CDN12] for the linear case, and to [MS19] for the analysis — based on the same arguments as in the present paper — of the (nonlinear) two dimensional Navier-Stokes equations.

As a consequence of the quasi-optimal *a priori* estimates introduced above and of the weighted analytic regularity of the ground state for up-to-quartic nonlinearities, we obtain exponential convergence of the numerical solution computed with the *hp* dG method. This is briefly discussed in Section 5 and presented in Theorem 3.

Finally, in Section 6, we investigate the performance of the scheme in two and three dimensional numerical tests. We confirm our theoretical estimates, while also showing the effect of sources of numerical error that have not been taken into consideration in the theoretical analysis.

2. Statement of the problem and notation

2.1. **Functional setting and notation.** Let $\Omega = (\mathbb{R}/L)^d$ be a periodic d-cube of edge $L < 1$. We use the standard notation for Sobolev spaces $W^{k,p}(\Omega)$, with $W^{k,2}(\Omega) = H^k(\Omega)$ and $W^{0,p}(\Omega) = L^p(\Omega)$. We denote the scalar product in $L^2(\Omega)$ as $(\cdot,\cdot)$ and the norm as $\|u\| = (u,u)$. For two quantities a and b, we write $a \lesssim b$ (respectively $a \gtrsim b$) if there exists $C > 0$ independent from the discretization, such that $a \leq Cb$ (resp. $a \geq Cb$). We write $a \simeq b$ if $a \lesssim b$ and $a \gtrsim b$.

We now recall the definition of the weighted Sobolev spaces, introduced in [Kon67], that will be central to our regularity analysis. Given a set of isolated points $\mathfrak{C} \subset \Omega$, the homogeneous Kontrat'ev-Babuška space $\mathcal{K}^{k,p}_\gamma(\Omega,\mathfrak{C})$ is defined as

$$\mathcal{K}^{k,p}_\gamma(\Omega,\mathfrak{C}) = \{u : r^{|\alpha|-\gamma}\partial^\alpha u \in L^p(\Omega)\ \forall \alpha \in \mathbb{N}^d : |\alpha| \leq k\},$$

where $r = r(x)$ is a smooth function which is, in the vicinity of every point $\mathfrak{c} \in \mathfrak{C}$, equal to the euclidean distance $d(x,\mathfrak{c})$ from the point. The nonhomogeneous Kondrat'ev-Babuška space is defined by

$$\mathcal{J}^{k,p}_\gamma(\Omega,\mathfrak{C}) = \{u \in H^{\lfloor \gamma - d/p \rfloor}(\Omega) : r^{|\alpha|-\gamma}\partial^\alpha u \in L^p(\Omega)\ \forall \alpha \in \mathbb{N}^d : \lfloor \gamma - d/p \rfloor + 1 \leq |\alpha| \leq k\},$$

for $\gamma > d/p$. We define the associated seminorm as $|u|^p_{\mathcal{J}^{k,p}_\gamma} = |u|^p_{\mathcal{K}^{k,p}_\gamma} = \sum_{\alpha=k} \|r^{|\alpha|-\gamma}\partial^\alpha u\|^p_{(\Omega)}$. We also introduce the spaces of regular function with weighted analytic type estimates as

$$\mathcal{K}^{\varpi,p}_\gamma(\Omega,\mathfrak{C}) = \{v \in \mathcal{K}^{\infty,p}_\gamma(\Omega,\mathfrak{C}) : |v|_{\mathcal{K}^{k,p}_\gamma} \leq CA^k k!\ \forall k\},$$

and

$$\mathcal{J}^{\varpi,p}_\gamma(\Omega,\mathfrak{C}) = \{v \in \mathcal{J}^{\infty,p}_\gamma(\Omega,\mathfrak{C}) : |v|_{\mathcal{K}^{k,p}_\gamma} \leq CA^k k!,\ \forall k > \lfloor \gamma - d/p \rfloor\},$$

where $\mathcal{K}^{\infty,p}_\gamma = \bigcap_k \mathcal{K}^{k,p}_\gamma$, $\mathcal{J}^{\infty,p}_\gamma(\Omega)$ defined similarly. To simplify the notation, we will suppose that the there is only one singular point, i.e., $\mathfrak{C} = \{\mathfrak{c}\}$ and omit $\mathfrak{C}$ from the notation of the spaces. Furthermore, we write $\mathcal{K}^k_\gamma(\Omega) = \mathcal{K}^{k,2}_\gamma(\Omega)$, $\mathcal{J}^k_\gamma(\Omega) = \mathcal{J}^{k,2}_\gamma(\Omega)$, $\mathcal{K}^\varpi_\gamma(\Omega) = \mathcal{K}^{\varpi,2}_\gamma(\Omega)$, and $\mathcal{J}^\varpi_\gamma(\Omega) = \mathcal{J}^{\varpi,2}_\gamma(\Omega)$. For a thorough treatment of Kondrat'ev-Babuška spaces, see [KMR97, CDN10a, CDN10b, CDN12]. Note that the results obtained in the sequel can be trivially extended to the case where $\mathfrak{C}$ contains more than one point, as long as $\mathfrak{C}$ is a finite set of isolated points.

Finally, let $X = \mathcal{J}^2_\gamma(\Omega)$, for $\gamma \in (d/2, d/2+\varepsilon)$, where $0 < \varepsilon < 1$ will be specified later, namely in hypothesis (7b).

2.2. **Statement of the problem.** We introduce the problem under consideration. From the "physical" point of view, it consists in the minimization of an energy composed by a kinetic term, an interaction with a singular potential V and a nonlinear self-interaction term. Under the unitary norm constraint, using Euler's equation, the energy minimization problem translates into a nonlinear elliptic eigenvalue problem. This is the form under which most of the analysis will be carried out.

We start therefore by introducing the bilinear form over $H^1(\Omega) \times H^1(\Omega)$

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v + \int_\Omega V uv$$

and a function $F : \mathbb{R}^+ \to \mathbb{R}$, whose properties will be listed later in this section for the sake of clarity. Let

$$E(v) = \frac{1}{2} a(v, v) + \frac{1}{2} \int_\Omega F(v^2) \tag{3}$$

Let us denote by u the minimizer of (3) (unique up to a sign change under the hypotheses that follow) over the space $\{v \in H^1(\Omega) : \|v\| = 1\}$: then, there exists $\lambda \in \mathbb{R}$ such that u is the solution of

$${}_{X'}\langle A^u u - \lambda u, v\rangle_X = 0 \quad \forall v \in H^1(\Omega) \tag{4}$$

where

$${}_{X'}\langle A^u v, w\rangle_X = a(v, w) + \int_\Omega f(u^2) vw,$$

with $f = F'$. We introduce also

$$\langle E''(u)v, w\rangle = \langle A^u v, w\rangle + 2 \int_\Omega f'(u^2) u^2 vw. \tag{5}$$

The properties of the function F will be similar to those in [CCM10], namely we suppose that

$$F \in C^1([0, +\infty), \mathbb{R}) \cap C^\infty((0, +\infty), \mathbb{R}) \text{ and } F'' > 0 \text{ in } (0, +\infty), \tag{6a}$$

$$\exists q \in [0, 2), \exists C \in \mathbb{R} : \forall t \geq 0, |F'(t)| \leq C(1 + t^q), \tag{6b}$$

$$F''(t)t \text{ locally bounded in } [0, +\infty), \tag{6c}$$

and we suppose that $\forall R > 0, \exists C_R \in \mathbb{R}_+ : \forall t_1 \in (0, R], \forall t_2 \in \mathbb{R}$,

$$|F'(t_2^2)t_2 - F'(t_1^2)t_2 - 2F''(t_1^2)(t_1^2)(t_2 - t_1)| \leq C_R(1 + |t_2|^s)|t_2 - t_1|^r \tag{6d}$$

for $r \in (1, 2]$ and $s \in [0, 5 - r)$. We will impose additional conditions on F in order to obtain some improved convergence estimates: those conditions will be specified when necessary. Finally, we suppose that the potential V is such that

$$V \in L^{p_V}(\Omega) \tag{7a}$$

with $p_V > \max(1, d/2)$ and that there exists $0 < \varepsilon < 1$ such that

$$V \in \mathcal{K}^{\varpi,\infty}_{-2+\varepsilon}(\Omega, \mathfrak{C}). \tag{7b}$$

For $d = 2, 3$, (7b) implies (7a) as long as $p_V < d/(2 - \varepsilon)$. A consequence of (7a) is, in particular, that for $u, v \in H^1(\Omega)$,

$$(Vu, v) \leq C\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}.$$

where the constant C depends on V and on the domain. We have also the following regularity result, which follows from (6b) and (7b) and the regularity result obtained in [MM19].

Lemma 1. *The solution u to* (4)*, under hypotheses* (6a) *to* (7b)*, belongs to the space*

$$u \in \mathcal{J}^2_{d/2+\alpha}(\Omega)$$

for any $\alpha < \varepsilon$.

Proof. We have $u \in L^\infty(\Omega)$, see [CCM10]. Hence, u is the solution of $(-\Delta + V)u = \lambda u + f(u^2)u$, with right hand side belonging to the space $\mathcal{J}^0_{d/2+\varepsilon-2}(\Omega)$. Since the operator $-\Delta + V$ is an isomorphism from $\mathcal{J}^2_{d/2+\alpha}(\Omega)$ to $\mathcal{J}^0_{d/2+\alpha-2}(\Omega)$ for $0 < \alpha < \varepsilon$, see [MM19], we obtain the thesis. □

2.3. **Numerical method.** In this section we introduce the hp discontinuous Galerkin method. Concerning the design of the hp space, the setting is the one from [GB86a, GB86b]. Let $\mathcal{T}$ be a triangulation of axiparallel quadrilateral ($d = 2$) or hexahedral ($d = 3$) elements of Ω, such that $\bigcup_{K\in\mathcal{T}} \overline{K} = \overline{\Omega}$, whose properties will be specified later. A $d-1$ dimensional face (edge, when $d = 2$) is defined as the nonempty interior of $\partial K_\sharp \cap \partial K_\flat$ for two adjacent elements $K_\flat$ and $K_\sharp$. Let $\mathcal{E}$ be the set of all faces/edges. We denote

$$(u, v)_{\mathcal{T}} = \sum_{K\in\mathcal{T}} (u, v)_K$$

and, similarly,

$$(u, v)_{\mathcal{E}} = \sum_{e\in\mathcal{E}} (u, v)_e.$$

We suppose that for any $K \in \mathcal{T}$ there exists an affine transformation $\Phi : K \to \hat{K}$ to the d-dimensional cube $\hat{K}$ such that $\Phi(K) = \hat{K}$, that the mesh is shape and contact regular[1] Let $\mathcal{T}$ be isotropically and geometrically graded around the points in $\mathfrak{C}$: consider the case where $\mathfrak{C} = \{\mathfrak{c}\}$. Then, we fix a refinement ratio $\sigma \in (0, 1/2)$ and partition the mesh $\mathcal{T}$ into disjoint mesh layers Ω_j, $j = 1, \dots, \ell$ such that $\mathcal{T} = \bigcup_j \Omega_j$ and

$$h_K \simeq h_j = \sigma^j \qquad d(\mathfrak{c}, K) \simeq h_K,$$

for all $K \in \mathcal{T}$ and with constants uniform in $\mathcal{T}$ and ℓ. The generalization to the case of $\mathfrak{C}$ containing multiple points follows from the construction of a graded mesh around each point.

We will allow for 1-irregular edges/faces, i.e., given two neighboring elements $K_\flat$ and $K_\sharp$, that share an edge/face $e = \partial K_\sharp \cap \partial K_\flat$, we require that e is an entire edge/face of at least one between $K_\sharp$ and $K_\flat$. We refer to Section 6 (specifically, to Figure 1a) for a visualization of such a mesh. We introduce on this mesh the hp space with linear polynomial slope s, i.e., for an element $K \in \mathcal{T}$ such that $K \in \Omega_j$,

$$p_K \simeq p_j = p_0 + \mathfrak{s}(\ell - j),$$

where h_K is the diameter of the element K and p_K is the polynomial order whose role will be specified in (8). We introduce the discrete space

$$X_\delta = \left\{ v_\delta \in L^2(\Omega) : (v_{|K} \circ \Phi^{-1}) \in \mathbb{Q}_{p_K}(\hat{K}),\ \forall K \in \mathcal{T} \right\}, \tag{8}$$

where $\mathbb{Q}_p$ is the space of polynomials of maximal degree p in any variable and denote

$$X(\delta) = X + X_\delta.$$

[1] If h_K is the diameter of an element $K \in \mathcal{T}$ and ρ_K is the radius of the largest ball inscribed in K, a mesh sequence is *shape regular* if there exists C independent of the refinement level such that $h_K \leq C r_K$ for all $K \in \mathcal{T}$. The mesh sequence is *contact regular* if for all $K \in \mathcal{T}$, the number of elements adjacent to K is uniformly bounded and there exists a constant C independent from the refinement level such that for every face/edge e of K, $h_K \leq C h_e$, where h_e is the diameter of e.

Then $\mathcal{E}$ is the set of the edges (for $d = 2$) or faces ($d = 3$) of the elements in $\mathcal{T}$ and

$$\mathtt{h_e} = \min_{K\in\mathcal{T}: e\cap\partial K\neq\varnothing} h_K$$
$$\mathtt{p_e} = \max_{K\in\mathcal{T}: e\cap\partial K\neq\varnothing} p_K.$$

On an edge/face between two elements $K_\sharp$ and $K_\flat$, i.e., on $e \subset \partial K_\sharp \cap \partial K_\flat$, the average $\{\!\{\cdot\}\!\}$ and jump $[\![\cdot]\!]$ operators for a function $w \in X(\delta)$ are defined by

$$\{\!\{w\}\!\} = \frac{1}{2}\left(w_{|K_\sharp} + w_{|K_\flat}\right), \qquad [\![w]\!] = w_{|K_\sharp}\mathbf{n}_\sharp + w_{|K_\flat}\mathbf{n}_\flat,$$

where $\mathbf{n}_\sharp$ (resp. $\mathbf{n}_\flat$) is the outward normal to the element $K_\sharp$ (resp. $K_\flat$). 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 now introduce the discrete versions of the operators defined in Section 2.2. First, the bilinear form a_δ over $X_\delta \times X_\delta$ is given by

$$\begin{aligned} a_\delta(u_\delta, v_\delta) = (\nabla u_\delta, \nabla v_\delta)_\mathcal{T} - (\{\!\{\nabla u_\delta\}\!\}, [\![v_\delta]\!])_\mathcal{E} - (\{\!\{\nabla v_\delta\}\!\}, [\![u_\delta]\!])_\mathcal{E} \\ + \sum_{e\in\mathcal{E}} \alpha_e \frac{\mathtt{p_e}^2}{\mathtt{h_e}}([\![u_\delta]\!], [\![v_\delta]\!])_e + \int_\Omega V u_\delta v_\delta. \end{aligned} \tag{9}$$

Furthermore,

$$E_\delta(v_\delta) = \frac{1}{2} a_\delta(v_\delta, v_\delta) + \frac{1}{2}\int_\Omega F(v_\delta^2). \tag{10}$$

Let u_δ be a minimizer of (10) over X_δ, with unitary norm constraint. Then, there exists an eigenvalue $\lambda_\delta \in \mathbb{R}$ such that

$$\langle A_\delta^{u_\delta} u_\delta - \lambda_\delta u_\delta, v_\delta\rangle = 0 \quad \forall v_\delta \in X_\delta \tag{11}$$

where

$$\langle A_\delta^{u_\delta} v_\delta, w_\delta\rangle = a_\delta(v_\delta, w_\delta) + \int_\Omega f(u_\delta^2) v_\delta w_\delta.$$

Finally, E''_δ, defined on X_δ, is obtained by replacing A^u with A^u_δ in (5).

Remark 1 (Simmetry of the numerical method)**.** *The dG method with bilinear form* (9) *is the symmetric interior penalty* (*SIP*) *method. The requirement of symmetry in the bilinear form of the numerical method is a strong one, and will be used without explicit mention throughout the proofs.*

This could be seen as a limitation; nonetheless, from a practical point of view, there is little interest in approximating a symmetric eigenvalue problem with a non symmetric numerical method. Non symmetric methods tend to exhibit, in the linear case, lower rates of convergence than symmetric ones [ABP06]*. Furthermore, the solution of the finite dimensional problem would be more problematic, since algebraic eigenvalue problems are more easily treated for symmetric matrices* [Saa11].

We introduce the mesh dependent norms that will be used in this section. First, for a $v \in X(\delta)$,

$$\|v\|_{\mathrm{DG}}^2 = \sum_{K\in\mathcal{T}} \|v\|^2_{\mathcal{J}^1_1(K)} + \sum_{e\in\mathcal{E}} \mathtt{h_e}^{-1}\mathtt{p_e}^2 \|[\![v]\!]\|^2_{L^2(e)}. \tag{12}$$

Remark that on X, this norm is equivalent to the $\mathcal{J}^1_1(\Omega) = H^1(\Omega)$ norm, since functions in X have no face discontinuity, implying $[\![v]\!] = 0$. Then, on $X(\delta)$ we introduce, when $d = 3$,

$$|\!|\!|u|\!|\!|_{\mathrm{DG}}^2 = \sum_{K\in\mathcal{T}} \|u\|^2_{H^1(K)} + \sum_{e\in\mathcal{E}} \mathtt{p_e}^2\mathtt{h_e}^{-1}\|[\![u]\!]\|^2_{L^2(e)} + \sum_{e\in\mathcal{E}} \mathtt{p_e}^{-2}\|r^{1/2}\nabla u\cdot n_e\|^2_{L^2(e)}, \tag{13}$$

where n_e denotes the normal to face e. If $d = 2$, we denote by $\mathcal{E}_c$ the set of edges abutting at the singularity, and write (note that on $\mathcal{E}_c$, $\mathtt{p_e} = p_0$)

$$\|\!|u|\!\|_{\mathrm{DG}}^2 = \sum_{K\in\mathcal{T}} \|u\|^2_{H^1(K)} + \sum_{e\in\mathcal{E}} \mathtt{p_e}^2 \mathtt{h_e}^{-1} \|[\![u]\!]\|^2_{L^2(e)} + \sum_{e\in\mathcal{E}\setminus\mathcal{E}_c} \mathtt{p_e}^{-2} \|r^{1/2}\nabla u\cdot n_e\|^2_{L^2(e)} + \sum_{e\in\mathcal{E}_c} \mathtt{h_e}^{q-1} \|\nabla u \cdot n_e\|^q_{L^q(e)}, \tag{14}$$

where q is fixed and such that $1 < q < 2/(3-\gamma)$, see Remark 2.

Let us also introduce the broken space

$$\mathcal{J}^{s,p}_\gamma(\Omega,\mathcal{T}) = \left\{ v : v \in \mathcal{J}^{s,p}_\gamma(K), \forall K \in \mathcal{T} \right\}.$$

Remark 2. *When $d = 3$, by the definition of the weighted spaces, see* [MR10], *for $e \subset \partial K$, and since $\gamma > 3/2$,*

$$\|r^{1/2}\nabla v\|_{L^2(e)} \le C\|v\|_{\mathcal{J}^2_\gamma(K)},$$

then $\|\!|v|\!\|_{\mathrm{DG}}$ (13) is bounded on $\mathcal{J}^2_\gamma(\Omega,\mathcal{T})$. Since furthermore $X(\delta) \subset \mathcal{J}^2_\gamma(\Omega,\mathcal{T})$, $\|\!|v|\!\|_{\mathrm{DG}}$ as defined in (13) is bounded when $d = 3$ for all $v \in X(\delta)$.

When $d = 2$, we consider the definition of the norm (14). Let $\gamma > 1$: then, for any $q < 2/(3-\gamma)$ and writing $t = (1/q - 1/2)^{-1}$, there exists C_1, C_2 such that

$$\sum_{|\alpha|=2} \|\partial^\alpha v\|_{L^q(\Omega)} \le C_1 \|r^{\gamma-2}\|_{L^t(\Omega)} \sum_{|\alpha|=2} \|r^{2-\gamma}\partial^\alpha v\|_{L^2(\Omega)} \le C_2 \|v\|_{\mathcal{J}^2_\gamma(\Omega)}.$$

Hence, $u \in W^{2,q}(\Omega)$, and $\|\nabla u \cdot n\|_{L^q(e)}$ is well defined. Therefore, $\|\!|v|\!\|_{\mathrm{DG}}$ as defined in (14) is bounded when $d = 2$ for all $v \in X(\delta)$.

Remark 3. *By proving continuity as in Lemma 3 (see the part of the proof referring to inequality (17a)) and thanks to Remark 2, it can be shown that the bilinear form a_δ defined in (9) over $X_\delta \times X_\delta$ can be extended over $X(\delta) \times X_\delta$.*

Remark 4. *Note that on X_δ and for $d \le 3$, the two norms (12) and (13) are equivalent, since for any $K \in \mathcal{T}$, $r_{|K} \lesssim h_K$ and thanks to the discrete trace inequality* [DPE12]

$$h_e^{(1-d)/p+d/2} \|w_\delta\|_{L^p(e)} \le C_{d,p} \|w_\delta\|_{L^2(K)},$$

valid for $e \in \partial K$ and for all $w_\delta \in X_\delta$. The constant $C_{d,p}$ depends on the dimension d, on p, and on the polynomial order $\mathtt{p_e}$, but is independent of $\mathtt{h_e}$. Furthermore, $C_{d,p}$ is bounded by $\mathtt{p_e}$ if $p = 2$.

We conclude this section by introducing the discrete approximation to the solution of the linear problem, i.e. the function $u^*_\delta \in X_\delta$ such that

$$\langle A^u_\delta u^*_\delta - \lambda^*_\delta u^*_\delta, v_\delta\rangle = 0 \quad \forall v_\delta \in X_\delta \tag{15}$$

for an eigenvalue λ^*_δ. Note that, since u is an eigenfunction of A^u and the associated eigenspace is of dimension 1 [CCM10], we have that

$$\|u^*_\delta - u\|_{\mathrm{DG}} \lesssim \inf_{v_\delta\in X_\delta} \|\!|u - v_\delta|\!\|_{\mathrm{DG}},$$

$$|\lambda^*_\delta - \lambda| \lesssim \inf_{v_\delta\in X_\delta} \|\!|u - v_\delta|\!\|^2_{\mathrm{DG}},$$

see [MM19], and the eigenspace associated with u^*_δ is of dimension one, for a sufficiently large number of degrees of freedom [ABP06].

The isotropically refined *hp* finite element space X_δ defined here provides approximations that converge with exponential rate to the function in the weighted analytic class, as stipulated in the following statement, see [SSW13].

Proposition 2. *Given a function $v \in \mathcal{J}^{\varpi}_{\gamma}(\Omega)$, for a $\gamma > d/2$, there exists two constants $C, b > 0$ independent of ℓ such that*

$$\inf_{v_\delta \in X_\delta} |||v - v_\delta|||_{\mathrm{DG}} \leq C e^{-b\ell}. \tag{16}$$

Here, ℓ is the number of refinement steps, and $\ell = N^{1/(d+1)}$, with N denoting the number of degrees of freedom of X_δ.

3. A priori estimates

In this section we prove some a priori estimates on the convergence of the numerical eigenfunction and eigenvalue. We start by giving some continuity and coercivity estimates, then we provide an auxiliary estimate on a scalar product where we construct an adjoint problem, and we conclude by proving convergence and quasi optimality for the eigenfunctions. The rate of convergence proven for the eigenvalues is smaller than what is obtained in the linear case: in the following it will be shown that under some additional hypothesis we can recover the rate typically obtained in the approximation of solutions to linear elliptic operators with singular potentials, see [MM19].

Since our main focus here is on isotropically refined hp methods, the approach we take uses the assumption that finite element space and the underlying mesh are those of an hp discontinuous Galerkin method, as described in the previous sections. It is important to remark, nonetheless, that the results of this section can be extended, with minimal effort, to the analysis of a general discontinuous Galerkin approximation. The novelty of the approach we use in this section lies, indeed, more into the treatment of the nonconformity of the method than in the aspects related to the hp space. The modification necessary to get a proof that applies to a classical h-type discontinuous Galerkin finite element method, for example, would be related to the continuity and coercivity estimates, since those would need not to use the hypothesis that $r \simeq h$.

For the aforementioned reason, and for the sake of generality, we prove our results for an F as general as possible, even though the hp method shows its full power (i.e., exponential rate of convergence) only in a less general setting.

To conclude, we mention the fact that we will mainly write our proofs so that they work for $d = 3$, even though this sometimes means using a suboptimal strategy for the case $d \leq 2$. Consider for example the bound

$$\|v\|_{L^p(\Omega)} \leq C\|v\|_{H^1(\Omega)},$$

for a $v \in H^1(\Omega)$: we will always use it for p such that $1 \leq p \leq 6$, even if for $d = 2$ any $1 \leq p < \infty$ would be acceptable.

3.1. Continuity and coercivity.

We start with an auxiliary lemma, where we prove the continuity, positivity and coercivity of some operators. As mentioned before, we use the numerical eigenvalue λ^*_δ obtained from the numerical approximation of the linear problem as a lower bound of the operators over the discrete space X_δ.

Lemma 3. *Given the definition of the operators A^u_δ and $E''_\delta(u)$, of the spaces X_δ and $X(\delta)$, and of λ^*_δ provided in Section 2, the following results hold*

$$|\langle (A^u_\delta - \lambda^*_\delta)\, v, v_\delta \rangle| \lesssim |||v|||_{\mathrm{DG}}\, \|v_\delta\|_{\mathrm{DG}} \qquad \forall v \in X(\delta), v_\delta \in X_\delta \tag{17a}$$

$$\langle (A^u_\delta - \lambda^*_\delta)\, v_\delta, v_\delta \rangle \geq 0 \qquad \forall v_\delta \in X_\delta. \tag{17b}$$

Furthermore,

$$\langle (A^u_\delta - \lambda^*_\delta)\,(u_\delta - u^*_\delta), (u_\delta - u^*_\delta) \rangle \gtrsim \|u_\delta - u^*_\delta\|^2_{\mathrm{DG}} \tag{18}$$

and

$$\langle (E''_\delta(u) - \lambda^*_\delta)\, v_\delta, v_\delta\rangle \gtrsim \|v_\delta\|^2_{\mathrm{DG}} \qquad \forall v_\delta \in X_\delta \tag{19a}$$

$$|\langle (E''_\delta(u) - \lambda^*_\delta)\, v, v_\delta\rangle| \lesssim |||v|||_{\mathrm{DG}}\, \|v_\delta\|_{\mathrm{DG}} \qquad \forall v \in X(\delta), v_\delta \in X_\delta. \tag{19b}$$

Proof. Let us first consider the continuity inequality (17a). The proof when $d = 2$ is classical, see in particular [DPE12, Lemma 4.30] for the bound on the edges in $\mathcal{E}_c$, and the same arguments that we use here for the bounds on the rest of the elements and edges. We restrict then ourselves here to the case $d = 3$, where we use a slightly different norm than usual. Consider a function $v \in X(\delta)$. We can decompose $v = \tilde{v} + \tilde{v}_\delta$, where $v \in X$ and $\tilde{v}_\delta \in X_\delta$. Consider an edge/face $e \in \mathcal{E}$. Then, $[\![v]\!]_{|e} = [\![\tilde{v}_\delta]\!]_{|e}$. If $\mathfrak{C} \cap \bar{e} = \varnothing$, then $\mathtt{h_e} \simeq r$; if instead there exists a $\mathfrak{c} \in \mathfrak{C}$ such that $\mathfrak{c}$ is one of the vertices of e, then $[\![\tilde{v}_\delta]\!]_{|e} \in \mathbb{Q}_{p_0}(e)$, which is a finite dimensional space of fixed size. Therefore on $X(\delta)$ we have the equivalency

$$\mathtt{h_e}^{-1} \|[\![\cdot]\!]\|^2_{L^2(e)} \simeq \|r^{-1/2}[\![\cdot]\!]\|^2_{L^2(e)} \tag{20}$$

if $d = 3$. The continuity estimate (17a) can be obtained through multiple applications of Hölder's inequality: we consider the terms in the bilinear form separately. First, on the broken space $H^1(\mathcal{T}) := \{v : v_{|K} \in H^1(K), \forall K \in \mathcal{T}\}$ we exploit the fact that, as shown in [LS03],

$$\|v\|_{L^q(\Omega)} \lesssim \|v\|_{\mathrm{DG}} \quad \forall v \in H^1(\mathcal{T}) \tag{21}$$

with $q \le 2d/(d-2)$ if $d \ge 3$ and $q \in [1, \infty)$ if $d = 2$. Note that $X(\delta) \subset H^1(\mathcal{T})$, thus

$$\left|\sum_{K\in\mathcal{T}} (\nabla v, \nabla v_\delta)_K + (Vv, v_\delta)_K\right| \lesssim \|v\|_{\mathrm{DG}}\|v_\delta\|_{\mathrm{DG}}$$

Secondly,

$$\begin{aligned}\left|\sum_e (\{\!\{\nabla v\}\!\}, [\![v_\delta]\!])_e\right| &\lesssim \sum_e \mathtt{p_e}^{-1}\|r^{1/2}\{\!\{\nabla v\}\!\}\|_{L^2(e)}\mathtt{p_e}\|r^{-1/2}[\![v_\delta]\!]\|_{L^2(e)}\\ &\lesssim \sum_e \mathtt{p_e}^{-1}\|r^{1/2}\{\!\{\nabla v\}\!\}\|_{L^2(e)}\mathtt{p_e}\mathtt{h_e}^{-1/2}\|[\![v_\delta]\!]\|_{L^2(e)}\\ &\lesssim \left(\sum_e \mathtt{p_e}^{-2}\|r^{1/2}\{\!\{\nabla v\}\!\}\|^2_{L^2(e)}\right)^{1/2}\left(\sum_e \mathtt{p_e}^2\mathtt{h_e}^{-1}\|[\![v_\delta]\!]\|^2_{L^2(e)}\right)^{1/2}\end{aligned}$$

where the second inequality follows from (20). Similarly,

$$\begin{aligned}\left|\sum_e (\{\!\{\nabla v_\delta\}\!\}, [\![v]\!])_e\right| &\lesssim \left(\sum_e \mathtt{p_e}^{-2}\mathtt{h_e}\|\{\!\{\nabla v_\delta\}\!\}\|^2_{L^2(e)}\right)^{1/2}\left(\sum_e \mathtt{p_e}^2\mathtt{h_e}^{-1}\|[\![v]\!]\|^2_{L^2(e)}\right)^{1/2}\\ &\lesssim \left(\sum_K \|\nabla v_\delta\|^2_{L^2(K)}\right)^{1/2}\left(\sum_e \mathtt{p_e}^2\mathtt{h_e}^{-1}\|[\![v]\!]\|^2_{L^2(e)}\right)^{1/2},\end{aligned}$$

using (4) in the second line. Then,

$$\begin{aligned}\left|\sum_{e\in\mathcal{E}} \alpha_e \frac{\mathtt{p_e}^2}{\mathtt{h_e}}([\![v]\!], [\![v_\delta]\!])_e\right| &\lesssim C\sum_e \left(\mathtt{p_e}\mathtt{h_e}^{-1/2}\|[\![u]\!]\|_{L^2(e)}\right)\left(\mathtt{p_e}\mathtt{h_e}^{-1/2}\|[\![v_\delta]\!]\|_{L^2(e)}\right)\\ &\lesssim C\left(\sum_e \mathtt{p_e}^2\mathtt{h_e}^{-1}\|[\![v]\!]\|^2_{L^2(e)}\right)^{1/2}\left(\sum_e \mathtt{p_e}^2\mathtt{h_e}^{-1}\|[\![v_\delta]\!]\|^2_{L^2(e)}\right)^{1/2}.\end{aligned}$$

Thanks to the Hölder inequality, Sobolev imbeddings, hypothesis (6b), and (21),

$$\begin{aligned}\left|\int_\Omega f(u^2)vv_\delta\right| &\lesssim \|1+u^{2q}\|_{L^{3/2}(\Omega)}\|v\|_{L^6(\Omega)}\|v_\delta\|_{L^6(\Omega)}\\ &\lesssim \|u\|^{2q}_{H^1(\Omega)}\|v\|_{H^1(\Omega)}\|v_\delta\|_{\mathrm{DG}}\\ &\lesssim \|v\|_{H^1(\Omega)}\|v_\delta\|_{\mathrm{DG}}.\end{aligned}$$

Since then $\lambda^*_\delta \to \lambda$, we have that $|\lambda^*_\delta(v, v_\delta)| \leq C\|v\|\|v_\delta\|$ and this, combined with the above inequalities, proves (17a).

We now consider (17b). As already stated, λ^*_δ is a simple eigenvalue for a sufficient number of degrees of freedom and therefore $A^u_\delta - \lambda^*_\delta$ is coercive on the subspace of X_δ L^2-orthogonal to u^*_δ. Hence, since $\|u^*_\delta\| = 1$ and A^u_δ is symmetric,

$$\begin{aligned}\langle(A^u_\delta - \lambda^*_\delta)\, v_\delta, v_\delta\rangle &= \langle(A^u_\delta - \lambda^*_\delta)\,(v_\delta - (v_\delta, u^*_\delta)_\Omega u^*_\delta), v_\delta - (v_\delta, u^*_\delta)_\Omega u^*_\delta\rangle\\ &\gtrsim \|v_\delta\|^2 - (u^*_\delta, v_\delta)^2 \geq 0,\end{aligned} \tag{22}$$

for all $v_\delta \in X_\delta$. We may then prove (18) following the same reasoning as in [CCM10]. We recall it here for ease of reading. We choose, without loss of generality, u^*_δ such that $(u^*_\delta, u_\delta) \geq 0$. From the above inequality we have (recall that $\|u^*_\delta\| = \|u_\delta\| = 1$)

$$\begin{aligned}\langle(A^u_\delta - \lambda^*_\delta)\,(u_\delta - u^*_\delta), (u_\delta - u^*_\delta)\rangle &\gtrsim \|u_\delta - u^*_\delta\|^2 - (u^*_\delta, u_\delta - u^*_\delta)^2\\ &= \|u_\delta - u^*_\delta\|^2 - (1 + (u^*_\delta, u_\delta)^2 - 2(u^*_\delta, u_\delta))\\ &= 1 - (u^*_\delta, u_\delta)^2\\ &\geq \frac{1}{2}\|u_\delta - u^*_\delta\|^2,\end{aligned}$$

and this proves (18). To prove (19a), we note that

$$\langle(E''_\delta(u) - \lambda^*_\delta)\, v_\delta, v_\delta\rangle \geq \langle(A^u_\delta - \lambda^*_\delta)\, v_\delta, v_\delta\rangle + \int_\Omega f'(u^2)u^2v_\delta^2.$$

Suppose we negate (19a): then, there has to be a sequence $\{v^j_\delta\}_j \subset X_\delta$ such that $\|v^j_\delta\| = 1$ and $\langle(E''_\delta(u) - \lambda^*_\delta)\, v^j_\delta, v^j_\delta\rangle \to 0$. Since $\int_\Omega f'(u^2)u^2(v^j_\delta)^2 > 0$, from (22) we have that

$$\begin{aligned}\frac{1}{2}\|v^j_\delta - u^*_\delta\|^2 &= \|v^j_\delta\|^2 - (v^j_\delta, u^*_\delta)^2\\ &\lesssim \langle(E''(u) - \lambda^*_\delta)v^j_\delta, v^j_\delta\rangle,\end{aligned}$$

thus, $v^j_\delta \to u^*_\delta$ in $L^2(\Omega)$. Now, since u^*_δ converges towards u in the DG norm, and using (6c) and the positivity of f', we can show that there exists an $\alpha > 0$ such that, for a sufficient number of degrees of freedom,

$$\int_\Omega f'(u^2)u^2(u^*_\delta)^2 > \alpha.$$

This negates the contradiction hypothesis that $\langle(E''_\delta(u) - \lambda^*_\delta)\, v^j_\delta, v^j_\delta\rangle \to 0$, hence

$$\langle(E''_\delta(u) - \lambda^*_\delta)\, v_\delta, v_\delta\rangle \geq C\|v_\delta\|^2 \tag{23}$$

for all $v_\delta \in X_\delta$. Then, using the classical result that

$$(\nabla v_\delta, \nabla v_\delta)_{\mathcal{T}} - (\{\!\{\nabla v_\delta\}\!\}, [\![v_\delta]\!])_{\mathcal{E}} - (\{\!\{\nabla v_\delta\}\!\}, [\![v_\delta]\!])_{\mathcal{E}} + \sum_{e\in\mathcal{E}} \alpha_e \frac{\mathtt{p_e}^2}{\mathtt{h_e}}([\![v_\delta]\!], [\![v_\delta]\!])_e \geq \|v_\delta\|^2_{\mathrm{DG}},$$

combined with the estimate from the proof of [CCM10, Lemma 1], we can show that

$$\langle (A_\delta^u - \lambda_\delta^*)\, v_\delta, v_\delta\rangle \geq \alpha \|v_\delta\|_{\mathrm{DG}}^2 - C\|v_\delta\|^2. \tag{24}$$

The coercivity estimate (19a) then follows from (23) and (24).

Finally, (19b) follows directly from the definition of $E''_\delta(u)$, the continuity estimate (17b) and the fact that $|f'(u^2)u^2| \leq C$. □

3.2. Estimates on the adjoint problem. In this section we develop an estimate on the scalar product between a function and the error $u - u_\delta$, whose interest lies mainly in the $L^2(\Omega)$ convergence estimate given in Theorem 1. The estimate is based on the introduction of the adjoint problem (25).

Lemma 4. *Let $u_\delta^{\star\perp} = \{v_\delta \in X_\delta : (v_\delta, u_\delta^*) = 0\}$ be the space of functions $L^2(\Omega)$-orthogonal to u_δ^* and let ψ_{w_δ} be the solution to the problem*

$$\begin{aligned} &\text{find } \psi_{w_\delta} \in u_\delta^{\star\perp} \text{ such that} \\ &\langle (E''_\delta(u) - \lambda_\delta^*)\, \psi_{w_\delta}, v_\delta\rangle = \langle w_\delta, v_\delta\rangle, \forall v_\delta \in u_\delta^{\star\perp} \end{aligned} \tag{25}$$

for w_δ in $L^2(\Omega)$. Then, if hypotheses (6a) *to* (6d) *hold,*

$$\begin{aligned} |\langle w_\delta, u_\delta - u_\delta^*\rangle| \lesssim{}& \|u_\delta - u_\delta^*\|^r_{L^{6r/(5-s)}} \|\psi_{w_\delta}\|_{\mathrm{DG}} + |\lambda_\delta - \lambda_\delta^*| \|u_\delta - u_\delta^*\| \|\psi_{w_\delta}\| + \|u - u_\delta^*\| \|\psi_{w_\delta}\| \\ &+ \|u_\delta - u_\delta^*\|^2 \|\psi_{w_\delta}\| + \|u_\delta - u_\delta^*\|^2 \|w_\delta\|, \end{aligned} \tag{26}$$

Proof. We break $u_\delta - u_\delta^*$ into two parts, one parallel to u_δ^* and one perpendicular to it. Those are given respectively by

$$(u_\delta - u_\delta^*, u_\delta^*)u_\delta^* = -\frac{1}{2}\|u_\delta - u_\delta^*\|^2 u_\delta^* \quad \text{and} \quad u_\delta - (u_\delta, u_\delta^*)u_\delta^* \in u_\delta^{\star\perp}.$$

Then,

$$\begin{aligned} \langle w_\delta, u_\delta - u_\delta^*\rangle &= (w_\delta, u_\delta - (u_\delta, u_\delta^*)u_\delta^*) - \frac{1}{2}\|u_\delta - u_\delta^*\|^2 (w_\delta, u_\delta^*) \\ &= \langle (E''_\delta(u) - \lambda_\delta^*)\, \psi_{w_\delta}, u_\delta - (u_\delta, u_\delta^*)u_\delta^*\rangle - \frac{1}{2}\|u_\delta - u_\delta^*\|^2 (w_\delta, u_\delta^*) \\ &= \langle (E''_\delta(u) - \lambda_\delta^*)\, (u_\delta - u_\delta^*), \psi_{w_\delta}\rangle - \frac{1}{2}\|u_\delta - u_\delta^*\|^2 \langle (E''_\delta(u) - \lambda)\, u_\delta^*, \psi_{w_\delta}\rangle \\ &\quad - \frac{1}{2}\|u_\delta - u_\delta^*\|^2 (w_\delta, u_\delta^*) \\ &= \langle (E''_\delta(u) - \lambda_\delta^*)\, (u_\delta - u_\delta^*), \psi_{w_\delta}\rangle - \|u_\delta - u_\delta^*\|^2 \int_\Omega f'(u^2)u^2 u_\delta^* \psi_{w_\delta} \\ &\quad - \frac{1}{2}\|u_\delta - u\|^2 (w_\delta, u_\delta^*). \end{aligned} \tag{27}$$

We consider the first term:

$$\begin{aligned} \langle (E''_\delta(u) - \lambda_\delta^*)\, (u_\delta - u_\delta^*), \psi_{w_\delta}\rangle &= \langle (A_\delta^u - \lambda_\delta^*) u_\delta, \psi_{w_\delta}\rangle + 2\int_\Omega f'(u^2)u^2 \psi_{w_\delta}(u_\delta - u_\delta^*) \\ &= -\int_\Omega \left[f({u_\delta}^2)u_\delta - f(u^2)u_\delta - 2f'(u^2)u^2(u_\delta - u)\right] \psi_{w_\delta} \\ &\quad + (\lambda_\delta - \lambda_\delta^*)(u_\delta - u_\delta^*, \psi_{w_\delta}) \\ &\quad + 2\int_\Omega f'(u^2)u^2 \psi_{w_\delta}(u - u_\delta^*). \end{aligned} \tag{28}$$

Thanks to (6d), combining (21), (27), and (28) we can infer that

$$|\langle w_\delta, u_\delta - u_\delta^* \rangle| \lesssim \|u_\delta - u_\delta^*\|^r_{L^{6r/(5-s)}} \|\psi_{w_\delta}\|_{\mathrm{DG}} + |\lambda_\delta - \lambda_\delta^*| \|u_\delta - u_\delta^*\| \|\psi_{w_\delta}\| + \|u - u_\delta^*\| \|\psi_{w_\delta}\| \\ + \|u_\delta - u_\delta^*\|^2 \int_\Omega \left| f'(u^2) u^2 u_\delta^* \psi_{w_\delta} \right| + \|u_\delta - u_\delta^*\|^2 \left| (w_\delta, u_\delta^*) \right|,$$

which gives the thesis. □

3.3. **Basic convergence.** At this stage, we are able to prove the first convergence result for the numerical eigenfunction and eigenvalue. We work mainly in the discrete setting, in order to avoid the issues due to the nonconformity of the method. The analysis is carried out for the symmetric interior penalty discontinuous Galerkin method, but it holds for any nonconforming symmetric method, as long as the results of Lemma 3 hold for such a method. Furthermore, the remark made at the beginning of Section 3 still holds, in that the result can be adapted with few modifications to a classical h-type discontinuous Galerkin finite element method.

In general, the goal is to prove that the numerical eigenvalue-eigenfunction couple obtained as solution to the nonlinear problem converges as fast as for linear elliptic operators. In this section, we obtain this result for the eigenfunction, which is shown to converge quasi optimally. The hypotheses on the function F are instead not strong enough to prove that the eigenvalue converges twice as fast as the eigenfunction in the $\|\ \|_{\mathrm{DG}}$ norm. We can nonetheless prove that the eigenvalue converges at least as fast as the eigenfunction; the doubling of the rate of convergence is deferred to the later Proposition 7, where we will have introduced additional hypotheses on F.

The following theorem gives then the above mentioned estimates on the convergence of the eigenfunction and eigenvalue. We start by showing the convergence to zero of the error, and use this result to show that the estimate is quasi optimal. We then show that the eigenvalue converges, with the basic rate mentioned above, and conclude by showing an estimate on the $L^2(\Omega)$ norm of the error.

Theorem 1. *If the hypotheses* (6a) *to* (6d) *on F hold and the hypotheses on the potential V* (7a), (7b) *hold, then*

$$\|u - u_\delta\|_{\mathrm{DG}} \to 0. \tag{29}$$

In particular, we have the quasi-optimal convergence

$$\|u - u_\delta\|_{\mathrm{DG}} \lesssim \inf_{v_\delta \in X_\delta} \|\!|\!| u - v_\delta |\!|\!|_{\mathrm{DG}}. \tag{30}$$

Furthermore,

$$|\lambda - \lambda_\delta| \lesssim \inf_{v_\delta \in X_\delta} |\!|\!| u - v_\delta |\!|\!|_{\mathrm{DG}}$$

and

$$\|u - u_\delta\| \lesssim \|u - u_\delta^*\|^r_{L^{6r/(5-s)}} + \|u - u_\delta\|^r_{L^{6r/(5-s)}} + \|u - u_\delta^*\|.$$

where r is defined in (6d) *and u_δ^* is the solution of the linear eigenvalue problem defined in* (15).

Proof. We start by proving (29), i.e. the convergence of the numerical solution towards the exact one. We have

$$\begin{aligned}2\left(E_\delta(u_\delta)-E(u)\right) &= \langle A^u_\delta u_\delta, u_\delta\rangle - \langle A^u u, u\rangle + \int_\Omega \left(F(u_\delta^2)-F(u^2)-f(u^2)(u_\delta^2-u^2)\right)\\ &= \langle (A^u_\delta-\lambda^*_\delta)\,(u_\delta-u^*_\delta), u_\delta-u^*_\delta\rangle - \lambda+\lambda^*_\delta\\ &\quad + \int_\Omega \left(F(u_\delta^2)-F(u^2)-f(u^2)(u_\delta^2-u^2)\right)\\ &\gtrsim \|u_\delta-u^*_\delta\|^2_{\mathrm{DG}} - |\lambda-\lambda^*_\delta| + \int_\Omega \left(F(u_\delta^2)-F(u^2)-f(u^2)(u_\delta^2-u^2)\right).\end{aligned}$$

Therefore, exploiting the convexity of F and the convergence of λ towards λ^*_δ, we have that

$$\begin{aligned}\|u_\delta-u^*_\delta\|^2_{\mathrm{DG}} &\lesssim E_\delta(u_\delta)-E(u)+|\lambda-\lambda^*_\delta|\\ &\leq E_\delta(\Pi_\delta u)-E_\delta(u)+|\lambda-\lambda^*_\delta|\to 0.\end{aligned} \tag{31}$$

Considering that u^*_δ converges towards u in the DG norm, (31) implies (29). Note then that

$$\begin{aligned}\lambda_\delta-\lambda^*_\delta &= \langle A^u_\delta u_\delta,u_\delta\rangle-\lambda^*_\delta+\int_\Omega \left[f(u_\delta^2)-f(u^2)\right]u_\delta^2\\ &= \langle (A^u_\delta-\lambda^*_\delta)\,(u_\delta-u^*_\delta)\,,u_\delta-u^*_\delta\rangle+\int_\Omega \left[f(u_\delta^2)-f(u^2)\right]u_\delta^2.\end{aligned} \tag{32}$$

Remarking, as in [CCM10, Proof of Theorem 1], that

$$\int_\Omega \left[f(u_\delta^2)-f(u^2)\right]u_\delta^2 \leq \|1+{u_\delta}^{2q+1}\|_{L^{6/(2q+1)}(\Omega)}\|u-u_\delta\|_{\mathrm{DG}}$$

and using (21) and (29), we can conclude that

$$|\lambda-\lambda_\delta| \lesssim |\lambda-\lambda^*_\delta|+\|u_\delta-u^*_\delta\|^2_{\mathrm{DG}}+\|u-u_\delta\|_{\mathrm{DG}}. \tag{33}$$

Now, from (19a) we have

$$\begin{aligned}\|u_\delta-u^*_\delta\|^2_{\mathrm{DG}} &\lesssim \langle (E''_\delta(u)-\lambda^*_\delta)\,(u_\delta-u^*_\delta),u_\delta-u^*_\delta\rangle\\ &= \langle (A^u_\delta-\lambda^*_\delta)\,(u_\delta-u^*_\delta),u_\delta-u^*_\delta\rangle+2\int_\Omega f'(u^2)u^2(u_\delta-u^*_\delta)^2\\ &= \langle (A^u_\delta-\lambda_\delta)\,u_\delta,u_\delta-u^*_\delta\rangle+(\lambda_\delta-\lambda^*_\delta)\|u_\delta-u^*_\delta\|^2+2\int_\Omega f'(u^2)u^2(u_\delta-u^*_\delta)^2\\ &= \int_\Omega \left[\left(f(u^2)-f({u_\delta}^2)\right)u_\delta+2f'(u^2)u^2(u_\delta-u^*_\delta)\right](u_\delta-u^*_\delta)+(\lambda_\delta-\lambda^*_\delta)\|u_\delta-u^*_\delta\|^2.\end{aligned}$$

Consider the first term: hypothesis (6c) gives

$$\int_\Omega f'(u^2)u^2(u_\delta-u^*_\delta)^2 \lesssim \int_\Omega f'(u^2)u^2(u_\delta-u)(u_\delta-u^*_\delta)+\|u-u^*_\delta\|\|u_\delta-u^*_\delta\|.$$

The two above equations and (6d) thus give

$$\|u_\delta-u^*_\delta\|^2_{\mathrm{DG}} \lesssim \|1+|u_\delta|^s\|_{L^{6/s}(\Omega)}\|u_\delta-u\|^r_{L^{6r/(5-s)}(\Omega)}\|u_\delta-u^*_\delta\|_{\mathrm{DG}}+|\lambda_\delta-\lambda^*_\delta|\|u_\delta-u^*_\delta\|^2+\|u-u^*_\delta\|\|u_\delta-u^*_\delta\|$$

and, since $r>1$ and $6r/(5-s)\leq 6$, we can conclude that

$$\|u-u_\delta\|_{\mathrm{DG}} \lesssim \|u-u^*_\delta\|_{\mathrm{DG}}.$$

The quasi optimality of u^*_δ then implies (30). Additionally, we can use this estimate in (33) and, considering that

$$|\lambda-\lambda^*_\delta| \lesssim |||u-u^*_\delta|||^2_{\mathrm{DG}} \lesssim \inf_{v_\delta\in X_\delta} |||u-v_\delta|||^2_{\mathrm{DG}}\,,$$

we conclude that

$$|\lambda - \lambda_\delta| \lesssim \inf_{v_\delta \in X_\delta} |||u - v_\delta|||_{\mathrm{DG}}\,.$$

Note that this result can be a bit sharper if q in (6b) is significantly smaller than 2; we write it this way for ease of reading. As already mentioned, we will prove a sharper result under some additional conditions in the following sections.

We finish by showing the estimate for the L^2 norm of the error. This follows from Lemma 4, since (26) implies

$$\begin{aligned}\|u_\delta - u_\delta^*\|^2 \lesssim \|u_\delta - u_\delta^*\|^r_{L^{6r/(5-s)}} \|\psi_{u_\delta - u_\delta^*}\|_{\mathrm{DG}} + |\lambda_\delta - \lambda_\delta^*|\|u_\delta - u_\delta^*\|\|\psi_{u_\delta - u_\delta^*}\| + \|u - u_\delta^*\|\|\psi_{u_\delta - u_\delta^*}\| \\ + \|u_\delta - u_\delta^*\|^2\|\psi_{u_\delta - u_\delta^*}\| + \|u_\delta - u_\delta^*\|^3\end{aligned} \tag{34}$$

for $\psi_{u_\delta - u_\delta^*} \in X_\delta$ defined as in (25), with $w_\delta = u_\delta - u_\delta^*$. Now, the coercivity of $\langle (E''(u) - \lambda_\delta^*)\cdot, \cdot\rangle$ over X_δ shown in (19a) and a Cauchy-Schwartz inequality imply

$$\|\psi_{u_\delta - u_\delta^*}\|_{\mathrm{DG}} \lesssim \|u_\delta - u_\delta^*\|. \tag{35}$$

Hence, from the combination of (34), (35), and the convergences of λ_δ towards λ_δ^* and of u_δ towards u_δ^* in the $L^2(\Omega)$ norm, we derive

$$\|u_\delta - u_\delta^*\| \lesssim \|u_\delta - u_\delta^*\|^r_{L^{6r/(5-s)}} + \|u - u_\delta^*\|.$$

Noting that

$$\|u - u_\delta\| \leq \|u - u_\delta^*\| + \|u_\delta - u_\delta^*\|$$

we conclude the proof. □

3.4. **Pointwise convergence.** We now wish to recover the doubling of the convergence rate normally obtained for the eigenvalue error, with respect to the eigenfunction. We need to introduce a conjecture on the local norm of discrete solutions to problems with localised right hand side, which will influence the convergence of the error $\|u - u_\delta\|_{L^\infty(\Omega)}$.

Assumption 1. *We assume that, if $g_\delta \in X_\delta$ is the solution to*

$$\langle (E''(u) - \lambda_\delta^*) g_\delta, v_\delta\rangle = (\rho, v_\delta) \qquad \textit{for all } v_\delta \in X_\delta \tag{36}$$

with ρ such that

$$\begin{aligned}&\mathrm{supp}(\rho) = \widetilde{K} \textit{ for a } \widetilde{K} \in \mathcal{T}\\ &\|\rho\|_{L^p(\Omega)} = h_{\widetilde{K}}^{d\frac{1-p}{p}} \textit{ for all } p \in [1,2]\end{aligned}$$

then,

$$\sum_{j=1}^{\ell} h_j^{1/2} \|g_\delta\|_{\mathrm{DG}(\Omega_j)} \leq C, \tag{37}$$

where the constant C does not depend on $h_{\widetilde{K}}$ and the sets Ω_j have been introduced in Section 2.3.

We furthermore introduce $\widetilde{u}_\delta \in X_\delta$ as the discontinuous Galerkin projection of u for the operator A^u, i.e., such that

$$\langle A^u_\delta \widetilde{u}_\delta, v_\delta\rangle = \langle A^u u, v_\delta\rangle \text{ for all } v_\delta \in X_\delta. \tag{38}$$

Denote also

$$p_{\max} = \max_{K\in\mathcal{T}} p_K\,.$$

Proposition 5. *Suppose that the hypotheses of Theorem 1 hold. Furthermore, suppose that* (37) *holds and that at least one of the following is true: either*

$$p_{\max}^d \|u - u_\delta\|_{\mathrm{DG}}^{r-1} \to 0, \tag{39}$$

or

$$s < 4 - r, \tag{40}$$

where s *and* r *are defined in* (6d)*. Then,*

$$\|u - u_\delta\|_{L^\infty(\Omega)} \lesssim p_{\max}^d \left(\|u - u_\delta\|_{\mathrm{DG}}^r + \|u - u_\delta\|_{L^2(\Omega)} + |\lambda - \lambda_\delta| + |\lambda - \lambda_\delta^*| + \|u - \widetilde{u}_\delta\|_{L^\infty(\Omega)} \right),$$

where $\widetilde{u}_\delta \in X_\delta$ *is defined as in* (38)*.*

Proof. We prove the theorem assuming that (39) holds; at the end we will delineate the necessary modifications in case only (40) holds. The $L^\infty(\Omega)$ error between u_δ and u can be split in two parts, as

$$\|u - u_\delta\|_{L^\infty(\Omega)} \leq \|u - \widetilde{u}_\delta\|_{L^\infty(\Omega)} + \|u_\delta - \widetilde{u}_\delta\|_{L^\infty(\Omega)} \tag{41}$$

The first term of the right hand side of the inequality above is the $L^\infty(\Omega)$ norm of the error for a linear problem. We now consider the second part of the right hand side of (41),

$$\|\widetilde{u}_\delta - u_\delta\|_{L^\infty(\Omega)} = \|\widetilde{u}_\delta - u_\delta\|_{L^\infty(\widetilde{K})}$$

for a $\widetilde{K} \in \mathcal{T}$. An inverse inequality gives

$$\begin{aligned}\|\widetilde{u}_\delta - u_\delta\|_{L^\infty(\Omega)} &\lesssim h_{\widetilde{K}}^{-d/2} p_{\widetilde{K}}^d \|\widetilde{u}_\delta - u_\delta\|_{L^2(\widetilde{K})} \\ &= p_{\widetilde{K}}^d (\rho, \widetilde{u}_\delta - u_\delta),\end{aligned}$$

where we have chosen ρ as

$$\rho = h_{\widetilde{K}}^{-d/2} \frac{\widetilde{u}_\delta - u_\delta}{\|\widetilde{u}_\delta - u_\delta\|_{L^2(\widetilde{K})}} \mathbb{1}_{\widetilde{K}}.$$

We now introduce the finite element function g_δ as the solution of (36).

Then, we have

$$(\rho, \widetilde{u}_\delta - u_\delta) = \langle (A_\delta^u - \lambda_\delta^*)\, g_\delta, \widetilde{u}_\delta - u_\delta \rangle + 2 \int_\Omega f'(u^2) u^2 g_\delta (\widetilde{u}_\delta - u_\delta). \tag{42}$$

Due to the definition of $\widetilde{u}_\delta$ and the symmetry of the bilinear form,

$$\langle (A_\delta^u - \lambda_\delta^*)\, g_\delta, \widetilde{u}_\delta - u_\delta \rangle = \lambda (u - \widetilde{u}_\delta, g_\delta) + (\lambda - \lambda_\delta^*)(\widetilde{u}_\delta, g_\delta) - \langle (A_\delta^u - \lambda_\delta^*)\, g_\delta, u_\delta \rangle \tag{43}$$

We can treat the second term by noting that

$$-\langle (A_\delta^u - \lambda_\delta^*)\, u_\delta, g_\delta \rangle = \int_\Omega \left[f({u_\delta}^2) - f(u^2) \right] g_\delta u_\delta + (\lambda_\delta^* - \lambda_\delta)(u_\delta, g_\delta) \tag{44}$$

We want to use (6d) on the integrals containing f and its derivative in (42) and (44). We start by noting that

$$2 \int_\Omega f'(u^2) u^2 g_\delta (\widetilde{u}_\delta - u_\delta) = 2 \int_\Omega f'(u^2) u^2 g_\delta (u - u_\delta) + 2 \int_\Omega f'(u^2) u^2 g_\delta (\widetilde{u}_\delta - u)$$

Therefore, by the equation above,

$$\int_\Omega \left[f({u_\delta}^2) - f(u^2)\right] g_\delta u_\delta + 2\int_\Omega f'(u^2)u^2 g_\delta(\widetilde{u}_\delta - u_\delta)$$
$$= \int_\Omega \left[f({u_\delta}^2)u_\delta - f(u^2)u_\delta - 2f'(u^2)u^2(u_\delta - u)\right] g_\delta + 2\int_\Omega f'(u^2)u^2 g_\delta(\widetilde{u}_\delta - u)$$

and by (6c), (6d), and the Cauchy-Schwartz inequality,

$$\left|\int_\Omega \left[f({u_\delta}^2) - f(u^2)\right] g_\delta u_\delta + 2\int_\Omega f'(u^2)u^2 g(\widetilde{u}_\delta - u_\delta)\right| \lesssim \int_\Omega (1 + |u_\delta|^s)|u - u_\delta|^r |g_\delta| + \|g_\delta\|\|u - \widetilde{u}_\delta\|.$$

Combining (42), (43), and (44) with the above equation gives

$$(\rho, \widetilde{u}_\delta - u_\delta) \lesssim \int_\Omega (1 + |u_\delta|^s)|u - u_\delta|^r |g_\delta| + \|g_\delta\|_{L^2(\Omega)}\|u - \widetilde{u}_\delta\| + (\lambda^*_\delta - \lambda_\delta)(u_\delta, g_\delta) + \lambda(\widetilde{u}_\delta - u_\delta, g_\delta). \tag{45}$$

A Hölder inequality and the condition $s < 5 - r$ imply that there exists an

$$0 < \alpha \leq \frac{15 - 3(s + r)}{7 - s - r}$$

such that

$$(\rho, \widetilde{u}_\delta - u_\delta) \lesssim \left(1 + \|u_\delta\|^s_{L^6(\Omega)}\right) \|u - u_\delta\|^{r-1}_{L^6(\Omega)}\|g_\delta\|_{L^{3-\alpha}(\Omega)}\|u - u_\delta\|_{L^\infty(\Omega)} + \|g_\delta\|_{L^2(\Omega)} \left(\|u - \widetilde{u}_\delta\|_{L^2(\Omega)} + \|u - u_\delta\|_{L^2(\Omega)} + |\lambda^*_\delta - \lambda_\delta|\right).$$

Consider now that

$$\mathcal{J}^1_{1/2}(\Omega) \hookrightarrow H^{1/2-\alpha'}(\Omega) \hookrightarrow L^{3-\alpha}(\Omega), \tag{46}$$

with $\alpha' = \alpha/(3 - \alpha)$, see [Nic97] for the first embedding; the second one is classical in Sobolev spaces. The double embedding (46) then implies

$$\begin{aligned}\|g_\delta\|_{L^{3-\alpha}(\Omega)} &\leq C\|g_\delta\|_{\mathcal{J}^1_{1/2}(\Omega)} \\ &\leq C\left(\sum_j \|g_\delta\|^2_{\mathcal{J}^1_{1/2}(\Omega_j)}\right)^{1/2} \\ &\leq C\left(\sum_j h_j\|g_\delta\|^2_{H^1(\Omega_j)}\right)^{1/2}\end{aligned}$$

where the second inequality follows from the fact that $r_{|\Omega_j}/h_j \leq C$ for all $j = 1, \dots, \ell$. Therefore, using (37) and noting that the $\ell^2(\{1, \dots, \ell\})$ norm is bounded by the $\ell^1(\{1, \dots, \ell\})$ norm with constants that do not depend on ℓ, we conclude that $\|g_\delta\|_{L^{3-\alpha}(\Omega)} \leq C$ for any positive α, thus,

$$(\rho, \widetilde{u}_\delta - u_\delta) \lesssim \|u - u_\delta\|^{r-1}_{L^6(\Omega)}\|u - u_\delta\|_{L^\infty(\Omega)} + \|u - \widetilde{u}_\delta\|_{L^2(\Omega)} + \|u - u_\delta\|_{L^2(\Omega)} + |\lambda - \lambda_\delta| + |\lambda - \lambda^*_\delta|.$$

If hypothesis (39) holds, we can conclude with the thesis. If (39) does not hold, then hypothesis (40) is necessary: the proof follows the same lines, though at (45) we use the inequality

$$\int_\Omega (1 + |u_\delta|^s)|u - u_\delta|^r |g_\delta| \leq C\|u_\delta\|^s_{L^6(\Omega)}\|u - u_\delta\|^r_{L^6(\Omega)}\|g_\delta\|_{L^{3-\alpha}(\Omega)}$$

for an
$$0 < \alpha \leq \frac{12 - 3s - 3r}{6 - s - r}.$$
Note that such an α exists thanks to (40). In this case we conclude
$$(\rho, \widetilde{u}_\delta - u_\delta) \lesssim \|u - u_\delta\|^r_{L^6(\Omega)} + \|u - \widetilde{u}_\delta\|_{L^2(\Omega)} + \|u - u_\delta\|_{L^2(\Omega)} + |\lambda - \lambda_\delta| + |\lambda - \lambda^*_\delta|,$$
hence the thesis. □

The following statement is then a direct consequence of Proposition 5.

Corollary 6. *Under the hypotheses of Proposition 5 and if the following hold:*

$$p^d_{\max} \inf_{v_\delta \in X_\delta} \|u - v_\delta\|_{\mathrm{DG}} \to 0 \tag{47a}$$

$$p^d_{\max} \|u - \widetilde{u}_\delta\|_{L^\infty(\Omega)} \to 0, \tag{47b}$$

then
$$\|u - u_\delta\|_{L^\infty(\Omega)} \to 0.$$

Remark 5. *We remark that* (47a) *depends on the regularity of* u *(which in turn depends on the regularity of* F*).*

Furthermore, in the case of isotropically refined hp *methods, assuming* (47b) *is pleonastic, as it is a consequence of* (37)*, by a similar argument to the one used in the proof of Proposition 5.*

3.5. **Convergence revisited.** In this section we finally show that, if the solution converges in the $L^\infty(\Omega)$ norm, we can prove that the eigenvalue converges with the same rate as the square of the eigenfunction.

When dealing with h type discontinuous finite element methods, $p_{\max}$ is bounded by some global constant, thus (47b) translates into the requirement of simple $L^\infty(\Omega)$ convergence for the linear problem. Furthermore, if the mesh is globally regular, (47b) and (37) can be proven if the solution is sufficiently regular, see [CC04]. Therefore, the following theorem can be extended to h-type finite element methods too, as long as the solution is sufficiently regular. We will not treat this case, as it it outside the focus of our analysis.

Under the forthcoming additional hypothesis (50) on the regularity of F, we prove that the eigenvalues converge as the square of the eigenfunction.

We introduce another adjoint problem: let $\psi \in u^\perp$ such that
$$\langle (E''(u) - \lambda)\, \psi, v\rangle = (f'(u^2)u^3, v) \tag{48}$$
for all $v \in \{v \in X(\delta) : (u, v) = 0\}$.

Remark 6. ψ *satisfies the equation*
$$\left(A^u + 2f'(u^2)u^2 - \lambda\right)\psi = 2\left(\int_\Omega f'(u^2)u^3\psi\right)u + f'(u^2)u^3 - (f'(u^2)u^3, u)u. \tag{49}$$
The regularity of ψ *depends therefore on the regularity of* f*,* f'*, and* u*.*

Proposition 7. *Suppose that the hypotheses of Theorem 1 and Proposition 5 hold, and that conditions* (47a) *and* (47b) *hold. Furthermore, suppose that* (6d) *holds with* $r = 2$ *and that*
$$F \in C^3((0, +\infty), \mathbb{R}), \text{ and } F'''(t)t^2 \text{ is locally bounded in } [0, +\infty). \tag{50}$$
Then
$$|\lambda - \lambda_\delta| \lesssim |||u - u_\delta|||_{\mathrm{DG}} \left(\inf_{v_\delta \in X_\delta} |||\psi - v_\delta|||_{\mathrm{DG}} + |||u - u_\delta|||_{\mathrm{DG}}\right),$$
where ψ *is defined in* (48) *above.*

Proof. The proof begins similarly to (32):

$$\lambda_\delta - \lambda = \langle (A^u_\delta - \lambda)(u - u_\delta), u - u_\delta \rangle + \int_\Omega \left[f(u_\delta^2) - f(u^2) \right] {u_\delta}^2$$

$$= \langle (A^u_\delta - \lambda)(u_\delta - u), u_\delta - u \rangle + \int_\Omega \left[f(u_\delta^2) - f(u^2) - f'(u^2)({u_\delta}^2 - u^2) \right] u_\delta^2 \tag{51a}$$

$$+ \int_\Omega f'(u^2) \left[{u_\delta}^2 (u + u_\delta) - 2u^3 \right] (u - u_\delta) \tag{51b}$$

$$+ \int_\Omega 2 f'(u^2) u^3 (u - u_\delta). \tag{51c}$$

We consider the three integrals in the last equation separately. Firstly, considering term (51b), we have

$$\int_\Omega f'(u^2) \left[{u_\delta}^2 (u + u_\delta) - 2u^3 \right] (u - u_\delta) = \int_\Omega f'(u^2)(u^2 + 2uu_\delta + 2{u_\delta}^2)(u - u_\delta)^2$$
$$\lesssim \int_\Omega \left(1 + \frac{u_\delta}{u} + \frac{{u_\delta}^2}{u^2} \right) (u - u_\delta)^2.$$

Thanks to the Cauchy-Schwartz inequality, to the assumed $L^\infty(\Omega)$ convergence of u_δ towards u (see Corollary 6), and to the fact that there exists $u_{\min}$ such that $u \geq u_{\min} > 0$, see [CCM10], the above inequality implies that (after a certain level of refinement)

$$\left| \int_\Omega f'(u^2) \left[{u_\delta}^2 (u + u_\delta) - 2u^3 \right] (u - u_\delta) \right| \lesssim \|u - u_\delta\|^2. \tag{52}$$

Integral (51c) is then treated by using (48)

$$\int_\Omega f'(u^2) u^3 (u - u_\delta) = (f'(u^2) u^3, (u - u_\delta)^{u^\perp}) + (f'(u^2)u^3, (u - u_\delta, u)u)$$
$$= \langle (E''(u) - \lambda)\,\psi, (u - u_\delta)^{u^\perp} \rangle + \frac{1}{2} \|u - u_\delta\|^2 \|f'(u^2)u^3\|_{L^2} \tag{53}$$

Consider the first term above: for any $\tilde{v}_\delta \in u_\delta^\perp$,

$$\langle (E''(u) - \lambda)\,\psi, (u - u_\delta)^{u^\perp} \rangle = \langle (E''_\delta(u) - \lambda)(\psi - \tilde{v}_\delta), (u - u_\delta)^{u^\perp} \rangle + \langle (E''_\delta(u) - \lambda)\,\tilde{v}_\delta, (u - u_\delta)^{u^\perp} \rangle, \tag{54}$$

Now,

$$\langle (E''_\delta(u) - \lambda)\,\tilde{v}_\delta, (u - u_\delta)^{u^\perp} \rangle = -\langle (A^u_\delta - \lambda)\, u_\delta, \tilde{v}_\delta \rangle + 2 \int_\Omega f'(u^2) u^2 (u - u_\delta)^{u^\perp} \tilde{v}_\delta$$
$$= \int_\Omega \left(f({u_\delta}^2) - f(u^2) \right) u_\delta \tilde{v}_\delta + (\lambda - \lambda_\delta)(u_\delta, \tilde{v}_\delta)$$
$$+ 2 \int_\Omega f'(u^2) u^2 (u - u_\delta) \tilde{v}_\delta + \|u - u_\delta\|^2 \int_\Omega f'(u^2) u^3$$

thus

$$\langle (E''_\delta(u) - \lambda)\,\tilde{v}_\delta, (u - u_\delta)^{u^\perp} \rangle \lesssim \|u - u_\delta\|^2 \|\tilde{v}_\delta\|_{L^\infty(\Omega)} + \|u - u_\delta\|^2,$$

where we have used the fact that (6d) holds with $r = 2$, the orthogonality between $\tilde{v}_\delta$ and u_δ, condition (6c), and the $L^\infty(\Omega)$ convergence of u_δ towards u. We now turn to the first term at the right hand side of (54). We have

$$\langle (E''_\delta(u) - \lambda)(\psi - \tilde{v}_\delta), (u - u_\delta)^{u^\perp} \rangle \lesssim \left(|||u - u_\delta|||_{\mathrm{DG}} + \|u - u_\delta\|^2\, |||u|||_{\mathrm{DG}} \right) |||\psi - \tilde{v}_\delta|||_{\mathrm{DG}}\,.$$

Note that the term in $\|u-u_\delta\|^2$ is of higher order, so we can omit it from the following estimates. We have therefore, from (54),

$$\begin{aligned}\langle (E''(u)-\lambda)\,\psi,(u-u_\delta)^{u^\perp}\rangle &\lesssim \inf_{\tilde v_\delta\in u_\delta^\perp}\left[\|u-u_\delta\|^2\left(\|\tilde v_\delta\|_{L^\infty(\Omega)}+1\right)+|||u-u_\delta|||_{\mathrm{DG}}\,|||\psi-\tilde v_\delta|||_{\mathrm{DG}}\right]\\ &\lesssim \inf_{v_\delta\in X_\delta}\left[\|u-u_\delta\|^2\left(\|v_\delta-(v_\delta,u_\delta)u_\delta\|_{L^\infty(\Omega)}+1\right)\right.\\ &\qquad\left.+\,|||u-u_\delta|||_{\mathrm{DG}}\,|||\psi-v_\delta+(v_\delta,u_\delta)u_\delta|||_{\mathrm{DG}}\right],\end{aligned}$$

where we have replaced $\tilde v_\delta$ by $v_\delta-(v_\delta,u_\delta)u_\delta$, thus being able to extend the inf over all v_δ in X_δ. Now,

$$\begin{aligned}|||\psi-v_\delta+(v_\delta,u_\delta)u_\delta|||_{\mathrm{DG}} &\le |||\psi-v_\delta|||_{\mathrm{DG}}+|||(\psi,u_\delta)u_\delta|||_{\mathrm{DG}}+|||(\psi-v_\delta,u_\delta)u_\delta|||_{\mathrm{DG}}\\ &\lesssim |||\psi-v_\delta|||_{\mathrm{DG}}+\|\psi-v_\delta\|+\|\psi\|\|u-u_\delta\|.\end{aligned}\tag{55}$$

Furthermore,

$$\|v_\delta-(v_\delta,u_\delta)u_\delta\|_{L^\infty(\Omega)}\le\|v_\delta\|_{L^\infty(\Omega)}+\|u_\delta\|_{L^\infty(\Omega)}\|v_\delta\|.\tag{56}$$

The norm $\|u_\delta\|_{L^\infty(\Omega)}$ is bounded due to the assumed $L^\infty(\Omega)$-convergence of u_δ towards u. The best approximation v_δ to ψ in the $|||\cdot|||_{\mathrm{DG}}$ norm is such that $\|\psi-v_\delta\|_{L^\infty\Omega}\to 0$: this follows from an inverse inequality and from (47a). Furthermore, the norm $\|\psi\|_{L^\infty(\Omega)}$ can be bounded by a constant depending on u by Remark 6 and elliptic regularity, hence $\|v_\delta-(v_\delta,u_\delta)u_\delta\|_{L^\infty(\Omega)}\le C$. Using these remarks, (55), and (56), inequality (54) can be rewritten as

$$\langle (E''_\delta(u)-\lambda)\,\psi,(u-u_\delta)^{u^\perp}\rangle\lesssim |||u-u_\delta|||_{\mathrm{DG}}\left(\inf_{v_\delta\in X_\delta}|||\psi-v_\delta|||_{\mathrm{DG}}+\|u-u_\delta\|\right),$$

where, once again, we have omitted the higher order terms. Going back to (53) we obtain

$$\left|\int_\Omega f'(u^2)u^3(u-u_\delta)\right|\lesssim |||u-u_\delta|||_{\mathrm{DG}}\left(\inf_{v_\delta\in X_\delta}|||\psi-v_\delta|||_{\mathrm{DG}}+\|u-u_\delta\|\right)\tag{57}$$

We finally consider the second term of line (51a). Under the additional hypotheses $F\in C^3$ and $t^2F'''(t)$ locally bounded in $[0,\infty)$, denoting $w=[f(u_\delta^2)-f(u^2)-f'(u^2)({u_\delta}^2-u^2)]$ and recalling that $u\ge u_{\min}>0$,

$$\begin{aligned}\int_\Omega w{u_\delta}^2 &= \int_\Omega\left(\int_0^1 tf''(u^2+t({u_\delta}^2-u^2))dt\right){u_\delta}^2({u_\delta}^2-u^2)^2\\ &\lesssim \int_\Omega\left(\int_0^1\frac{t}{(u^2+t({u_\delta}^2-u^2))^2}dt\right){u_\delta}^2({u_\delta}^2-u^2)^2\\ &= \int_\Omega\left|{u_\delta}^2\log\left(\frac{u^2}{{u_\delta}^2}\right)+{u_\delta}^2-u^2\right|\left|{u_\delta}^2-u^2\right|\end{aligned}$$

Under the hypothesis of $L^\infty(\Omega)$ convergence given in Proposition 5, then,

$$\left|\int_\Omega\left[f(u_\delta^2)-f(u^2)-f'(u^2)({u_\delta}^2-u^2)\right]u_\delta^2\right|\lesssim\|u-u_\delta\|^2.\tag{58}$$

The thesis follows from (51a)–(51c), (52), (57), and (58). □

4. Weighted analytic regularity estimates for polynomial nonlinearities

This section is centered on the proof of analytic-type estimates on the norms of the solution to nonlinear elliptic problems. Specifically, we consider the nonlinear Schrödinger equation and prove that, under some conditions on the coefficients of the operator, the solution belongs to $\mathcal{J}^\varpi_\gamma(\Omega)$, for the same γ as in the linear case seen in [MM19]. Since the singularities we consider are internal to the domain, we suppose that the domain Ω is a compact domain without boundary, e.g., $\Omega = (-1,1)^d/2\mathbb{Z}$. The extension of the theory to the case of a bounded domain with smooth boundary can be done using the classical tools used in the analysis of elliptic problems in Sobolev spaces, as long as $r_{|\partial\Omega} \simeq 1$, i.e., the singularity is bounded away from the boundary.

First, in Section 4.1 we prove the local elliptic estimate in weighted Sobolev spaces that will allow for the derivation of the bounds on higher order derivatives from those obtained on lower order ones. Then, in order to estimate the norms of the nonlinear terms, we follow the proof technique used in [DFØSS12]. The idea is to proceed by induction and to consider L^p norms in nested balls and with a big enough p. Let L_{lin} be an elliptic linear operator and consider the operator $L_u u = -L_{\mathrm{lin}} u + |u|^{\delta-1}u$, where $\delta = 2,3,4$: the L^p norms of the nonlinear terms can then be broken up into products of $L^{\delta p}$ norms by a Cauchy-Schwartz inequality. In order to get back to L^p norms, in [DFØSS12] the authors use an interpolation inequality where $L^{\delta p}$ is contained in an interpolation space between L^p and $W^{1,p}$. Since in our case we need to deal with weighted spaces, in Section 4.2 we derive the weighted version of this inequality, via a dyadic decomposition of the domain near the singular points.

The proof of the analytic bound on a nonlinear scalar elliptic eigenvalue problem is then given in Section 4.3, in the case of the nonlinear Schrödinger equation up to a quartic nonlinear term. Starting from a basic regularity assumption, we are able to treat the potential and the nonlinear term thanks to the results presented in the preceding sections.

We suppose, for the sake of simplicity, the presence of a single point singularity, i.e., that $\mathfrak{C} = \{\mathfrak{c}\}$.

We denote the commutator by square brackets, i.e., we write

$$[A,B] = AB - BA.$$

4.1. Local elliptic estimate.

We start by proving a local seminorm estimate in weighted Sobolev spaces. This has been already established in [CDN12], as an intermediate estimate leading to the proof of another regularity result. We restate it here fully, in the specific form that will be needed in the sequel. The goal is to control the weighted norm of a higher order derivative of a function with the weighted norm of its Laplacian and of lower order derivatives in a bigger domain, while giving an explicit dependence of the constants on the distance between the domains.

Proposition 8. *Let $1 < p < \infty$, $R > 0$, $k \in \mathbb{N}$ and $\rho \in (0, \frac{R}{2(k+1)})$. Furthermore, let $\gamma \in \mathbb{R}$ and $j \in \mathbb{N}$ such that $1 \le j \le k$. Then,*

$$\sum_{|\alpha|=k+1} \|r^{k+1-\gamma}\partial^\alpha u\|_{L^p(B_{R-(j+1)\rho})} \le C_{\mathrm{reg}} \left(\sum_{|\beta|=k-1} \|r^{k+1-\gamma}\partial^\beta(\Delta u)\|_{L^p(B_{R-j\rho})} + \sum_{|\alpha|=k} \rho^{-1}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-j\rho})} + \sum_{|\alpha|=k-1} \rho^{-2}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-j\rho})} \right), \tag{59}$$

with C_{reg} independent of u, k, j, and ρ.

In order to prove this Proposition we introduce a smooth cutoff function $\psi \in C_0^\infty(B_{R-j\rho})$ such that for any $\alpha \in \mathbb{N}^d$, $|\alpha| \leq 2$,

$$0 \leq \psi \leq 1, \qquad \psi = 1 \text{ on } B_{R-(j+1)\rho}, \qquad |\partial^\alpha \psi| \leq C\rho^{-|\alpha|}, \tag{60}$$

and we derive an auxiliary estimate.

Lemma 9. *Let $\beta \in \mathbb{N}^d$, $1 < p < \infty$, $R > 0$, and $\rho \in (0, \frac{R}{2(|\beta|+2)})$. Then, for any $j \in \mathbb{N}$ such that $1 \leq j \leq |\beta| + 1$, there exists $C = C(R)$*

$$\sum_{|\alpha|=2} \| \left[\partial^\alpha, r^{|\beta|+2-\gamma}\right] \psi \partial^\beta u\|_{L^p(B_{R-j\rho})} \leq C \sum_{|\alpha|\leq 1} \rho^{-2+|\alpha|} \|r^{|\beta|+|\alpha|-\gamma}\partial^{\alpha+\beta}u\|_{L^p(B_{R-j\rho})} \tag{61}$$

Proof. First, let us fix $i, k \in \{1, \dots, d\}$ such that $\partial^\alpha = \dfrac{\partial^2}{\partial x_i \partial x_k} = \partial_{ik} = \partial_i\partial_k$. In the proof, we indicate by $a \lesssim b$ when there exists a constant C independent of u, j, $|\beta|$, and ρ such that $a \leq Cb$. Then,

$$\left[\partial^\alpha, r^{|\beta|+2-\gamma}\right] \psi\partial^\beta u = \left(\partial_{ik} r^{|\beta|+2-\gamma}\right) \psi \partial^\beta u + \left(\partial_i r^{|\beta|+2-\gamma}\right) \partial_k \left(\psi \partial^\beta u\right) + \left(\partial_k r^{|\beta|+2-\gamma}\right) \partial_i \left(\psi\partial^\beta u\right)$$

Writing $(\star) = \| \left[\partial^\alpha, r^{|\beta|+2-\gamma}\right] \psi\partial^\beta u\|_{L^p(B_{R-j\rho})}$, we have that

$$\begin{aligned}
(\star) &\leq \| \left(\partial_{ik} r^{|\beta|+2-\gamma}\right) \psi\partial^\beta u\|_{L^p(B_{R-j\rho})} + \| \left(\partial_i r^{|\beta|+2-\gamma}\right) \partial_k \left(\psi\partial^\beta u\right) \|_{L^p(B_{R-j\rho})} \\
&\qquad + \| \left(\partial_k r^{|\beta|+2-\gamma}\right) \partial_i \left(\psi\partial^\beta u\right) \|_{L^p(B_{R-j\rho})} \\
&\lesssim \left(|\beta|^2 + \delta_{ik}|\beta|\right) \|r^{|\beta|-\gamma}\psi\partial^\beta u\|_{L^p(B_{R-j\rho})} + |\beta|\|r^{|\beta|+1-\gamma}\partial_k \left(\psi\partial^\beta u\right) \|_{L^p(B_{R-j\rho})} \\
&\qquad + |\beta|\|r^{|\beta|+1-\gamma}\partial_i \left(\psi\partial^\beta u\right) \|_{L^p(B_{R-j\rho})},
\end{aligned}$$

where $\delta_{ik} = 1$ if $i = k$, $\delta_{ik} = 0$ otherwise. Now, for $\iota \in \{i, k\}$,

$$\begin{aligned}
&|\beta|\|r^{|\beta|+1-\gamma}\partial_\iota \left(\psi\partial^\beta u\right) \|_{L^p(B_{R-j\rho})} \\
&\qquad \lesssim |\beta|\|r^{|\beta|+1-\gamma}\psi\partial_\iota\partial^\beta u\|_{L^p(B_{R-j\rho})} + |\beta|\|r^{|\beta|+1-\gamma} \left[\psi, \partial_\iota\right] \partial^\beta u\|_{L^p(B_{R-j\rho})} \\
&\qquad \lesssim |\beta|\|r^{|\beta|+1-\gamma}\psi\partial_\iota\partial^\beta u\|_{L^p(B_{R-j\rho})} + |\beta|\|r^{|\beta|+1-\gamma} \left(\partial_\iota\psi\right) \partial^\beta u\|_{L^p(B_{R-j\rho})}.
\end{aligned}$$

By the definition of ψ given in (60), we obtain

$$(\star) \lesssim \left(|\beta|^2 + \delta_{ik}|\beta| + |\beta|\rho^{-1}\right) \|r^{|\beta|-\gamma}\partial^\beta u\|_{L^p(B_{R-j\rho})} + \sum_{\iota\in\{i,k\}} |\beta|\|r^{|\beta|+1-\gamma}\partial^\beta\partial_\iota u\|_{L^p(B_{R-j\rho})}.$$

Summing over all multi indices $|\alpha| = 2$,

$$\begin{aligned}
\sum_{|\alpha|=2} \| \left[\partial^\alpha, r^{|\beta|+2-\gamma}\right] \psi\partial^\beta u\|_{L^p(B_{R-j\rho})} &\lesssim \left(|\beta|^2 + |\beta|\rho^{-1}\right) \|r^{|\beta|-\gamma}\partial^\beta u\|_{L^p(B_{R-j\rho})} \\
&\qquad + \sum_{|\alpha|=1} |\beta|\|r^{|\beta|+1-\gamma}\partial^{\beta+\alpha}u\|_{L^p(B_{R-j\rho})}.
\end{aligned}$$

Since $\rho \in (0, \frac{R}{2(|\beta|+2)})$ implies $|\beta| \leq R\rho^{-1}$, we can conclude with (61). ☐

We can now prove estimate (59).

Proof of Proposition 8. Let us consider a multi index β such that $|\beta| = k-1$. We indicate by $a \lesssim b$ when there exists a constant C independent of u, j, k, and ρ such that $a \leq Cb$. First,

$$\sum_{|\alpha|=2} \|r^{|\beta|+2-\gamma}\partial^{\alpha+\beta}u\|_{L^p(B_{R-(j+1)\rho})} \leq \sum_{|\alpha|=2}\Big\{\|\partial^\alpha\left(r^{|\beta|+2-\gamma}\partial^\beta u\right)\|_{L^p(B_{R-(j+1)\rho})} + \|\left[\partial^\alpha, r^{|\beta|+2-\gamma}\right]\partial^\beta u\|_{L^p(B_{R-(j+1)\rho})}\Big\}. \tag{62}$$

We consider the first term at the right hand side: using (60)

$$\sum_{|\alpha|=2} \|\partial^\alpha\left(r^{|\beta|+2-\gamma}\partial^\beta u\right)\|_{L^p(B_{R-(j+1)\rho})} \leq \sum_{|\alpha|=2} \|\partial^\alpha\left(r^{|\beta|+2-\gamma}\psi\partial^\beta u\right)\|_{L^p(B_{R-j\rho})}$$

and by elliptic regularity and using the triangular inequality

$$\begin{aligned}\sum_{|\alpha|=2} \|\partial^\alpha\left(r^{|\beta|+2-\gamma}\psi\partial^\beta u\right)\|_{L^p(B_{R-j\rho})} &\leq C\|\Delta\left(r^{|\beta|+2-\gamma}\psi\partial^\beta u\right)\|_{L^p(B_{R-j\rho})}\\ &\leq C\|r^{|\beta|+2-\gamma}\psi\Delta\partial^\beta u\|_{L^p(B_{R-j\rho})}\\ &\quad + C\|\left[\Delta, r^{|\beta|+2-\gamma}\right]\psi\partial^\beta u\|_{L^p(B_{R-j\rho})}\\ &\quad + C\|r^{|\beta|+2-\gamma}\left[\Delta,\psi\right]\partial^\beta u\|_{L^p(B_{R-j\rho})}.\end{aligned}$$

Combining the last inequality with (62) we obtain

$$\begin{aligned}\sum_{|\alpha|=2} \|r^{|\beta|+2-\gamma}\partial^{\alpha+\beta}u\|_{L^p(B_{R-(j+1)\rho})} &\lesssim \|r^{|\beta|+2-\gamma}\psi\partial^\beta\left(\Delta u\right)\|_{L^p(B_{R-j\rho})}\\ &\quad + \sum_{i=1}^d\Big\{\|r^{|\beta|+2-\gamma}\left(\partial_{ii}\psi\right)\partial^\beta u\|_{L^p(B_{R-j\rho})} + \|r^{|\beta|+2-\gamma}\left(\partial_i\psi\right)\partial^\beta\partial_i u\|_{L^p(B_{R-j\rho})}\Big\}\\ &\quad + \sum_{|\alpha|=2}\|\left[\partial^\alpha, r^{|\beta|+2-\gamma}\right]\psi\partial^\beta u\|_{L^p(B_{R-j\rho})}.\end{aligned}$$

The bounds on the derivatives of ψ given in (60) and (61) then give

$$\begin{aligned}\sum_{|\alpha|=2} \|r^{|\beta|+2-\gamma}\partial^{\alpha+\beta}u\|_{L^p(B_{R-(j+1)\rho})} &\lesssim \|r^{|\beta|+2-\gamma}\psi\partial^\beta\left(\Delta u\right)\|_{L^p(B_{R-j\rho})}\\ &\quad + \sum_{|\alpha|\leq 1}\rho^{-2+|\alpha|}\|r^{|\beta|+|\alpha|-\gamma}\partial^{\alpha+\beta}u\|_{L^p(B_{R-j\rho})}.\end{aligned}$$

We can now sum over all multi indices β such that $|\beta| = k-1$ to obtain the thesis (59). □

4.2. **Weighted interpolation estimate.**

Lemma 10. *Let $R > 0$ such that $B_R \subset \Omega$, $\beta \in \mathbb{N}_0^d$, $u \in \mathcal{K}_\gamma^{|\beta|+1,p}(B_R)$, $\delta > 1$, $\gamma - d/p \geq 2/(1-\delta)$, and $p \geq d(1-1/\delta)$. Then, the following "interpolation" estimate holds*

$$\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_R)} \leq C\|r^{|\beta|-\gamma}\partial^\beta u\|_{L^p(B_R)}^{1-\vartheta}\Bigg\{(|\beta|+1)^\vartheta\|r^{|\beta|-\gamma}\partial^\beta u\|_{L^p(B_R)}^\vartheta + \sum_{|\alpha|=1}\|r^{|\beta|+1-\gamma}\partial^{\beta+|\alpha|}u\|_{L^p(B_R)}^\vartheta\Bigg\}, \tag{63}$$

with $\vartheta = \frac{d}{p}\left(1-\frac{1}{\delta}\right)$.

Proof. Consider a dyadic decomposition of Ω given by the sets

$$V^j = \left\{ x \in \Omega : 2^{-j} \leq |x| \leq 2^{-j+1} \right\},\ j = 1, 2, \dots$$

and decompose the ball B_R into its intersections with the sets belonging to the decomposition, i.e., into $B^j = B_R \cap V^j$. Let us introduce the linear maps $\varphi_j : V^1 \to V^j$ and write with a hat the pullback of functions by φ_j^{-1}, e.g, $\hat{r} = r \circ \varphi_j^{-1}$ and $\widehat{B}^j = \varphi_j^{-1}(B^j)$. Then,

$$\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B^j)} \leq 2^{\frac{j}{\delta}(\gamma-2-d/p)}\|\hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^\beta \hat{u}\|_{L^{\delta p}(\widehat{B}^j)}$$

We can now use the interpolation inequality

$$\|v\|_{L^{\delta p}(B)} \leq C \|v\|^{1-\vartheta}_{L^p(B)} \|v\|^{\vartheta}_{W^{1,p}(B)},$$

for $B \subset \mathbb{R}^d$, $v \in W^{1,p}(B)$ and with ϑ defined as above, see [DFØSS12]. Therefore,

$$\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B^j)} \leq C 2^{\frac{j}{\delta}(\gamma-2-d/p)}\|\hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^\beta \hat{u}\|^{1-\vartheta}_{L^p(\widehat{B}^j)} \sum_{|\alpha|=1} \|\hat{\partial}^\alpha \hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^\beta \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)}. \tag{64}$$

Let us now consider the first norm in the product above. Since $\hat{r} \in (1/2, 1)$, we can inject in the norm a term $\hat{r}^{\gamma(1-1/\delta)} \leq \max(1, 2^{|\gamma|(1-1/\delta)}) = C(\gamma, \delta)$, i.e.,

$$\|\hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^\beta \hat{u}\|^{1-\vartheta}_{L^p(\widehat{B}^j)} \leq C \|\hat{r}^{|\beta|-\gamma}\hat{\partial}^\beta \hat{u}\|^{1-\vartheta}_{L^p(\widehat{B}^j)}.$$

We now compute more explicitly the second norm in the product in (64):

$$\sum_{|\alpha|=1} \|\hat{\partial}^\alpha \hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^\beta \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)} \leq \left(|\beta| + \frac{2-\gamma}{\delta}\right)^{\vartheta} \|\hat{r}^{\frac{2-\gamma}{\delta}+|\beta|-1}\hat{\partial}^\beta \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)} + \sum_{|\alpha|=1} \|\hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^{\beta+\alpha} \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)}$$

and we may adjust the exponents of $\hat{r}$ and the term in $\frac{2-\gamma}{\delta}$ introducing a constant that depends on γ, δ, d and p, obtaining

$$\sum_{|\alpha|=1} \|\hat{\partial}^\alpha \hat{r}^{\frac{2-\gamma}{\delta}+|\beta|}\hat{\partial}^\beta \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)} \leq C \left(|\beta|+1\right)^{\vartheta} \|\hat{r}^{|\beta|-\gamma}\hat{\partial}^\beta \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)} + \sum_{|\alpha|=1} \|\hat{r}^{|\beta|-\gamma+1}\hat{\partial}^{\beta+\alpha} \hat{u}\|^{\vartheta}_{L^p(\widehat{B}^j)}.$$

Scaling everything back to B^j and adjusting the exponents,

$$\begin{aligned}\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B^j)} \leq C 2^{\frac{j}{\delta}((\gamma-d/p)(1-\delta)-2)}\|r^{|\beta|-\gamma}\partial^\beta u\|^{1-\vartheta}_{L^p(B^j)} \Bigg\{ &\left(|\beta|+1\right)^{\vartheta} \|r^{|\beta|-\gamma}\partial^\beta u\|^{\vartheta}_{L^p(B^j)} \\ &+ \sum_{|\alpha|=1} \|r^{|\beta|-\gamma+1}\partial^{\beta+\alpha} u\|^{\vartheta}_{L^p(B^j)} \Bigg\}.\end{aligned}$$

If $\gamma - d/p \geq 2/(1-\delta)$ and therefore $2^{\frac{j}{\delta}((\gamma-d/p)(1-\delta)-2)} \leq 1$, we can sum over all $j = 1, 2, \dots$ thus obtaining the estimate (59) on the whole ball B_R. □

4.3. **Analyticity of solutions.** We now consider the nonlinear Schrödinger eigenvalue problem (4) with polynomial nonlinearity, given by

$$Lu = -\Delta u + Vu + |u|^{\delta-1}u = \lambda u. \tag{65}$$

We suppose that the potential V is singular on a finite set of discrete points and consider the case of an up-to-quartic nonlinearity (i.e., $\delta \in \mathbb{N}$ and $\delta \leq 4$). We show, in the following theorem, that the results on the regularity of the solution that can be obtained in the linear case [MM19] can be extended to the nonlinear one. We recall that $\varepsilon \in (0, 1)$ so that $V \in \mathcal{K}^{\varpi,\infty}_{\varepsilon-2}(\Omega)$ and we suppose, from now on, that $\delta \in \{2, 3, 4\}$.

Theorem 2. *Let u be the solution to (65), let $\varepsilon \in (0,1)$ such that $V \in \mathcal{K}^{\varpi,\infty}_{\varepsilon-2}(\Omega)$, and $\delta = 2,3,4$. Then,*

$$u \in \mathcal{J}^{\varpi}_{\gamma}(\Omega) \tag{66}$$

for any $\gamma < \varepsilon + d/2$.

In order to prove the analyticity in weighted spaces of the function u we need to bound the nonlinear term. We will introduce some preliminary lemmas and proceed by induction: let us specify the induction hypothesis. We suppose, here and in the sequel, that we have fixed a nonempty ball $B_R \subset \Omega$.

Induction Assumption. *For $k \in \mathbb{N}$, $1 < p < \infty$, $C_u, A_u > 0$, $\gamma \in \mathbb{R}$, we say that $H_u(k,p,\gamma,C_u,A_u)$ holds in B_R if for all $\rho \in (0, R/(2k)]$, $u \in \mathcal{J}^{k,p}_{\gamma}(B_R)$ and*

$$\sum_{|\alpha|=j} \|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-k\rho})} \le C_u A_u^j (k\rho)^{-j} j^j \quad \text{for all } \lfloor \gamma - d/p \rfloor + 1 \le j \le k. \tag{67}$$

Lemma 11. *Let $k \in \mathbb{N}$ and let $H_u(k,p,\gamma,C_u,A_u)$ hold for p,γ such that $2/(1-\delta) \le \gamma - d/p < \min(\varepsilon, 2)$ and $p \ge d(1-1/\delta)$. Then, there exists $C > 0$ dependent on C_u and independent of A_u, j, k, and ρ such that*

$$\sum_{|\beta|=j} \|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})} \le C A_u^{|\beta|+\vartheta}(k\rho)^{-|\beta|-\vartheta}|\beta|^{|\beta|}(|\beta|+1)^{\vartheta}$$

for $1 \le j \le k-1$ and with $\vartheta = \frac{d}{p}\left(1 - \frac{1}{\delta}\right)$.

Proof. First, we use (63) in order to go back to integrals in L^p:

$$\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})} \le C \|r^{|\beta|-\gamma}\partial^\beta u\|^{1-\vartheta}_{L^p(B_{R-k\rho})} \Bigg\{ (|\beta|+1)^{\vartheta} \|r^{|\beta|-\gamma}\partial^\beta u\|^{\vartheta}_{L^p(B_{R-k\rho})} + \sum_{i=1}^{d} \|r^{|\beta|+1-\gamma}\partial^\beta\partial_i u\|^{\vartheta}_{L^p(B_{R-k\rho})} \Bigg\}.$$

Then, hypothesis (67) implies, for $j = 1, \dots, k-1$,

$$\sum_{|\beta|=j} \|r^{|\beta|-\gamma}\partial^\beta u\|^{1-\vartheta}_{L^p(B_{R-k\rho})} \le C A_u^{j(1-\vartheta)} \rho^{-j(1-\vartheta)} \left(\frac{j}{k}\right)^{j(1-\vartheta)}$$

and

$$\sum_{|\beta|=j} (|\beta|+1)^{\vartheta} \|r^{|\beta|-\gamma}\partial^\beta u\|^{\vartheta}_{L^p(B_{R-k\rho})} + \sum_{i=1}^{d} \|r^{|\beta|+1-\gamma}\partial^\beta\partial_i u\|^{\vartheta}_{L^p(B_{R-k\rho})}$$
$$\le C(j+1)^{\vartheta} A_u^{j\vartheta}\rho^{-j\vartheta}\left(\frac{j}{k}\right)^{j\vartheta} + C A_u^{(j+1)\vartheta}\rho^{-(j+1)\vartheta}\left(\frac{j+1}{k}\right)^{(j+1)\vartheta}.$$

Therefore, multiplying the right hand sides of the last two inequalities,

$$\sum_{|\beta|=j} \|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})} \le C A_u^{j+\vartheta}(k\rho)^{-j-\vartheta} j^{j(1-\vartheta)}(j+1)^{(j+1)\vartheta}.$$

We finally need to bound the last two terms in the multiplication above. By Stirling's formula,

$$j^{j(1-\vartheta)}(j+1)^{(j+1)\vartheta} \le C j! j^{-1/2} e^j (j+1)^{\vartheta/2} j^{\vartheta/2}$$

and another application of Stirling's formula gives the thesis. □

In order to estimate the L^p weighted norms of derivatives of u^δ we will use Leibniz's rule and break the L^p norms into multiple $L^{\delta p}$ norms. Lemma 11 then allows to go back to the induction hypothesis. We continue by estimating the weighted norms of u^2 through the procedure we just outlined. For two multi indices $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ and $\beta = (\beta_1, \dots, \beta_d) \in \mathbb{N}^d$, we write $\alpha! = \alpha_1! \cdots \alpha_d!$, $\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_d + \beta_d)$, and

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha-\beta)!}.$$

Furthermore, recall [Kat96] that

$$\sum_{\substack{|\beta|=n\\ \beta\le\alpha}} \binom{\alpha}{\beta} = \binom{|\alpha|}{n}. \tag{68}$$

Note also that for $k \in \mathbb{N}$ and two quantities a, b depending on a multi index,

$$\sum_{|\beta|=k}\sum_{0<\zeta<\beta} a(\zeta)b(\beta-\zeta) = \sum_{j=1}^{k-1}\sum_{|\zeta|=j}\sum_{\substack{\beta:\beta>\zeta\\ |\beta|=k}} a(\zeta)b(\beta-\zeta) = \sum_{j=1}^{k-1}\sum_{|\zeta|=j}\sum_{|\xi|=k-j} a(\zeta)b(\xi). \tag{69}$$

Lemma 12. *Let $k \in \mathbb{N}$ and let $H_u(k, p, \gamma, C_u, A_u)$ hold for p, γ such that $2/(1-\delta) \le \gamma - d/p\varepsilon$ and $p \ge d(1 - 1/\delta)$. Then, for all $1 \le j \le k-1$ and for all $0 < \rho \le \frac{R}{2k}$,*

$$\sum_{|\alpha|=j} \|r^{2\frac{2-\gamma}{\delta}+|\alpha|}\partial^\alpha(u^2)\|_{L^{\delta p/2}(B_{R-k\rho})} \le CA_u^{j+2\vartheta}\rho^{-j-2\vartheta}\left(\frac{j}{k}\right)^j j^{1/2}. \tag{70}$$

Here, C depends on C_u and is independent of A_u, j, k, and ρ.

Proof. By Leibniz's rule and the Cauchy-Schwartz inequality,

$$\begin{aligned}
&\sum_{|\alpha|=j} \|r^{2\frac{2-\gamma}{\delta}+|\alpha|}\partial^\alpha(u^2)\|_{L^{\delta p/2}(B_{R-k\rho})} \\
&\quad\le \sum_{|\alpha|=j}\sum_{0<\beta<\alpha}\binom{\alpha}{\beta}\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})}\|r^{\frac{2-\gamma}{\delta}+|\alpha|-|\beta|}\partial^{\alpha-\beta}u\|_{L^{\delta p}(B_{R-k\rho})} \\
&\quad\quad + 2\sum_{|\alpha|=j}\|r^{2\frac{2-\gamma}{\delta}+|\alpha|}\partial^\alpha u\|_{L^{\delta p/2}(B_{R-k\rho})}\|u\|_{L^\infty(B_{R-k\rho})}
\end{aligned} \tag{71}$$

Considering the sum over $0 < \beta < \alpha$, Lemma 11, (69), and Stirling's inequality give

$$\begin{aligned}
&\sum_{|\alpha|=j}\sum_{0<\beta<\alpha}\binom{\alpha}{\beta}\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})}\|r^{\frac{2-\gamma}{\delta}+|\alpha|-|\beta|}\partial^{\alpha-\beta}u\|_{L^{\delta p}(B_{R-k\rho})} \\
&\quad\le CA_u^{j+2\vartheta}\rho^{-j-2\vartheta}\sum_{i=1}^{j-1}\binom{j}{i}i!(j-i)!e^j\,\frac{(i+1)^\vartheta(j-i+1)^\vartheta}{k^{j+2\vartheta}}\,\frac{1}{\sqrt{i(j-i)}} \\
&\quad\le CA_u^{j+2\vartheta}\rho^{-j-2\vartheta}\frac{j!e^j}{k^j} \\
&\quad\le CA_u^{j+2\vartheta}\rho^{-j-2\vartheta}\left(\frac{j}{k}\right)^j j^{1/2}.
\end{aligned}$$

The second term at the right hand side of (71) is controlled using Lemma 11, since $\gamma - d/p < 2$, and the injection $\mathcal{J}^2_{d/2+\eta}(\Omega) \hookrightarrow L^\infty(\Omega)$, valid for any $\eta > 0$ [KMR97]. □

With the same proof as above, we can deal with a cubic nonlinear term, as we show in the following lemma.

Lemma 13. *Under the same hypotheses as in Lemma 12, for all* $1 \leq j \leq k-1$ *and for all* $0 < \rho \leq \frac{R}{2k}$*,*

$$\sum_{|\alpha|=j} \|r^{3\frac{2-\gamma}{\delta}+|\alpha|}\partial^\alpha(u^3)\|_{L^{\delta p/3}(B_{R-k\rho})} \leq CA_u^{j+3\vartheta}\rho^{-j-3\vartheta}\left(\frac{j}{k}\right)^j j$$

C depends on C_u *and is independent of* A_u*,* j*,* k*, and* ρ*.*

Proof. We have

$$\begin{aligned}\sum_{|\alpha|=j} &\|r^{3\frac{2-\gamma}{\delta}+|\alpha|}\partial^\alpha(u^3)\|_{L^{\delta p/3}(B_{R-k\rho})} \\ &\leq C\sum_{|\alpha|=j}\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})}\|r^{2\frac{2-\gamma}{\delta}+|\alpha|-|\beta|}\partial^{\alpha-\beta}(u^2)\|_{L^{\delta p/2}(B_{R-k\rho})}.\end{aligned}$$

Using (70) we follow the same procedure as in the proof of Lemma 12. When $0 < \beta < \alpha$ in the sum above,

$$\begin{aligned}\sum_{|\alpha|=j}\sum_{0<\beta<\alpha}&\binom{\alpha}{\beta}\|r^{\frac{2-\gamma}{\delta}+|\beta|}\partial^\beta u\|_{L^{\delta p}(B_{R-k\rho})}\|r^{2\frac{2-\gamma}{\delta}+|\alpha|-|\beta|}\partial^{\alpha-\beta}(u^2)\|_{L^{\delta p/2}(B_{R-k\rho})} \\ &\leq CA_u^{j+3\vartheta}\rho^{-j-3\vartheta}\sum_{i=1}^{j-1}\binom{j}{i}i!(j-i)!e^j\frac{(i+1)^\vartheta(j-i+1)^\vartheta}{k^{j+2\vartheta}}\frac{\sqrt{j-i}}{\sqrt{i(j-i)}} \\ &\leq CA_u^{j+3\vartheta}\rho^{-j-3\vartheta}\frac{j!e^j\sqrt{j}}{k^j} \\ &\leq CA_u^{j+3\vartheta}\rho^{-j-3\vartheta}\left(\frac{j}{k}\right)^j j.\end{aligned}$$

As before, the terms in the sum where $\beta = 0$ and $\beta = \alpha$ give the same bound. ☐

The proof of the next lemma, in which we control a quartic term, amounts to a repetition of the arguments above; we show its proof for completeness.

Lemma 14. *Under the same hypotheses as in Lemma 12, for all* $1 \leq j \leq k-1$ *and for all* $0 < \rho \leq \frac{R}{2k}$*,*

$$\sum_{|\alpha|=j} \|r^{2-\gamma+|\alpha|}\partial^\alpha(u^4)\|_{L^p(B_{R-k\rho})} \leq CA^{j+4\vartheta}\rho^{-j-4\vartheta}\left(\frac{j}{k}\right)^j j^{3/2}.$$

C depends on C_u *and is independent of* A_u*,* j*,* k*, and* ρ*.*

Proof. There holds

$$\begin{aligned}\sum_{|\alpha|=j} &\|r^{2-\gamma+|\alpha|}\partial^\alpha(u^4)\|_{L^p(B_{R-k\rho})} \\ &\leq C\sum_{|\alpha|=j}\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}\|r^{\frac{2-\gamma}{4}+|\beta|}\partial^\beta u\|_{L^{4p}(B_{R-k\rho})}\|r^{3\frac{2-\gamma}{4}+|\alpha|-|\beta|}\partial^{\alpha-\beta}(u^3)\|_{L^{4p/3}(B_{R-k\rho})}.\end{aligned}$$

We can now use the result of Lemma 13 with $\delta = 4$. When $0 < \beta < \alpha$ in the sum above,

$$\sum_{|\alpha|=j} \sum_{0<\beta<\alpha} \binom{\alpha}{\beta} \|r^{\frac{2-\gamma}{4}+|\beta|}\partial^\beta u\|_{L^{4p}(B_{R-k\rho})} \|r^{3\frac{2-\gamma}{4}+|\alpha|-|\beta|}\partial^{\alpha-\beta}(u^3)\|_{L^{4p/3}(B_{R-k\rho})}$$
$$\leq C A_u^{j+4\vartheta}\rho^{-j-4\vartheta}\sum_{i=1}^{j-1}\binom{j}{i} i!(j-i)!e^j \frac{(i+1)^\vartheta(j-i+1)^\vartheta}{k^{j+2\vartheta}}\frac{\sqrt{j-i}}{\sqrt{i(j-i)}}$$
$$\leq C A_u^{j+4\vartheta}\rho^{-j-4\vartheta}\left(\frac{j}{k}\right)^j j.$$

The direct application of Lemmas 11 and 13 let us obtain the estimate for the terms in the sum where $\beta = \alpha$ and $\beta = 0$, respectively. ☐

The proof of (66) is now complete: we just need to bring the estimates together.

Proof of Theorem 2. The operator $L_{\text{lin}} = -\Delta + V$ is an isomorphism

$$\mathcal{J}^{k+2}_\xi(\Omega) \to \mathcal{J}^k_{\xi-2}(\Omega)$$

for any $0 < \xi - d/2 < \varepsilon$ and all $k \in \mathbb{N}$, see [MM19]. Since we can also show that $u \in L^\infty(\Omega)$ [Sta65], the solution to (65) is such that $u \in \mathcal{J}^2_\xi(\Omega)$. Iterating this line of reasoning, we can show that $u \in \mathcal{J}^3_\xi(\Omega)$ for all $0 < \xi - d/2 < \varepsilon$. We now claim that, by injection, for all $p > 1$ and all $0 < \gamma - d/p < \varepsilon$ there holds $u \in \mathcal{J}^{1,p}_\gamma(\Omega)$. Remark that, since $u \in \mathcal{J}^3_\xi(\Omega)$, we have [CDN10b] that $u - u(\mathfrak{c}) \in \mathcal{K}^3_\xi(\Omega)$. Then, by [MR10, Lemma 1.2.2], $u - u(\mathfrak{c}) \in \mathcal{K}^{1,p}_\gamma(\Omega)$ for all $p > 1$ and for $\gamma = \xi - d/2 + d/p$. Since furthermore $|u(\mathfrak{c})| \leq C\|u\|_{\mathcal{J}^2_\eta(\Omega)}$ for a C independent of u and for any $\eta > d/2$ [KMR97], we obtain that $u \in \mathcal{J}^{1,p}_\gamma(\Omega)$.

Therefore, for all $p > 1$ and all respective $0 < \gamma - d/p < \varepsilon$, the induction assumption $H_u(1, p, \gamma, C_u, A_u)$ is verified for some constants $C_u, A_u > 0$. We proceed by induction and impose a restriction on p; specifically,

$$p \geq \min\left(2, 2d\frac{\delta - 1}{5-\delta}\right). \tag{72}$$

The role of this condition on p will be clearer in the sequel. Let us now fix $\gamma \in (d/p, d/p + \varepsilon)$, suppose without loss of generality that $C_u A_u \geq \|u\|_{\mathcal{J}^{1,p}_\gamma(\Omega)}$ and that

$$C_u \geq C_{\text{reg}}(4 + |\lambda|R^2), \tag{73}$$

where C_{reg} is the constant in (59).

Suppose now that $H_u(k, p, \gamma, C_u, A_u)$ holds for $k \in \mathbb{N}$, with p subject to (72), C_u such that (73) holds, and a constant A_u independent of k whose lower bound will be specified throughout the proof. We will show that $H_u(k+1, p, \gamma, C_u, A_u)$ holds. Start by considering that, for all $\rho \in (0, \frac{R}{2(k+1)}]$ there exists $\widetilde{\rho} = \frac{k+1}{k}\rho$ such that $\widetilde{\rho} \in (0, \frac{R}{2k}]$ and

$$\sum_{|\alpha|=j}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-(k+1)\rho})} = \sum_{|\alpha|=j}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-k\widetilde{\rho}})} \qquad j = 1, \dots, k.$$

Hence, since $H_u(k, p, \gamma, C_u, A_u)$ holds, we have that

$$\sum_{|\alpha|=j}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-(k+1)\rho})} \leq C_u A_u^j (k\widetilde{\rho})^{-j} j^j \leq C_u A_u^j ((k+1)\rho)^{-j} j^j \quad \text{for all } 1 \leq j \leq k,$$

for all $0 < \rho \leq \frac{R}{2(k+1)}$. It remains to prove that, for all $0 < \rho \leq \frac{R}{2(k+1)}$,

$$\sum_{|\alpha|=k+1}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-(k+1)\rho})} \leq C_u A_u^{k+1}((k+1)\rho)^{-(k+1)}(k+1)^{k+1} \tag{74}$$

holds for $|\alpha| = k+1$ under condition (72). From (59) and (65), there exists a constant C_{reg} independent of u, k, and ρ such that

(75)
$$\sum_{|\alpha|=k+1} \|r^{k+1-\gamma}\partial^\alpha u\|_{L^p(B_{R-(k+1)\rho})} \leq C_{\mathrm{reg}} \left(\sum_{|\beta|=k-1} \|r^{k+1-\gamma}\partial^\beta(Vu + |u|^{\delta-1}u + \lambda u)\|_{L^p(B_{R-k\rho})} + \sum_{|\alpha|=k-1,k} \rho^{|\alpha|-k-1}\|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(B_{R-|\alpha|\rho})} \right).$$

We consider the term containing the potential V:

$$(76)\quad \sum_{|\beta|=k-1} \|r^{k+1-\gamma}\partial^\beta(Vu)\|_{L^p(B_{R-k\rho})}$$
$$\leq \sum_{|\beta|=k-1}\sum_{0<\zeta<\beta} \binom{\beta}{\zeta} \|r^{2-\varepsilon+|\zeta|}\partial^\zeta V\|_{L^\infty(B_{R-k\rho})} \|r^{\varepsilon-\gamma+|\beta|-|\zeta|}\partial^{\beta-\zeta}u\|_{L^p(B_{R-k\rho})}$$
$$+ \|r^{2-\varepsilon}V\|_{L^\infty(B_{R-k\rho})} \sum_{|\beta|=k-1} \|r^{\varepsilon-\gamma+|\beta|}\partial^\beta u\|_{L^p(B_{R-k\rho})}$$
$$+ \sum_{|\beta|=k-1} \|r^{2-\varepsilon+|\beta|}\partial^\beta V\|_{L^\infty(B_{R-k\rho})} \|r^{\varepsilon-\gamma}u\|_{L^p(B_{R-k\rho})}$$

Since $V \in \mathcal{K}^{\varpi,\infty}_{\varepsilon-2}(\Omega)$, we denote by C_V, A_V the constants such that $\max_{|\alpha|=i} \|r^{|\alpha|-\varepsilon+2}\partial^\alpha V\|_{L^\infty(\Omega)} \leq C_V A_V^i i!$ for all $i \in \mathbb{N}$ and suppose $A_u \geq A_V$. Using (68) and (69), we have, for a constant C independent of A_u

$$(77)\quad \sum_{|\beta|=k-1}\sum_{0<\zeta<\beta} \binom{\beta}{\zeta} \|r^{2-\varepsilon+|\zeta|}\partial^\zeta V\|_{L^\infty(B_{R-k\rho})} \|r^{\varepsilon-\gamma+|\beta|-|\zeta|}\partial^{\beta-\zeta}u\|_{L^p(B_{R-k\rho})}$$
$$\leq C \sum_{j=1}^{k-2} \binom{k-1}{j} A_u^{k-1} j! (k\rho)^{-(k-j-1)} (k-j-1)^{k-j-1}$$
$$\leq C A_u^{k-1}(k\rho)^{-(k-1)}(k-1)! e^{k-1} \sum_{j=1}^{k-2} \frac{(k\rho/e)^j}{\sqrt{k-1-j}}$$
$$\leq A_u^k (k\rho)^{-(k-1)}(k-1)^{k-1}$$

where we have concluded supposing, without loss of generality, that $A_u \geq C$ and $k\rho/e \leq 1$. The bound on the second to last term in (76) is straightforward, while for the last term we note that $\varepsilon - \gamma > -d/p$ thus $\|r^{\varepsilon-\gamma}u\|_{L^p(\Omega)} \leq C$.

We now consider the nonlinear term: in the lemmas above we have shown that, for $\delta = 2,3,4$ (recall that $\vartheta = \frac{d}{p}\left(1-\frac{1}{\delta}\right)$), for C independent of A_u,

$$\sum_{|\beta|=k-1} \|r^{2-\gamma+|\beta|}\partial^\beta u^\delta\|_{L^p(B_{R-k\rho})} \leq C A_u^{k-1+\delta\vartheta} \rho^{-(k-1)-\delta\vartheta} \left(\frac{k-1}{k}\right)^{k-1} (k-1)^{(\delta-1)/2}.$$

In addition, $|\beta| \leq C\rho^{-1}$, therefore, for C independent of A_u,

$$\sum_{|\beta|=k-1} \|r^{2-\gamma+|\beta|}\partial^\beta u^\delta\|_{L^p(B_{R-k\rho})} \leq C A_u^{k-1+\delta\vartheta} \rho^{-(k-1)-\delta\vartheta-(\delta-1)/2} \left(\frac{k-1}{k}\right)^{k-1}.$$

If (72) holds, then $\delta\vartheta + (\delta-1)/2 \leq 2$, hence, since also $\left(\frac{k-1}{k}\right)^{k-1} \leq e$,

$$\sum_{|\beta|=k-1} \|r^{2-\gamma+|\beta|}\partial^\beta u^\delta\|_{L^p(B_{R-k\rho})} \leq C A_u^{k-1+\delta\vartheta}\rho^{-(k-1)-\delta\vartheta-(\delta-1)/2} \leq A_u^{k+1}((k+1)\rho)^{-(k+1)}(k+1)^{k+1}, \tag{78}$$

where we have supposed that $A_u^{2-\delta\vartheta} \geq C$ (which is possible since $2-\delta\vartheta > 0$ and C is independent of A_u). Note that for all d and δ considered, (72) is stronger than the hypothesis $p \geq d(1-1/\delta)$ of Lemma 10. The bound on the term in λu and on the second sum of the right hand side of (75) can be obtained straightforwardly from the induction hypothesis. Hence, from (75), (77), and (78),

$$\begin{aligned}\sum_{|\alpha|=k+1} &\|r^{k+1-\gamma}\partial^\alpha u\|_{L^p(B_{R-(k+1)\rho})}\\ &\leq C_{\mathrm{reg}}\Big\{A_u^k(k\rho)^{-(k-1)} + A_u^{k+1}((k+1)\rho)^{-(k+1)}(k+1)^{k+1} + C_u A_u^k (k\rho)^{-(k+1)}k^{k+1}\\ &\quad + C_u A_u^{k-1}((k-1)\rho)^{-(k-1)}(k-1)^{k-1} + |\lambda|R^2 C_u A_u^{k-1}(k-1)^{k-1}(k\rho)^{-(k-1)}\Big\}\\ &\leq C_{\mathrm{reg}}(4+|\lambda|R^2)A_u^{k+1}((k+1)\rho)^{-k-1}(k+1)^{k+1}\end{aligned}$$

where C_{reg} is independent of k, ρ, C_u, and A_u, and we have supposed $A_u \geq C_u$ to obtain the last line. Since $C_u \geq C_{\mathrm{reg}}(4+|\lambda|R^2)$,we have proven the induction step, i.e., that for any fixed p such that (72) holds, and any fixed $\gamma - d/p \in (0,\varepsilon)$ there exist $C_u, A_u > 0$, such that for all $k \in \mathbb{N}$,

$$H_u(k,p,\gamma,C_u,A_u) \implies H_u(k+1,p,\gamma,C_u,A_u).$$

Therefore, (74) holds for all $k \in \mathbb{N}$; furthermore, since $R - k\rho \geq R/2$, we can find a covering of Ω that gives, for a constant $\widetilde{C}_u > 0$,

$$\sum_{|\alpha|=k} \|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(\Omega)} \leq \widetilde{C}_u A_u^k k^k,$$

for all $|\alpha| \geq 1$. Thanks to Stirling's formula, this is equivalent to (increasing the constant A_u to $\widetilde{A}_u$ in order to absorb the exponential and square root terms)

$$\sum_{|\alpha|=k} \|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(\Omega)} \leq \widetilde{C}_u \widetilde{A}_u^k k!. \tag{79}$$

Then, let q be such that $1/q = 1/2 - 1/p$. For any $\tilde\gamma < \gamma + d/q = \gamma - d/p + d/2$, $\|r^{\gamma-\tilde\gamma}\|_{L^q(\Omega)}$ is bounded, and

$$\sum_{|\alpha|=k} \|r^{|\alpha|-\gamma}\partial^\alpha u\|_{L^p(\Omega)} \leq \|r^{\gamma-\tilde\gamma}\|_{L^q(\Omega)} \sum_{|\alpha|=k} \|r^{|\alpha|-\tilde\gamma}\partial^\alpha u\|_{L^2(\Omega)}. \tag{80}$$

From (79) and (80) we infer (66). □

5. Exponential convergence for polynomial nonlinearities

In this section, we make the same hypotheses on F as in Section 4, i.e., we consider the concrete case where $f(u^2)$ is a polynomial. Let then

$$f(u^2) = |u|^{\delta-1} \tag{81}$$

for $\delta = 2, 3, 4$ (the case $\delta = 1$ is the linear one). Remark that this class of functions satisfies (6a) to (6d), with in particular $r = 2$ in (6d).

We also remark that since in this instance $f'(u^2)u^2 = \delta/2 f(u^2)$, the non scalar coefficients in (49) are the same that we find in (4). Hence, using elliptic regularity in weighted Sobolev spaces and the proof of Theorem 2, we obtain

$$\psi \in \mathcal{J}^{\varpi}_{\gamma}(\Omega)$$

for any $\gamma < d/2 + \varepsilon$, and in particular for all $s \geq 2$, $\|\psi\|_{\mathcal{J}^s_\gamma(\Omega)} \leq C\|u\|_{\mathcal{J}^s_\gamma(\Omega)}$. Therefore, thanks to (16),

$$\inf_{v_\delta \in X_\delta} |||\psi - v_\delta|||_{\mathrm{DG}} \leq C e^{-b\ell}. \tag{82}$$

We can regroup the results of the previous sections, applied to the case where (81) holds, in the following theorem.

Theorem 3. *Let u, λ be the solution to* (4) *and u_δ, λ_δ be the solution to* (11). *Suppose that* (7a), (7b), *and* (81) *hold. Then, for a space X_δ with N degrees of freedom, there exists $b > 0$ such that*

$$\|u - u_\delta\|_{\mathrm{DG}} \leq C e^{-bN^{1/(d+1)}} \tag{83}$$

and

$$|\lambda - \lambda_\delta| \leq C e^{-bN^{1/(d+1)}}. \tag{84}$$

Furthermore, if (37) *holds, then,*

$$|\lambda - \lambda_\delta| \leq C e^{-2bN^{1/(d+1)}}. \tag{85}$$

Proof. Bounds (83) and (84) are a consequence of the weighted analytic regularity of u given by Theorem 2, of the exponential approximation properties of X_δ stated in (16), and of Theorem 1. Equation (85) follows from Proposition 7 and (82). ☐

6. Numerical results

In this section, we show some results obtained in the approximation of the problem that, in its continuous form, reads: find the eigenpair $(\lambda, u) \in \mathbb{R} \times H^1(\Omega)$ such that $\|u\|_{L^2(\Omega)} = 1$ and

$$\begin{aligned} -\Delta u + Vu + |u|^2 u &= \lambda u \text{ in } \Omega \\ u &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{86}$$

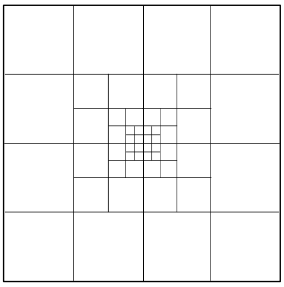
In particular, we focus on the computation of the *lowest eigenvalue* and of its associated eigenvector, corresponding, from a physical point of view, to the ground state of the system. The domain is given by the d-dimensional cube of unitary edge $(-1/2, 1/2)^d$.

Remark 7. *We consider different boundary conditions with respect to the setting of Sections 4 and 5, i.e., we consider here homogeneous Dirichlet boundary conditions instead of periodic ones. The theoretical analysis of this case is more complex, due to the fact that the ground state is not bounded from below, but the behavior of the method with these boundary condition is of computational interest, since homogeneous Dirichlet boundary conditions can be used to approximate physical systems in the whole space $\mathbb{R}^d$. Our numerical results indicate that the theoretical behavior shown for the periodic case extends to that of homogeneous Dirichlet boundary conditions.*

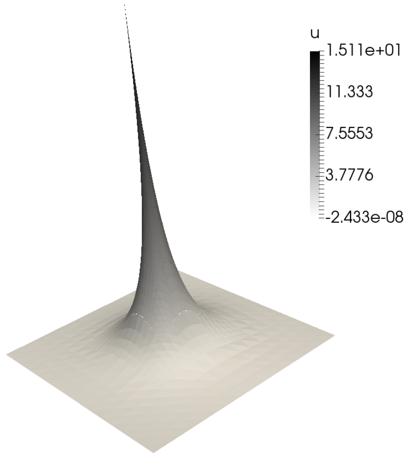
We take potentials of the form $V(x) = -r^{-\alpha}$, for $\alpha = -1/2, -1, -3/2$. We use a SIP method, and solve the nonlinearity by fixed point iterations. The stopping criterion on the nonlinear iterations is residual based, i.e., we stop iterating when

$$\langle (A^{u_\delta} - \lambda_\delta) u_\delta, u_\delta \rangle \leq \varepsilon_{\mathrm{tol}}$$

for a given computed solution $u_\delta \in X_\delta$ and a given tolerance $\varepsilon_{\mathrm{tol}}$. We will indicate the tolerance we use, on a case by case basis, in the following sections.



(A)



(B)

FIGURE 1. Left: mesh for the two dimensional approximation at a fixed refinement step. Right: Numerical solution to (86) with $V(x) = -r^{-3/2}$.

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

$\mathfrak{s}$	b_{L^2}	b_{DG}	b_{L^∞}	b_λ
0.125	0.72	0.73	0.74	1.24
0.25	0.92	0.94	0.94	1.5
0.5	1.06	1	0.98	1.25

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

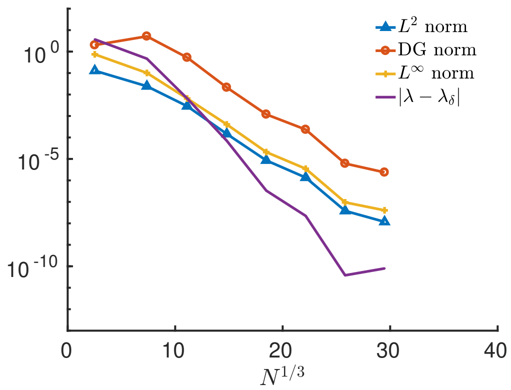
$\mathfrak{s}$	b_{L^2}	b_{DG}	b_{L^∞}	b_λ
0.125	0.59	0.6	0.58	1.05
0.25	0.72	0.72	0.7	1.01
0.5	0.68	0.68	0.58	0.65

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

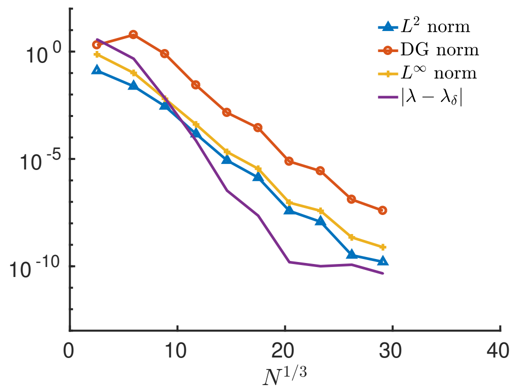
$\mathfrak{s}$	b_{L^2}	b_{DG}	b_{L^∞}	b_λ
0.062	0.43	0.45	0.5	0.8
0.125	0.56	0.52	0.65	0.76
0.25	0.48	0.47	0.43	0.47

6.1. **Two dimensional case.** In the two dimensional case, we compute the numerical solutions on meshes built with refinement ratio $\sigma = 1/2$, see Figure 1a. A visualization of the solution (in the most singular problem we analyse) is given in Figure 1b.

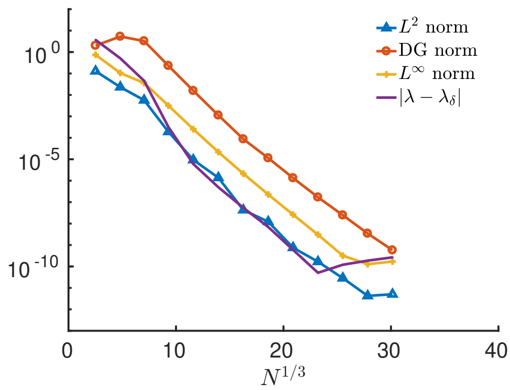
Writing $V(x) = -r^{-\alpha}$, we plot the curves of the errors in Figures 2 ($\alpha = 1/2$), 3 ($\alpha = 1$), and 4 ($\alpha = 3/2$). In the case of the approximations with low polynomial slopes, all errors converge



(A)



(B)



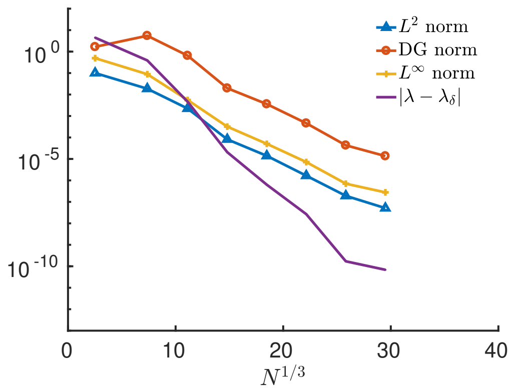
(C)

FIGURE 2. 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.

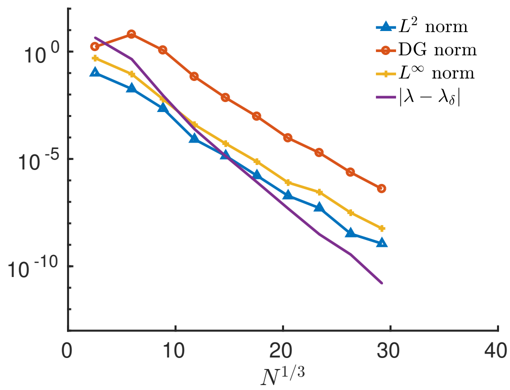
exponentially in the number of refinement steps, with the eigenvalue error converging faster than the norms of the eigenfunction error. A plateau due to the algebraic error is evident around 10^{-10}. When the polynomial slopes are higher, the quadrature error — not analyzed here, see [CCM10] for the analysis for h-type FE — manifests itself more strongly and causes, in extreme cases, the total loss of the doubling of the convergence rate.

The coefficients b_X, for $X = L^2(\Omega)$, DG, $L^\infty(\Omega)$ and λ are shown in Tables 1 to 3. As already discussed in [MM19], the higher the slope, the biggest the quadrature error and the furthest the estimated coefficients b_λ is from the double of the one for the DG norm.

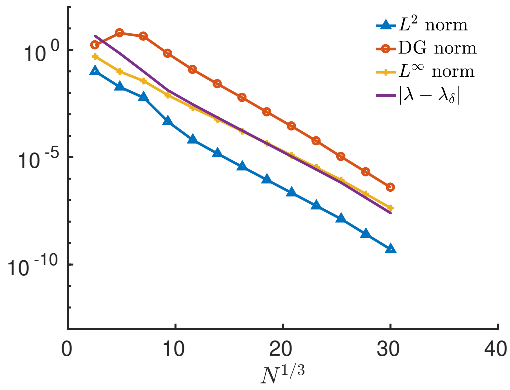
6.2. **Three dimensional problem.** In the three dimensional setting, we consider the domain $(-1/2, 1/2)^3$, and a mesh exemplified with refinement ratio $\sigma = 1/2$. The numerical solution of the problem with $V(x) = r^{-3/2}$ is shown in Figure 5. The solution shown is obtained at one of the highest degrees of refinement. The algebraic eigenproblem solver uses the Jacobi-Davidson



(A)



(B)



(C)

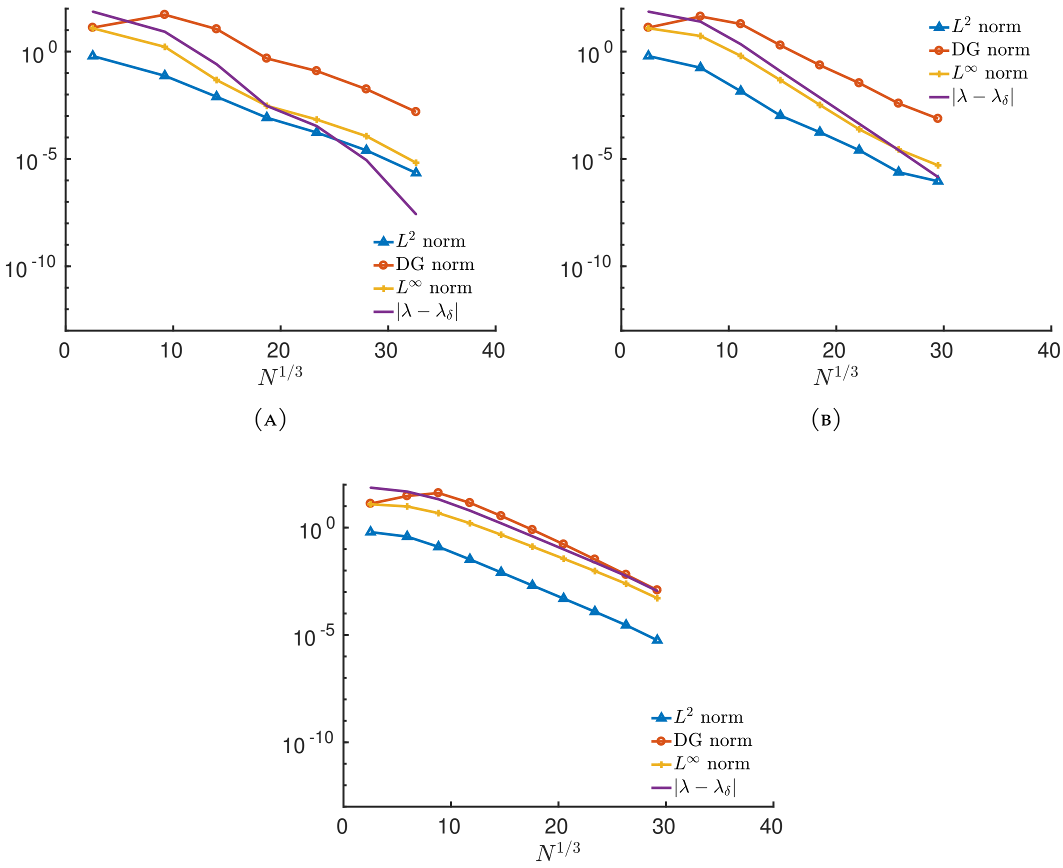
FIGURE 3. Errors for the numerical solution with potential $V(x) = -r^{-1}$. Polynomial slope: $\mathfrak{s} = 1/8$ in Figure a; $\mathfrak{s} = 1/4$ in Figure b and $\mathfrak{s} = 1/2$ in Figure c.

TABLE 4. Estimated coefficients. Potential: $r^{-1/2}$.

$\mathfrak{s}$	b_{L^2}	b_{DG}	b_{L^∞}	b_λ
0.125	0.73	0.74	0.81	1.28
0.25	0.82	0.82	0.85	1.3

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

$\mathfrak{s}$	b_{L^2}	b_{DG}	b_{L^∞}	b_λ
0.125	0.82	0.8	0.83	1.39
0.25	0.82	0.81	0.86	1.44



(A) (B) (C)

FIGURE 4. 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 6. Estimated coefficients. Potential: $r^{-3/2}$.

$\mathfrak{s}$	b_{L^2}	b_{DG}	b_{L^∞}	b_λ
0.125	0.69	0.67	0.71	1.29
0.25	0.8	0.73	0.52	1.3

method [SvdV96], with a biconjugate gradient method [vdV92, SvdVF94] as the linear algebraic system solver. The fixed point nonlinear iteration are set to a tolerance $\varepsilon_{\mathrm{tol}} = 10^{-7}$.

The algebraic and quadrature errors are not as evident as in the two dimensional case, and it can clearly be seen that an optimal slope can be chose to better approximate the eigenvalue. The nonlinearity does not seem to influence the rate of convergence; this is expected, since the source of the loss of regularity — the factor that most influences the rate of convergence — is primarily due to the potential.

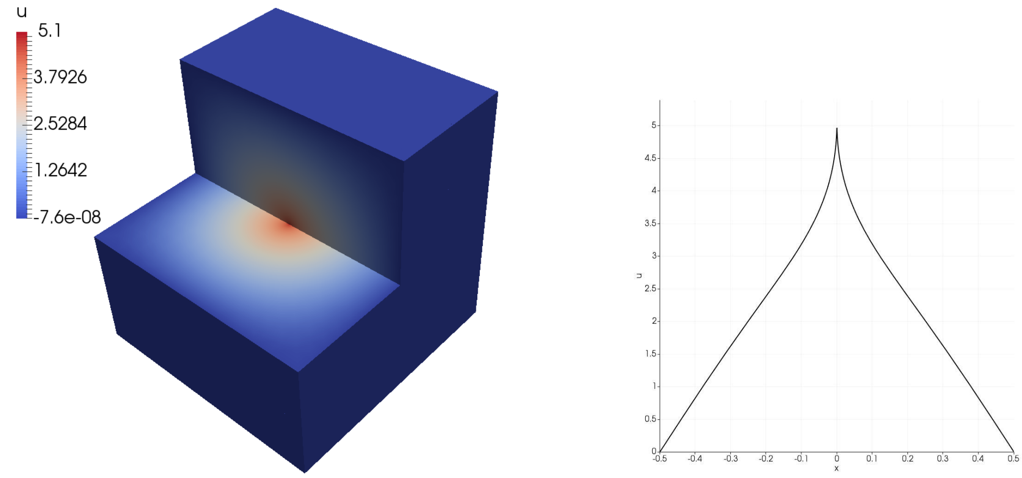


FIGURE 5. Numerical solution in the three dimensional case: solution in the cube, left, and close up near the origin of the restriction to the line $\{y = z = 0\}$, right

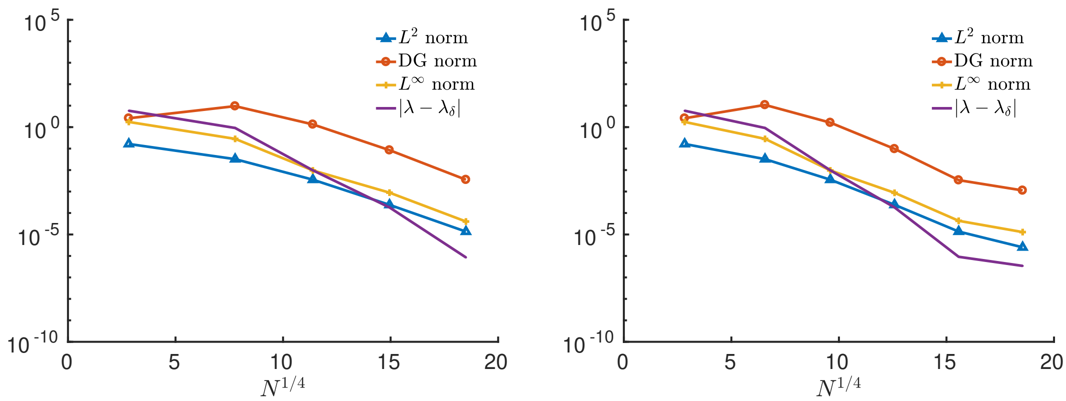


FIGURE 6. Errors of the numerical solution for $V(x) = r^{-1/2}$. Polynomial slope $\mathfrak{s} = 1/8$, left and $\mathfrak{s} = 1/4$, right.

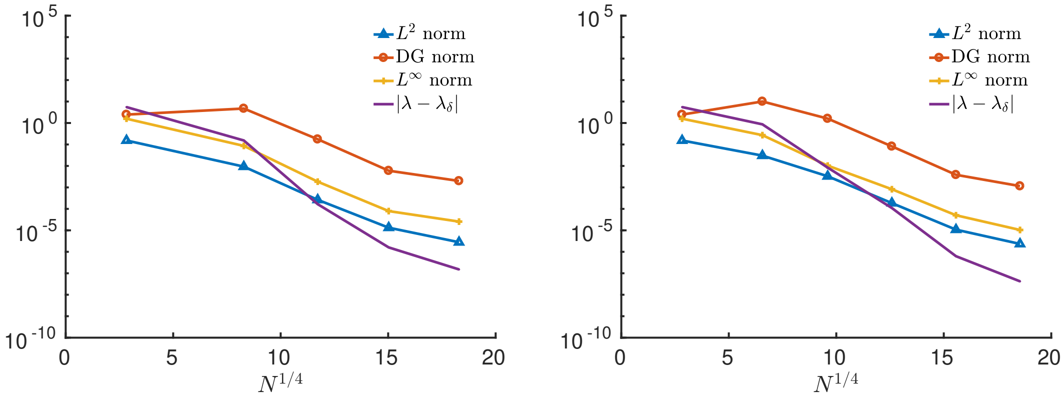


FIGURE 7. Errors of the numerical solution for $V(x) = r^{-1}$. Polynomial slope $\mathfrak{s} = 1/8$, left and $\mathfrak{s} = 1/4$, right.

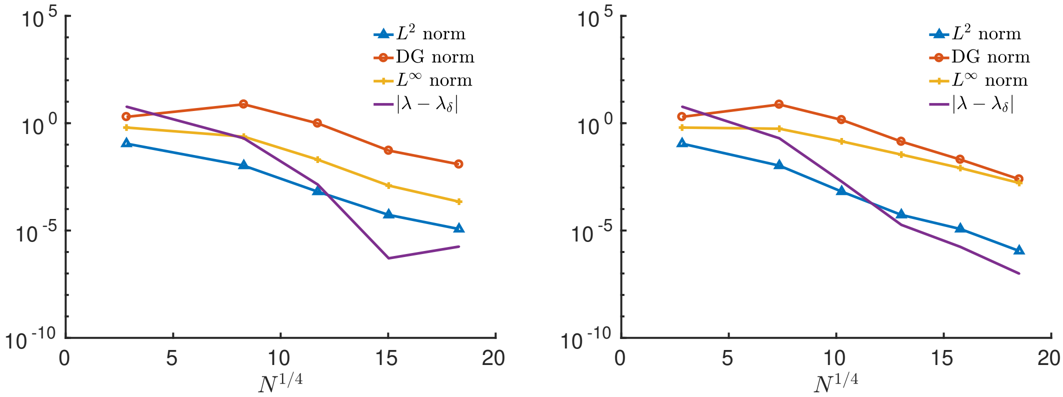


FIGURE 8. Errors of the numerical solution for $V(x) = r^{-3/2}$. Polynomial slope $\mathfrak{s} = 1/8$, left and $\mathfrak{s} = 1/4$, right.

References

[ABP06] Paola F. Antonietti, Annalisa Buffa, and Ilaria Perugia, *Discontinuous Galerkin approximation of the Laplace eigenproblem*, Comput. Methods Appl. Mech. Engrg. **195** (2006), no. 25-28, 3483–3503.

[BO91] Ivo Babuška and John Osborn, *Eigenvalue problems*, Handbook of numerical analysis, Vol. II, Handb. Numer. Anal., II, North-Holland, Amsterdam, 1991, pp. 641–787.

[CC04] Zhangxin Chen and Hongsen Chen, *Pointwise error estimates of discontinuous Galerkin methods with penalty for second-order elliptic problems*, SIAM J. Numer. Anal. **42** (2004), no. 3, 1146–1166.

[CCM10] Eric Cancès, Rachida Chakir, and Yvon Maday, *Numerical analysis of nonlinear eigenvalue problems*, J. Sci. Comput. **45** (2010), no. 1-3, 90–117.

[CDN10a] Martin Costabel, Monique Dauge, and Serge Nicaise, *Corner singularities and analytic regularity for linear elliptic systems*, 2010, Book in preparation.

[CDN10b] ———, *Mellin analysis of weighted Sobolev spaces with nonhomogeneous norms on cones*, Around the research of Vladimir Maz'ya. I, Int. Math. Ser. (N. Y.), vol. 11, Springer, New York, 2010, pp. 105–136.

[CDN12] ———, *Analytic regularity for linear elliptic systems in polygons and polyhedra*, Math. Models Methods Appl. Sci. **22** (2012), no. 8, 1250015, 63.

[DFØSS12] Anna Dall'Acqua, Søren Fournais, Thomas Østergaard Sørensen, and Edgardo Stockmeyer, *Real analyticity away from the nucleus of pseudorelativistic Hartree-Fock orbitals*, Anal. PDE **5** (2012), no. 3, 657–691.

[DPE12] Daniele A. Di Pietro and Alexandre Ern, *Mathematical aspects of discontinuous Galerkin methods*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 69, Springer, Heidelberg, 2012.

[GB86a] Benqi Guo and Ivo Babuška, *The h-p version of the finite element method - Part 1: The basic approximation results*, Computational Mechanics **1** (1986), no. 1, 21–41.

[GB86b] ———, *The h-p version of the finite element method - Part 2: General results and applications*, Computational Mechanics **1** (1986), no. 3, 203–220.

[GS06] Benqi Guo and Christoph Schwab, *Analytic regularity of Stokes flow on polygonal domains in countably weighted Sobolev spaces*, J. Comput. Appl. Math. **190** (2006), no. 1-2, 487–519.

[Kat96] Keiichi Kato, *New idea for proof of analyticity of solutions to analytic nonlinear elliptic equations*, SUT J. Math. **32** (1996), no. 2, 157–161.

[KMR97] Vladimir A. Kozlov, Vladimir Maz'ya, and Jürgen Rossmann, *Elliptic boundary value problems in domains with point singularities*, Mathematical Surveys and Monographs, vol. 52, American Mathematical Society, Providence, RI, 1997.

[Kon67] Vladimir A. Kondrat'ev, *Boundary value problems for elliptic equations in domains with conical or angular points*, Trudy Moskov. Mat. Obšč. **16** (1967), 209–292.

[LS03] Andris Lasis and Endre Suli, *Poincaré-type inequalities for broken Sobolev spaces*, Tech. Report 03/10, Oxford University Computing Laboratory, 2003.

[MM19] Yvon Maday and Carlo Marcati, *Regularity and hp discontinuous Galerkin finite element approximation of linear elliptic eigenvalue problems with singular potentials*, Math. Models Methods Appl. Sci. **29** (2019), no. 8, 1585–1617.

[MR10] Vladimir Maz'ya and Jürgen Rossmann, *Elliptic equations in polyhedral domains*, Mathematical Surveys and Monographs, vol. 162, American Mathematical Society, Providence, RI, 2010.

[MS19] Carlo Marcati and Christoph Schwab, *Analytic regularity for the Navier-Stokes equations in polygons*, Tech. Report 2019-12, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2019.

[Nic97] Serge Nicaise, *Regularity of the solutions of elliptic systems in polyhedral domains*, Bull. Belg. Math. Soc. Simon Stevin **4** (1997), no. 3, 411–429.

[Saa11] Yousef Saad, *Numerical methods for large eigenvalue problems*, Classics in Applied Mathematics, vol. 66, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[SSW13] Dominik Schötzau, Christoph Schwab, and Thomas P. Wihler, *hp-DGFEM for second order elliptic problems in polyhedra II: Exponential convergence*, SIAM J. Numer. Anal. **51** (2013), no. 4, 2005–2035.

[Sta65] Guido Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble) **15** (1965), no. 1, 189–258.

[SvdV96] Gerard L. G. Sleijpen and Henk A. van der Vorst, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 2, 401–425.

[SvdVF94] Gerard L. G. Sleijpen, Henk A. van der Vorst, and Diederik R. Fokkema, $\mathrm{BiCGstab}(l)$ *and other hybrid Bi-CG methods*, Numer. Algorithms **7** (1994), no. 1, 75–109.

[vdV92] Henk. A. van der Vorst, *Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput. **13** (1992), no. 2, 631–644.