



**HAL**  
open science

# Polygonal Faber-Krahn inequality: Local minimality via validated computing

Benjamin Bogosel, Dorin Bucur

► **To cite this version:**

Benjamin Bogosel, Dorin Bucur. Polygonal Faber-Krahn inequality: Local minimality via validated computing. 2024. hal-04613443

**HAL Id: hal-04613443**

**<https://hal.science/hal-04613443>**

Preprint submitted on 16 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# POLYGONAL FABER-KRAHN INEQUALITY: LOCAL MINIMALITY VIA VALIDATED COMPUTING.

BENIAMIN BOGOSEL, DORIN BUCUR

**ABSTRACT.** The main result of the paper shows that the regular  $n$ -gon is a local minimizer for the first Dirichlet-Laplace eigenvalue among  $n$ -gons having fixed area for  $n \in \{5, 6\}$ . The eigenvalue is seen as a function of the coordinates of the vertices in  $\mathbb{R}^{2n}$ . Relying on fine regularity results of the first eigenfunction in a convex polygon, an explicit *a priori* estimate is given for the eigenvalues of the Hessian matrix associated to the discrete problem, whose coefficients involve the solutions of some Poisson equations with singular right hand sides. The *a priori* estimates, in conjunction with certified finite element approximations of these singular PDEs imply the local minimality for  $n \in \{5, 6\}$ . All computations, including the finite element computations, are realized using interval arithmetic.

## 1. INTRODUCTION

In the previous work [7] it is shown that a local minimality result for a classical discrete shape optimization problem due to Polyà and Szegő, still unsolved today, depends on the positivity of a number of constants. Theory alone is not able to answer this question, for now. Using explicit finite element estimates for solutions of some singular partial differential equations together with interval arithmetic in Intlab [42], controlling all roundoff errors in the computations, allows to establish the local minimality results.

More precisely, for a bounded open set  $\Omega \subset \mathbb{R}^2$  let us consider the eigenvalue problem for the Laplace operator with Dirichlet boundary conditions

$$(1) \quad \begin{cases} -\Delta u = \lambda u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

The spectrum consists only on eigenvalues, counted with corresponding multiplicity,

$$0 < \lambda_1(\Omega) \leq \lambda_2(\Omega) \leq \dots \leq \lambda_k(\Omega) \dots \rightarrow +\infty.$$

Lord Rayleigh conjectured in 1877 that the first eigenvalue is minimal on the disc, among all other planar domains of the same area. Faber and Krahn gave rigorous proofs in 1923 (see [12] for a description of the history of the problem and [22, 21] for a survey of the topic).

Drawing a parallel with the isoperimetric inequality, it is natural to assert that the polygonal shape with  $n$  sides and fixed area which minimizes the fundamental Dirichlet Laplace eigenvalue is the regular polygon. This is also in accord with the intuition that a "rounder" shape takes longer to cool down from any initial state, keeping in mind that  $\lambda_1(\Omega)$  is the dominant rate of the vanishing exponential involved.

In their book of 1951, Pólya and Szegő conjecture a polygonal version of this inequality (see [41, page 158]). Denoting by  $\mathcal{P}_n$  the family of simple polygons with  $n$  sides in  $\mathbb{R}^2$  and for every  $n \geq 3$  consider the problem

$$(2) \quad \min_{P \in \mathcal{P}_n} \lambda_1(P) |P|.$$

Formulation (2) is one of the multiple equivalent statements of this problem and has the advantage of being scale invariant:  $\lambda_1(tP) = \lambda_1(P)/t^2$  for any  $t > 0$ .

**Pólya-Szegő Conjecture (1951).** *The unique solution to problem (2) is the regular polygon with  $n$  sides.*

The conjecture holds true for  $n = 3$  and  $n = 4$ : see [21, Chapter 3] where the Steiner symmetrization principle is used. Steiner symmetrization does not work for the case  $n \geq 5$  since, the number of vertices could possibly increase after symmetrization. A new approach,

which applies only to triangles, was proposed by Fragalà and Velichkov in [15], establishing that equilateral triangles are the only critical points for the first eigenvalue. Indrei finds in [25] a manifold in which the regular polygon is optimal. Numerical simulations suggesting that the Polyà-Szegő conjecture is true are given in [14], [2], [6].

In [7] the authors show that the proof of optimality of the regular  $n$ -gon for (2) can be reduced to a finite number of validated numerical computations. This is achieved by the following results:

- computation of second order shape derivatives of simple Dirichlet-Laplacian eigenvalues on Lipschitz domains (including polygons)
- computation of the Hessian matrix of  $\lambda_1(P)$  in terms of the coordinates of the vertices of  $P$ . Taking into account the symmetry, the eigenvalues of the Hessian matrix for the regular  $n$ -gon are characterized using solutions of three PDEs.
- Explicit a priori estimates for piece-wise affine finite elements approximation for these three PDEs are given (eigenvalue equation and material derivatives for the eigenfunction with respect to one vertex). Numerical computations justify, *up to machine errors*, that for  $5 \leq n \leq 8$  the regular  $n$ -gon is a local minimizer in (2).
- Qualitative results are given concerning the stability of the Hessian eigenvalues under vertex perturbations and bounds on geometric quantities of the optimal polygon are given.

In [30] the local maximality of the Robin eigenvalue problem for the Laplacian is established for quadrilaterals, using similar lines and exploiting the fact that in the case of the square, the PDEs involved in the Hessian computation become more explicit.

In order to conclude with local minimality, the missing step in [7] was precisely the certification of the machine errors, certification which could not be performed because of the complexity and size of the required computations.

Proof strategies based on certified computations have already been used successfully in articles related to spectral theory. In the paper [38], Schiffer's conjecture is proved for the regular pentagon: a Neumann eigenfunction that is positive on the boundary and not identically constant is found numerically through certified computations. In [10] a counter example to a famous conjecture of Payne is given through validated numerics: an example of two dimensional domain with 6 holes is given for which the nodal line of the second Dirichlet eigenfunction is closed and does not touch the boundary of the domain. In [17] a conjecture stated in [2] is settled, proving that there exist triangles which are not isometric and for which the first, second and fourth eigenvalues coincide.

The numerical framework chosen for performing the certified computations in this article uses finite elements. Let us briefly justify this choice. There exist precise numerical methods that compute eigenvalues and eigenfunctions for the Laplace operator on smooth domains or on polygons. For example [11] uses particular solutions coupled with interval arithmetic to obtain guaranteed enclosures for the Laplace eigenvalues which is then used in [10] to answer a question of Payne regarding the nodal line of the first eigenfunction on a domain with holes. In [5] the authors give tight inclusions for the Neumann eigenvalues and in [18] a fast numerical method is used for solving Laplace problems. While the results in [18] are promising in singular contexts, there are no explicit quantifiable error bounds that can be applied directly for some of the problems of interest in this work. We also point out [3] and [4] for a numerical analysis of general Poisson equations with distributional data.

Several non-trivial differences, with respect to the case of the Laplace operator treated in the references above, occur in our case: the operator appearing in our PDEs is  $-\Delta - \lambda_1(\mathbb{P}_n)I$ , it acts on a strict subspace contained in  $H_0^1$  and the right-hand side is singular, so that solutions are not globally in  $H^2$ . Here is the first comment: the general approximation results for singular data available in the literature do apply, but the (optimal) approximation order is too small and, in view of the requirements of the *a priori* estimates, the linear systems are too large to be solved with interval arithmetic software like Intlab. A second comment concerns the use of spectral methods. Because of the singularity of the right hand side, the solutions we deal with have discontinuous normal derivatives across segments inside the regular polygon  $\mathbb{P}_n$ . Adapting

the spectral methods enumerated above to this case is not straightforward. Piecewise affine finite element methods, on the other hand, can capture singular behavior if the singularities themselves are captured exactly by the underlying mesh. One main contribution of the paper is to quantify precisely the corresponding error estimates in such singular contexts.

Numerical computations performed in [7] correspond to meshes having triangle sizes small enough such that the *a priori* error estimates guarantee that all the non-zero Hessian eigenvalues are positive. In order to turn a numerical computation in a mathematical proof, all computations need to be certified, taking into account even the errors coming from floating point computations. Many such proof techniques emerged recently, using specialized software like Intlab [42], which replaces floating point variables with intervals. Operations performed with intervals guarantee that the output interval associated to a particular mathematical operation contains the desired result. All errors are taken into account, including roundoff errors present in floating point arithmetic.

Certified computation of eigenvalues problems in the  $\mathbf{P}_1$  finite elements context is described in [33]. The same paper allows to handle the context corresponding to Homogeneous Dirichlet and Neumann boundary conditions. For non-homogeneous Neumann boundary conditions the paper [32] applies. In Section 3 an alternative formulation is proposed for a particular problem involving non-homogeneous Neumann boundary conditions, allowing the validation of transmission problems, in particular. This framework should extend to other situations where the boundary data implies the  $H^2$  regularity of solutions, following classical results in [20].

The results of this paper finalize the proof of the local minimality of the regular  $n$ -gon for problem (2) for  $n \in \{5, 6\}$  by performing all computations (including the finite element computations) using interval arithmetic in Intlab [42]. To achieve this goal, in order to make the computation effort accessible, significant improvements over the *a priori* error estimates given in [7] are done. They rely on the following theoretical issues:

- The regularity of the material derivatives of the first eigenfunction of the Dirichlet-Laplace operator on the regular  $n$ -gon (see the next section), is characterized (and improved) locally, on a suitable partition of the polygon. The piece-wise  $H^2$  regularity is established and explicit error estimates for the piecewise affine finite elements are given. Because of the localization procedure which isolates the singularities, the order of the estimates and the constants are significantly improved compared with [7].
- The new error estimate allows meshes which have significantly fewer elements to be considered for approximating the eigenvalues (1) and the material derivatives. The size of the linear generalized eigenvalue problem and linear systems involved becomes small enough such that Intlab [42] can be used to certify the corresponding computations.

**Structure of the paper.** Section 2 introduces preliminary results from [7] regarding shape derivatives, polygonal perturbations and the Hessian matrix of the first eigenvalue on the regular  $n$ -gon. Explicit *a priori* estimates regarding finite element approximations of the eigenvalues are recalled.

Section 3 deals with explicit estimates for solutions to PDEs involving the Laplace operator with a singular source term, namely a measure with density in  $H^{1/2}$  supported on a segment, which gives rise to transmission conditions. These PDEs appear in the expression of the eigenvalues of the Hessian matrix described above. If the singularities are meshed exactly, the optimal order of convergence  $O(h)$  can still be reached in the approximation process. The main tool is proving that the solution is piece-wise  $H^2$ , adapting results from [20, Chapter 4]. The method proposed can extend to other cases where solutions are piecewise  $H^2$  and the the mesh captures the singularities exactly.

Section 4 gives various results regarding error estimates related to linear algebra, which justify the choices made in the validation process. Section 5 presents the validation strategy and the corresponding results concerning the local minimality of the regular  $n$ -gon in the Polyà-Szegő conjecture for  $n \in \{5, 6\}$ . The validation strategy illustrates the potential sources of errors: error between continuous problems and exact finite element solutions, errors in the discrete eigenvalue problems and linear systems, roundoff errors appearing when working in floating point precision.

For the sake of completeness, developments related to particular aspects in the code, possibly relevant for other applications, are detailed in the Appendices. Appendix A describes a possible implementation for the assembly of Morley finite elements which helps computing tight certified upper bounds for the interpolation constant for  $\mathbf{P}_1$  finite elements, based on the method proposed in [29]. Appendix B describes the main functions in the code allowing to verify the local minimality. The corresponding code is provided at the repository:

<https://github.com/bbogo/PolyaSzego>

allowing to verify and reproduce the results proved in this paper. Appendix C details a slight modification in the function `verifyeig` from Intlab (related to certification of discrete eigenvalue problems), replacing a matrix inversion with validated linear systems. For sparse matrices of large size this improves significantly the performance of the process of certifying solutions to generalized eigenvalue problems.

## 2. PRELIMINARIES

**2.1. Dirichlet-Laplacian eigenvalues on polygons: first and second derivatives.** Shape derivatives are introduced to study the behavior of shape functionals with respect to boundary variations. Given  $\xi \in W^{1,\infty}(\mathbb{R}^2, \mathbb{R}^2)$  Lipschitz vector fields the shape derivative verifies

$$J((I + \xi)(\Omega)) = J(\Omega) + J'(\Omega)(\xi) + o(\|\xi\|_{W^{1,\infty}}),$$

where  $\xi \mapsto J'(\Omega)(\xi)$  is a linear form in  $\xi$ .

Shape derivatives for a simple Dirichlet-Laplace eigenvalue are well known in the literature when the domain  $\Omega$  is smooth ([23], [43], [13]). For a simple eigenvalue one has

$$(3) \quad \lambda'(\Omega)(\xi) = - \int_{\partial\Omega} (\partial_n u)^2 \xi \cdot \mathbf{n},$$

where  $\partial_n$  denotes the normal derivative,  $\mathbf{n}$  is the outer normal to the domain  $\Omega$  and  $u$  is an  $L^2$  normalized eigenfunction associated to  $\lambda$ . The previous formula extends to all cases where  $u \in H^2(\Omega)$ . It can be noted that in this case the shape derivative depends only on  $\xi \cdot \mathbf{n}$  on the boundary  $\partial\Omega$ . This structure is verified in most cases where the functional and the domain are smooth enough [23, Chapter 5].

In situations where the domain is less smooth, known results do not apply directly. In [7], based on ideas developed in [31], the following volume form of the shape derivative is derived

$$(4) \quad \lambda'(\Omega)(\xi) = \int_{\Omega} \mathbf{S}_1 : D\xi,$$

where

$$(5) \quad \mathbf{S}_1 = (|\nabla u|^2 - \lambda u^2) \mathbf{Id} - 2(\nabla u \otimes \nabla u).$$

The matrix scalar product is denoted by  $:$ ; the identity matrix by  $\mathbf{Id}$  and  $D\xi$  is the Jacobian of  $\xi$ . This formula applies whenever  $u \in H^1(\Omega)$ , thus including less regular contexts. By analyzing singularities of the eigenfunction  $u$  on polygons in [7] it is shown that (3) remains valid on polygons.

An advantage of the volumic formulation (4) is that it can be further differentiated without additional regularity assumptions. Thus, second order shape derivatives for simple Dirichlet-Laplace eigenvalues are computed in [7] for all domains where  $u \in H^1(\Omega)$  and for all Lipschitz perturbation fields. The polygonal framework is thus included. For brevity, the formula for the second shape derivative is omitted here, since all details can be found in [7]. In the following, the polygonal case which is of interest in this work is detailed.

**Polygons and polygonal perturbations.** Suppose  $\Omega$  is a  $n$ -gon and follow the strategy described in [7], [31]. Starting from a perturbation of the vertices, the perturbation field  $\xi \in W^{1,\infty}(\mathbb{R}^2)$  will be built as follows. Denote the vertices of the polygon by  $\mathbf{a}_i \in \mathbb{R}^2$ ,  $i = 0, \dots, n-1$  and for each vertex consider the vector perturbation  $\xi_i \in \mathbb{R}^2$ ,  $i = 0, \dots, n-1$ . Whenever necessary, we suppose that the indices are considered modulo  $n$ . Consider a triangulation  $\mathcal{T}$  of  $\Omega$  such that the edges of the polygon are complete edges of some triangles in this triangulation. Moreover, consider the following globally Lipschitz functions  $\varphi_i$  for  $0 \leq i \leq n-1$  that are piecewise affine

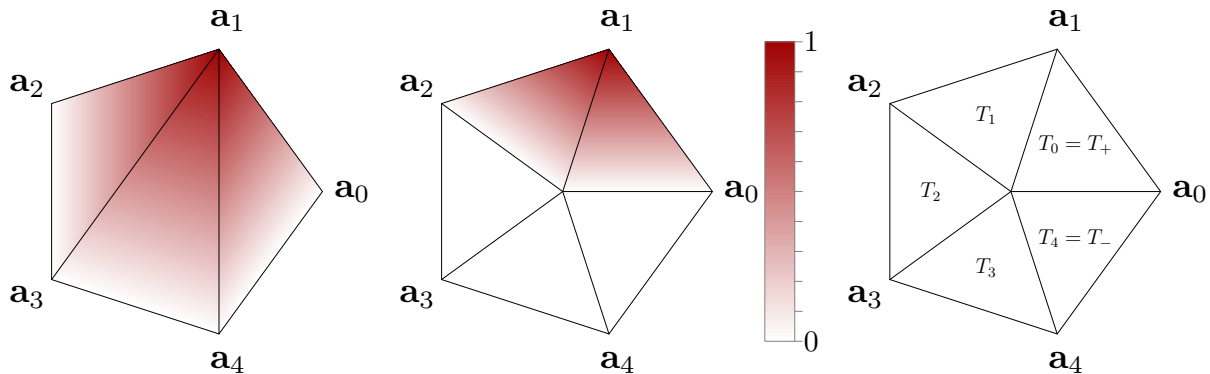


FIGURE 1. Examples of admissible triangulations used for defining perturbations on a polygon and graphical view of the function  $\varphi_1$ .

on each triangle of  $\mathcal{T}$  and satisfy

$$(6) \quad \varphi_i(\mathbf{a}_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Several choices are possible, as the two examples of Figure 1 show, their extension outside the polygon being irrelevant. Then, we build a global perturbation in  $\mathbb{R}^2$  given by

$$(7) \quad \xi = \sum_{i=0}^{n-1} \zeta_i \xi_i \in W^{1,\infty}(\mathbb{R}^2).$$

Plugging the polygonal perturbation field  $\xi$  given in (7) into the shape derivative (4) shows that the gradient of a simple eigenvalue with respect to the vertex coordinates  $\mathbf{x} = (\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_{n-1})$  is given by

$$\nabla \lambda(\mathbf{x}) = \left( \int_{\Omega} \mathbf{S}_1 \nabla \varphi_i \right)_{i=0}^{n-1},$$

where  $\varphi_i$  are the piecewise affine functions on the triangulation  $\mathcal{T}$  verifying (6) and  $\mathbf{S}_1$  is the matrix defined in (5). For simplicity,  $\varphi_0$  which will appear often in the following is simply denoted by  $\varphi$ .

The area of a polygon is explicit in terms of vertex coordinates and, assuming the vertices are ordered in the anti-clockwise direction, it is given by:

$$\mathcal{A}(\mathbf{x}) = \frac{1}{2} \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i).$$

Consider the case of the regular  $n$ -gon  $\mathbb{P}_n$ . In the following we use the notation  $\theta = 2\pi/n$ . Assume that  $\mathbb{P}_n$  is inscribed in the unit disk with  $\mathbf{a}_0 = (1, 0)$ . Consider the symmetric triangulation  $\mathcal{T}$  illustrated in Figure 1 with an extra vertex at the center of  $\mathbb{P}_n$ . Denote by  $T_i$  the triangles in the triangulation  $\mathcal{T}$  like in center image from Figure 1. The triangles  $T_0$  and  $T_{n-1}$  will sometimes be denoted by  $T_+, T_-$ , respectively and correspond to the support of the  $\mathbf{P}_1$  function  $\varphi$ .

The gradient of  $\lambda(\mathbf{x})$  with respect to perturbations in the first vertex  $\mathbf{a}_0$  of the regular polygon  $\mathbb{P}_n$  is

$$\int_{\mathbb{P}_n} \mathbf{S}_1 \nabla \varphi_0 = \begin{pmatrix} -2\lambda_1/n \\ 0 \end{pmatrix}.$$

Plugging polygonal perturbation fields into the second shape derivative formula obtained in [7] gives an explicit formula for the Hessian matrix  $D^2 \lambda_1(\mathbf{x})$  for  $\lambda_1(\mathbf{x})$  on the regular  $n$ -gon. In [7, Section 4] explicit formulas for the eigenvalues of the Hessian are given for the regular  $n$ -gon.

Denote by  $a(u, v)$  the bilinear form

$$a(u, v) = \int_{\mathbb{P}_n} \nabla u \cdot \nabla v - \lambda_1(\mathbb{P}_n) uv,$$

which is positive semi-definite on  $H_0^1(\mathbb{P}_n)$ . The only functions which cancel  $a(\cdot, \cdot)$  are multiples of the first eigenfunction  $u_1$ .

First, the material derivatives corresponding to perturbations of the first vertex  $\mathbf{a}_0$  of the regular  $n$ -gon  $\mathbb{P}_n$  are defined by  $\mathbf{U} = (U^1, U^2) \in H_0^1(\mathbb{P}_n)^2$  verifying:

$$(8) \quad \begin{aligned} a(U^1, v) &= \int_{\mathbb{P}_n} (\nabla u \cdot \nabla \varphi) \partial_x v + \int_{\mathbb{P}_n} (\nabla v \cdot \nabla \varphi) \partial_x u - \frac{2\lambda_1}{n} \int_{\mathbb{P}_n} uv, \forall v \in H_0^1(\mathbb{P}_n) \\ a(U^2, v) &= \int_{\mathbb{P}_n} (\nabla u \cdot \nabla \varphi) \partial_y v + \int_{\mathbb{P}_n} (\nabla v \cdot \nabla \varphi) \partial_y u, \forall v \in H_0^1(\mathbb{P}_n) \end{aligned}$$

Since the equations for  $U^1, U^2$  only characterize solutions up to the addition of a multiple of  $u_1$ , the orthogonality conditions

$$(9) \quad \int_{\mathbb{P}_n} U^i u_1 = 0$$

are considered, to fix unique solutions in (8). In [7] it is shown that the Hessian matrix  $D^2\lambda_1(\mathbf{x})$  does not depend on the chosen normalization. Consider the following notations from [7, Prop. 4.10] for  $0 \leq k \leq n-1$ , recalling that  $\theta = 2\pi/n$ :

$$\begin{aligned} A_k &= \sum_{j=0}^{n-1} (\cos(j+1)k\theta + \cos jk\theta) \int_{T_j} \nabla u_1 \cdot \nabla U^1 \\ &\quad + \sum_{j=0}^{n-1} \frac{\cos(j+1)k\theta - \cos jk\theta}{\sin \theta} \int_{T_j} \begin{pmatrix} -\sin(2j+1)\theta & \cos(2j+1)\theta \\ \cos(2j+1)\theta & \sin(2j+1)\theta \end{pmatrix} \nabla u_1 \cdot \nabla U^1 \\ B_k &= \frac{\cos \theta}{\sin \theta} \sum_{j=0}^{n-1} (\cos(j+1)k\theta - \cos jk\theta) \int_{T_j} \nabla u_1 \cdot \nabla U^2 \\ &\quad + \sum_{j=0}^{n-1} \frac{\cos(j+1)k\theta - \cos jk\theta}{\sin \theta} \int_{T_j} \begin{pmatrix} -\cos(2j+1)\theta & -\sin(2j+1)\theta \\ -\sin(2j+1)\theta & \cos(2j+1)\theta \end{pmatrix} \nabla u_1 \cdot \nabla U^2 \\ C_k &= \frac{\cos \theta}{\sin \theta} \sum_{j=0}^{n-1} (\sin(j+1)k\theta - \sin jk\theta) \int_{T_j} \nabla u_1 \cdot \nabla U^1 \\ &\quad + \sum_{j=0}^{n-1} \frac{\sin(j+1)k\theta - \sin jk\theta}{\sin \theta} \int_{T_j} \begin{pmatrix} -\cos(2j+1)\theta & -\sin(2j+1)\theta \\ -\sin(2j+1)\theta & \cos(2j+1)\theta \end{pmatrix} \nabla u_1 \cdot \nabla U^1 \\ D_k &= \sum_{j=0}^{n-1} (\sin(j+1)k\theta + \sin jk\theta) \int_{T_j} \nabla u_1 \cdot \nabla U^2 \\ &\quad + \sum_{j=0}^{n-1} \frac{\sin(j+1)k\theta - \sin jk\theta}{\sin \theta} \int_{T_j} \begin{pmatrix} -\sin(2j+1)\theta & \cos(2j+1)\theta \\ \cos(2j+1)\theta & \sin(2j+1)\theta \end{pmatrix} \nabla u_1 \cdot \nabla U^2 \end{aligned}$$

Then, according to [7, Section 4], the eigenvalues of the Hessian matrix  $D^2(\lambda_1(\mathbf{x})\mathcal{A}(\mathbf{x}))$  associated to the scale invariant formulation (2) are given by the following result.

**Theorem 2.1.** *For  $0 \leq k \leq n-1$ ,  $\theta = 2\pi/n$  denote*

$$\begin{aligned} \alpha_k &= \frac{2n(1 - \cos(k\theta))}{\sin \theta} \int_{T_0} (\partial_x u_1)^2 - 2|\mathbb{P}_n|A_k \\ \beta_k &= \frac{2n(1 - \cos(k\theta))}{\sin \theta} \int_{T_0} (\partial_y u_1)^2 - 2|\mathbb{P}_n|B_k \\ \gamma_k &= -2|\mathbb{P}_n|C_k = 2|\mathbb{P}_n|D_k \end{aligned}$$



and

$$\mu_{2k} = 0.5(\alpha_k + \beta_k - \sqrt{(\alpha_k - \beta_k)^2 + 4\gamma_k^2}), \quad \mu_{2k+1} = 0.5(\alpha_k + \beta_k + \sqrt{(\alpha_k - \beta_k)^2 + 4\gamma_k^2}).$$

The eigenvalues of the Hessian matrix of  $\lambda_1(\mathbf{x})\mathcal{A}(\mathbf{x})$  given in are given by  $\mu_j$ ,  $j = 0, \dots, 2n - 1$ .

Note that  $k = 0$  trivially gives two zero eigenvalues in the formula above. Moreover, in [7, Proposition 4.12] it is shown that the Hessian of  $\lambda(\mathbf{x})\mathcal{A}(\mathbf{x})$  has two additional zero eigenvalues. This is in agreement with the fact that  $\lambda(\Omega)|\Omega|$  is scale invariant: translations, rotations and homotheties do not change the objective function. Proving that the regular  $n$ -gon is a local minimum amounts to the following (see [7, Section 4]).

**Objective.** Given  $n \geq 5$  prove that  $2n - 4$  of the eigenvalues of the Hessian matrix of  $\mathbf{x} \mapsto \lambda_1(\mathbf{x})\mathcal{A}(\mathbf{x})$  are strictly positive.

Theorem 2.1 shows that eigenvalues of the Hessian matrix are explicit in terms of the following quantities:

$$\begin{aligned} & \int_{T_0} (\partial_x u_1)^2, \int_{T_0} (\partial_y u_1)^2, \\ & \int_{T_j} \partial_x u_1 \partial_x U^1, \int_{T_j} \partial_x u_1 \partial_y U^1, \int_{T_j} \partial_y u_1 \partial_x U^1, \int_{T_j} \partial_y u_1 \partial_y U^1, 0 \leq j \leq n - 1, \\ & \int_{T_j} \partial_x u_1 \partial_x U^2, \int_{T_j} \partial_x u_1 \partial_y U^2, \int_{T_j} \partial_y u_1 \partial_x U^2, \int_{T_j} \partial_y u_1 \partial_y U^2, 0 \leq j \leq n - 1. \end{aligned}$$

Since purely theoretical results allowing to prove the positivity of eigenvalues in Theorem 2.1 are not available, numerical computations are used in the proof for  $n \in \{5, 6\}$ . Explicitly quantified error estimates for finite element computations together with interval arithmetic for controlling floating point errors are used.

**2.2. General approximation framework. a) Interpolation errors.** Consider  $\Omega$  a polygon meshed exactly with a triangulation  $\mathcal{T}_h$ . Consider the  $\mathbf{P}_1$  finite element space on  $\mathcal{T}_h$  consisting of piece-wise affine functions on triangles in  $\mathcal{T}_h$ . The norms  $L^2, H^1, H^2$  are considered on the full domain, unless specified otherwise.

In the following denote  $\Pi_{1,h}$  the interpolation operator associated to the  $\mathbf{P}_1$  Lagrange finite element. More precisely,  $\Pi_{1,h}w$  is the piecewise affine function in  $\mathcal{V}_h$  such that  $w$  and  $\Pi_{1,h}w$  have the same values at the nodes of the mesh. Results of [33] show that whenever  $w \in H^2(\Omega)$  we have

$$(10) \quad \|\nabla w - \nabla \Pi_{1,h}(w)\|_{L^2} \leq C_1 h \|D^2 w\|_{L^2}.$$

The above inequality becomes explicit when an upper bound for  $\|D^2 w\|_{L^2}$  is known. The constant  $C_1$  is explicit, depending on the triangles of the mesh  $\mathcal{T}_h$ . We shortly recall of the results of [34, Theorem 4.3].

In each triangle  $T_i \in \mathcal{T}_h$ , the ratio between the smallest edge and the middle one  $L_i$  is denoted  $\alpha_i$  and the angle between these two edges is  $\tau_i$ . Then, we denote

$$C(T_i) := 0.493 L_i \frac{1 + \alpha_i^2 + \sqrt{1 + 2\alpha_i^2 \cos(2\tau_i) + \alpha_i^4}}{\sqrt{2(1 + \alpha_i^2 - \sqrt{1 + 2\alpha_i^2 \cos(2\tau_i) + \alpha_i^4})}}.$$

Following [34, Section 2], we introduce the constant

$$(11) \quad C_1 = \sup_h \frac{C(T_i)}{h},$$

where the parameter  $h$  dictating the size of the mesh is the size of the median edge. In the applications concerning regular polygons it is useful to consider  $\mathcal{T}_h$  consisting of congruent triangles each one similar to  $T_0$  in Figure 1. This renders the constant (11) completely explicit in terms of  $\theta$  and  $h$ .



Since the mesh used of the regular polygon used in the computations is constructed using congruent triangles with angles  $(2\pi/n, \pi/2 - \pi/n, \pi/2 - \pi/n)$  more precise estimates for the interpolation constant can be found through direct estimation. We use the technique described in [29], which shows that discrete problems related to the Morley finite element gives explicit bounds for the constant  $C_1$  in (10). A description of the procedure for obtaining certified upper bounds for constant  $C_1$  defined in (10), giving better bounds than (11) for triangles of interest in our case, is described in Appendix A. The constants used in the validation computations are shown in Table 2.

**b) General Poisson problems.** For  $\Omega$  a polygon in the plane, consider  $\Gamma \subset \partial\Omega$  a subset of its boundary and  $\Gamma_N = \partial\Omega \setminus \Gamma$ . Consider the problem

$$-\Delta w = f, \quad \partial_n w = 0 \text{ on } \Gamma_N, w = 0 \text{ on } \Omega \setminus \Gamma_N$$

where  $f \in H^{-1}(\Omega)$ . The associated variational formulation reads:  $w \in H_\Gamma(\Omega) := \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma \subset \partial\Omega\}$ ,

$$(12) \quad \int_{\Omega} \nabla w \cdot \nabla v = (f, v)_{H^{-1}, H^1}, \text{ for every } v \in H_\Gamma(\Omega).$$

Assume that the solution is  $H^2$  regular, for example if  $\Omega$  is convex.

Suppose  $\Omega$  is polygonal and triangulated, obtaining a mesh  $\mathcal{T}_h$  and an associated space of piecewise affine functions  $v_h \in \mathcal{V}_h$  such that  $v_h$  vanishes on nodes in  $\Gamma$ . Assume  $\mathcal{V}_h \subset H_\Gamma(\Omega)$  and consider  $p_h$  the solution to the problem

$$(13) \quad p_h \in \mathcal{V}_h : \int_{\Omega} \nabla p_h \cdot \nabla v_h = (f, v_h)_{H^{-1}, H^1}, \text{ for every } v_h \in \mathcal{V}_h.$$

Then  $p_h$  is the projection of  $u$ , solution of (12), on  $\mathcal{V}_h$ . In particular, for every  $v_h \in \mathcal{V}_h$  we have

$$\|\nabla w - \nabla v_h\|_{L^2}^2 = \|\nabla w - \nabla p_h\|_{L^2}^2 + \|\nabla p_h - \nabla v_h\|_{L^2}^2 \geq \|\nabla w - \nabla p_h\|_{L^2}^2$$

since  $\int_{\Omega} \nabla(w - p_h) \cdot \nabla(p_h - v_h) = 0$ .

Taking  $v_h = \Pi_{1,h}(w)$  shows that if  $w \in H^2(\Omega)$  then

$$\|\nabla w - \nabla p_h\| \leq C_1 h \|D^2 w\|_{L^2}.$$

If one discretizes also the distribution  $f$ , i.e. consider the problem

$$(14) \quad w_h \in \mathcal{V}_h : \int_{\Omega} \nabla w_h \cdot \nabla v_h = (f_h, v_h)_{H^{-1}, H^1}, \text{ for every } v_h \in \mathcal{V}_h.$$

then  $w_h$  and  $v_h$  verify the same equation with different right hand sides. It is immediate to observe that

$$\|\nabla w_h - \nabla v_h\|_{L^2} \leq \|f - f_h\|_{H^{-1}}.$$

Thus, we obtain

$$\|\nabla w - \nabla w_h\|_{L^2} \leq \|f - f_h\|_{H^{-1}} + C_1 h \|D^2 w\|_{L^2}.$$

**c) Dirichlet Laplacian eigenvalues and eigenfunctions.** Assume the Dirichlet boundary conditions are applied on the whole boundary  $\Gamma = \partial\Omega$  in the following. Given a triangulation  $\mathcal{T}_h$  of the polygon  $\Omega$  and  $\mathcal{V}_h \subset H_0^1(\Omega)$  the  $\mathbf{P}_1$  finite elements space on  $\mathcal{T}_h$ , denote by  $\lambda_{k,h}, u_{k,h}$  the  $k$ -th eigenvalue of  $\Omega$  and its associated eigenfunction approximated in  $\mathcal{V}_h$ , solving

$$(15) \quad u_{k,h} \in \mathcal{V}_h, \int_{\Omega} \nabla u_{k,h} \cdot \nabla v_h = \lambda_{k,h} \int_{\Omega} u_{k,h} v_h, \quad \forall v_h \in \mathcal{V}_h.$$

This problem is a generalized eigenvalue problem.

Results of [33, Theorem 4.3] show that for  $\Omega = \mathbb{P}_n$

$$\forall k \geq 1, \quad \lambda_{k,h} \geq \lambda_k \geq \frac{\lambda_{k,h}}{1 + C_1^2 h^2 \lambda_{k,h}^2},$$

since the associated eigenfunctions are in  $H^2(\Omega)$ . Therefore the following completely explicit estimate holds

$$(16) \quad |\lambda_k - \lambda_{k,h}| \leq \lambda_{k,h}^3 C_1^2 / (1 + C_1^2 h^2 \lambda_{k,h}^2) h^2,$$

with  $C_1$  verifying (10).

For bounds concerning the approximation of the eigenfunction  $u_1$  associated to the simple eigenvalue  $\lambda_1$  on the regular polygon  $\mathbb{P}_n$ , the results in [7, Section 5] are used. The key ingredients are the following:

- Projection of  $u_1$  on  $\mathcal{V}_h$ , defining

$$p_h \in \mathcal{V}_h, \int_{\mathbb{P}_n} \nabla p_h \cdot \nabla v = \lambda_1 \int_{\mathbb{P}_n} u_1 v, \forall v \in \mathcal{V}_h.$$

Like in the estimate for the general Laplace problem it follows that  $\|\nabla u_1 - \nabla p_h\|_{L^2} \leq C_1 h \|D^2 u_1\|_{L^2(\mathbb{P}_n)} = C_1 \lambda_1 h$ , since  $\|D^2 u_1\|_{L^2} = \|\Delta u_1\|_{L^2} = \lambda_1$  [20, Chapter 4]. The convexity of  $\mathbb{P}_n$  and the Aubin-Nitsche lemma implies also that  $\|u_1 - p_h\| \leq C_1^2 \lambda_1 h^2$ .

- Project  $p_h$  on the orthogonal of  $u_{1,h}$  in  $\mathcal{V}_h$ :  $p_h = \alpha u_{1,h} + \bar{p}_h$  with  $\int_{\Omega} \bar{p}_h u_{1,h} = 0$ . Without loss of generality assume  $\alpha > 0$ . This shows that

$$\|\nabla p_h - \nabla u_{1,h}\|_{L^2} = \|(\alpha - 1)\nabla u_h + \nabla \bar{p}_h\|_{L^2} \leq |\alpha - 1| \lambda_{1,h}^{1/2} + \|\nabla \bar{p}_h\|_{L^2}.$$

It follows that

$$\int_{\mathbb{P}_n} (\nabla \bar{p}_h \cdot \nabla v_h - \lambda_{1,h} \bar{p}_h v_h) = \int_{\mathbb{P}_n} (\lambda_1 u_1 - \lambda_{1,h} p) v_h, \forall v_h \in \mathcal{V}_h.$$

Poincaré's inequality in the orthogonal of  $u_{1,h}$  in  $\mathcal{V}_h$  gives

$$(17) \quad \lambda_{2,h}^{1/2} \|\bar{p}_h\|_{L^2} \leq \|\nabla \bar{p}_h\|_{L^2} \leq \frac{\lambda_{2,h}^{1/2}}{(\lambda_{2,h} - \lambda_{1,h})} \left( |\lambda_1 - \lambda_{1,h}| + \lambda_{1,h} \|u_1 - p_h\|_{L^2} \right).$$

Moreover, since  $\|p_h\|_{L^2}^2 = \alpha^2 + \|\bar{p}_h\|_{L^2}^2$  it follows that

$$|1 - \alpha| \leq |1 - \alpha^2| \leq \int_{\mathbb{P}_n} \bar{p}_h^2 + \int_{\mathbb{P}_n} (u_1^2 - p_h^2) \leq \|\bar{p}_h\|_{L^2}^2 + \|u_1 - p_h\|_{L^2} (2 + \|u_1 - p_h\|_{L^2}).$$

Combining these estimates leads to

$$(18) \quad \|\nabla u_1 - \nabla u_{1,h}\|_{L^2} \leq C_1 \lambda_1 h + O(h^2),$$

where the term  $O(h^2)$  is explicit using the inequalities proved above. A similar estimate using  $\|u_1 - u_h\|_{L^2} \leq \|u_1 - p_h\|_{L^2} + \|p_h - u_{1,h}\|_{L^2}$  gives

$$(19) \quad \|u_1 - u_{1,h}\|_{L^2} \leq C_1^2 \lambda_1 h^2 + O(h^2),$$

where all terms are explicit.

The three estimates  $|\lambda_1 - \lambda_{1,h}|$ ,  $\|\nabla u_1 - \nabla u_{1,h}\|_{L^2}$ ,  $\|u_1 - u_h\|_{L^2}$  are linked through the following relation

$$(20) \quad \begin{aligned} \|\nabla u_1 - \nabla u_{1,h}\|_{L^2}^2 - \lambda_1 \|u_1 - u_{1,h}\|_{L^2}^2 &= \int_{\mathbb{P}_n} |\nabla u_1|^2 - 2 \int_{\mathbb{P}_n} \nabla u_1 \cdot \nabla u_{1,h} + \int_{\mathbb{P}_n} |\nabla u_{1,h}|^2 \\ &\quad - \lambda_1 \int_{\mathbb{P}_n} u_1^2 + 2\lambda_1 \int_{\mathbb{P}_n} u_1 u_{1,h} - \lambda_1 \int_{\mathbb{P}_n} u_{1,h}^2 \\ &= \lambda_{1,h} - \lambda_1, \end{aligned}$$

where the weak formulations for  $u_1, u_{1,h}$  are used. Relation (20) is also stated in [8, Equation (3.14)]. One may observe that the three error estimates are related: knowing two of them gives the third one. Since the error estimates for  $\|\nabla u_1 - \nabla u_{1,h}\|_{L^2}$  and  $\|u_1 - u_{1,h}\|_{L^2}$  already contain  $|\lambda_{1,h} - \lambda_1|$ , the estimation procedure can be iterated, replacing a weaker estimate with a stronger one.

It can be observed that even though the relation (20) is stated for the regular polygon  $\mathbb{P}_n$ , it remains valid for every polygon  $\Omega$ , when meshed exactly. It is also verified for higher eigenpairs since the only result used is the variational formulation for  $u_1, u_{1,h}$ .

### 3. MATERIAL DERIVATIVE DECOMPOSITION: PIECEWISE $H^2$ REGULARITY AND ESTIMATES

Similar estimates are needed for solutions of (8), but it turns out that  $U^1, U^2$  do not belong to  $H^2(\mathbb{P}_n)$ . Indeed, the equations solved by  $U^1, U^2$  have singular right hand sides which can be

written as distributions of the form

$$(f^{1,2}, v)_{H^{-1}, H^1} = \int_{\mathbb{P}_n} f_{\text{reg}}^{1,2} v + \sum_{i=0}^{N-1} \int_{S_i} g_i v,$$

where  $f_{\text{reg}}^{1,2} \in L^2(\mathbb{P}_n)$ ,  $S_i$  represent the rays connecting the origin to the vertices  $\mathbf{a}_i$  of  $\mathbb{P}_n$ ,  $g_i \in H^{1/2}(S_i)$ . This motivated the estimates in interpolated spaces from [7], based on the fact that the eigenfunction  $u_1$  belongs to  $H^{2+s}(\mathbb{P}_n)$  for some  $s > 0$ , small.

The main observation is that the solutions of (21) are piece-wise  $H^2$  on every triangle  $T_i$ ,  $i = 0, \dots, n-1$  defined in Figure 1 and the  $H^2$  norms on these triangles can be controlled. This significantly improves the convergence rate and the constants involved, as shown in the remaining of this section.

Consider the context in [7, Section 5]. We propose to give a notably better estimate than the one in Lemma 5.1 of [7], taking into account the particular structure of the problem.

Using the notations from [7], we have for  $\mathbf{U} = \mathbf{U}_0$

$$(21) \quad a(\mathbf{U}, v) = (\mathbf{f}_{\text{reg}}, v) + (\mathbf{f}_{\text{sing}}, v), \text{ for every } v \in H_0^1(\mathbb{P}_n)$$

with  $a(u, v) = \int_{\mathbb{P}_n} (\nabla u \cdot \nabla v - \lambda_1 uv)$ , the equality in (21) being understood in a vectorial sense. Integrating by parts in the right hand side of (21), the regular part has the form

$$(22) \quad (\mathbf{f}_{\text{reg}}, v) = \begin{pmatrix} -2 \int_{T_+ \cup T_-} (\nabla \varphi \cdot \nabla (\partial_x u_1)) v - \frac{2\lambda_1}{n} \int_{\mathbb{P}_n} u_1 v \\ -2 \int_{T_+ \cup T_-} (\nabla \varphi \cdot \nabla (\partial_y u_1)) v \end{pmatrix}.$$

Denoting by  $\partial_r u_1$  the derivative of  $u_1$  in the radial direction. The singular part is given by

$$(23) \quad (\mathbf{f}_{\text{sing}}, v) = \begin{pmatrix} - \int_{S_+} \frac{1}{\tan \theta} \partial_r u_1 v - \int_{S_-} \frac{1}{\tan \theta} \partial_r u_1 v + \int_{S_0} \frac{2}{\tan \theta} \partial_r u_1 v \\ - \int_{S_+} \partial_r u_1 v + \int_{S_-} \partial_r u_1 v \end{pmatrix},$$

where  $\theta = 2\pi/n$ . It is straightforward to see that since  $u_1 \in H^2(\mathbb{P}_n)$ , the distribution  $\mathbf{f}_{\text{reg}}$  can be identified with an element in  $L^2(\mathbb{P}_n)$ . Moreover, a straightforward computation using integration by parts shows that  $(\mathbf{f}_{\text{reg}}, u_1) = 0$ .

The regular part of the distribution, being orthogonal on  $u_1$ , there exists a unique solution for

$$(24) \quad \mathbf{U}_{\text{reg}} \in H_0^1(\mathbb{P}_n)^2, \quad a(\mathbf{U}_{\text{reg}}, v) = (\mathbf{f}_{\text{reg}}, v), \forall v \in H_0^1(\mathbb{P}_n), \int_{\mathbb{P}_n} \mathbf{U}_{\text{reg}} u_1 = 0.$$

Consider now the solution of

$$(25) \quad a(U_{S_0}, v) = \int_{S_0} \partial_r u_1 v - 2c_0 \int_{\mathbb{P}_n} u_1 v, \forall v \in H_0^1(\mathbb{P}_n), \int_{\mathbb{P}_n} U_{S_0} u_1 = 0.$$

where  $2c_0 = \int_{S_0} \partial_r u_1 u_1$  ensures that the right hand side in (25) is orthogonal on  $u_1$ . A simple integration by parts shows that  $c_0 = u_1(0)^2/4$ . In a similar way one can define  $U_{S_-}$  and  $U_{S_+}$  with the right hand side as integrals on the segments  $S_-$  and  $S_+$ . Then by linearity we have

$$\mathbf{U} = \mathbf{U}_{\text{reg}} + \begin{pmatrix} -\frac{1}{\tan \theta} U_{S_+} - \frac{1}{\tan \theta} U_{S_-} + \frac{2}{\tan \theta} U_{S_0} \\ -U_{S_+} + U_{S_-} \end{pmatrix}.$$

The singular behavior of the functions  $(U^1, U^2)$  is underlined in the illustrations given in Figure 2. Normal derivatives of  $U^{1,2}$  across certain segments inside  $\mathbb{P}_n$  are discontinuous. Nevertheless, the numerical computation suggests that piecewise  $H^2$  regularity holds on triangles defined by the symmetric triangulation shown in Figure 1.

Using results in [20, Chapter 2] it is classical that since  $\mathbb{P}_n$  is convex,  $\mathbf{U}_{\text{reg}} \in H^2(\mathbb{P}_n)$ . Following [20, Lemma 4.3.1.3] we know that

$$\|D^2 \mathbf{U}_{\text{reg}}\|_{L^2} = \|\Delta \mathbf{U}_{\text{reg}}\|_{L^2},$$

allowing us to obtain an *a priori* error bound for the associated finite element problem. The key idea is that for functions  $v, w \in H^1(\mathbb{P}_n)$  such that  $(v, w)$  is orthogonal to a fixed vector for every

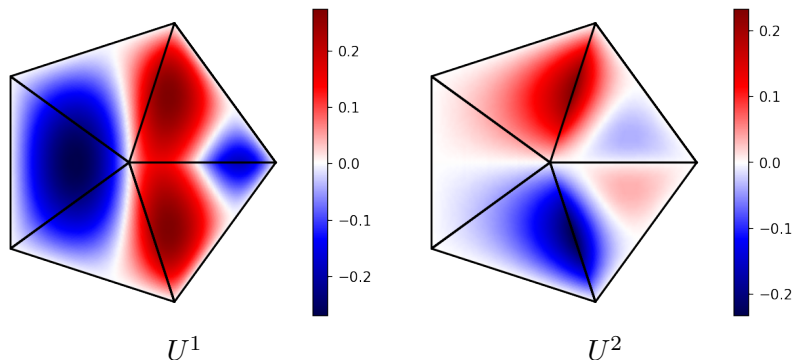


FIGURE 2. Graphical representation of finite element solutions for problems (21). Discontinuities of normal derivatives are visible across certain rays connecting the center of the regular  $n$ -gon to the vertices.

side of  $\mathbb{P}_n$ , it is possible to interchange derivatives in the integration:

$$\int_{\mathbb{P}_n} \partial_x v \partial_y w = \int_{\mathbb{P}_n} \partial_y v \partial_x w.$$

Since  $\mathbf{U}_{\text{reg}}$  is zero on the boundary of  $\mathbb{P}_n$ , its tangential gradient vanishes and applying the previous equality to  $v = \partial_y \mathbf{U}_{\text{reg}}$ ,  $w = \partial_x \mathbf{U}_{\text{reg}}$  gives the  $H^2(\mathbb{P}_n)$  identity stated above.

For the singular part the estimate is more involved. The functions  $U_{S_0}, U_{S_-}, U_{S_+}$  defined by (25) do not belong to  $H^2(\mathbb{P}_n)$  as their gradients are discontinuous on the segments  $S_0, S_-, S_+$ , respectively. More precisely, the tangential gradients are continuous across these segments, while the normal components have jumps. One can notice that  $U_{S_+}$  and  $U_{S_-}$  are rotations of  $U_{S_0}$  with an angle equal to  $\pm 2\pi/n$ . Therefore, initially, it is enough to concentrate the analysis on  $U_{S_0}$ . First, note that  $U_{S_0}$  is even with respect to the  $y$  variable, in view of the symmetry of the domain  $\mathbb{P}_n$ , the orthogonality constraint and the fact that  $S_0$  lies on the horizontal axis, which is a symmetry line of  $\mathbb{P}_n$ .

Even though  $U_{S_0}$  is not in  $H^2(\mathbb{P}_n)$ , it is  $H^2$  regular on both sides of the symmetry axis. Let us denote  $\mathbb{P}_+$  the part of  $\mathbb{P}_n$  contained in the upper half plane  $\{y > 0\}$ . Following results in [20] summarized in [37, Theorem 2.1] it is immediate that  $U_{S_0}|_{\mathbb{P}_+}$  is indeed  $H^2(\mathbb{P}_+)$ , when restricted to  $\mathbb{P}_+$ .

In order to be more precise, consider the problem on half the polygon. The restriction of  $U_{S_0}$  on  $\mathbb{P}_+$ , denoted by  $\overline{U_{S_0}}$  verifies the equation

$$(26) \quad \int_{\mathbb{P}_+} \nabla \overline{U_{S_0}} \nabla v - \lambda_1 \int_{\mathbb{P}_+} \overline{U_{S_0}} v = \frac{1}{2} \int_{S_0} \partial_r u_1 v - c_0 \int_{\mathbb{P}_+} u_1 v,$$

$$(27) \quad \forall v \in H^1(\mathbb{P}_n), v = 0 \text{ on } \partial\mathbb{P}_n \cap \mathbb{P}_+, \int_{\mathbb{P}_+} \overline{U_{S_0}} u_1 = 0.$$

Note that on  $\mathbb{P}_+$ , considering Neumann boundary condition on the symmetry axis,  $\lambda_1$  is still an eigenvalue with the same eigenfunction  $u_1$  (keeping the same notations for restrictions). Moreover,  $\lambda_2(\mathbb{P}_n)$  is also the second eigenvalue of  $\mathbb{P}_+$  with the Dirichlet-Neumann conditions described above, corresponding to a symmetric eigenfunction for  $\lambda_2(\mathbb{P}_n)$  with respect to  $y$ .

Indeed, this can be justified as follows: take a second eigenfunction for the mixed Dirichlet-Neumann problem on  $\mathbb{P}_+$ . By reflection across the horizontal axis, we build an eigenfunction on the full polygon  $\mathbb{P}_n$ , changing sign. This means that  $\lambda_2(\mathbb{P}_+)$  is not smaller than the second eigenvalue of the polygon  $\mathbb{P}_n$ . On the other hand, take a second eigenfunction  $u_2$  for the Dirichlet problem of the polygon  $\mathbb{P}_n$ . For convenience, we see  $u_2$  in polar coordinates  $(\rho, \theta)$ . Then  $\tilde{u}_2(\rho, \theta) := u_2(\rho, -\theta)$  is also a second eigenfunction of the polygon. If  $u_2 + \tilde{u}_2 \not\equiv 0$ , we have build a new second eigenfunction having 0 normal derivative on the horizontal line. So the second eigenvalue of the polygon is also an eigenvalue for the mixed problem on  $\mathbb{P}_+$ , of index higher than 1 since it has to change sign, hence we conclude. If  $u_2 + \tilde{u}_2 \equiv 0$ , then  $u_2$

is a second eigenfunction of the polygon vanishing on the horizontal line. Let us consider the function  $\bar{u}_2(\rho, \theta) = u_2(\rho, \theta + \frac{2\pi}{n})$ . This is another second eigenfunction which cannot vanish on the horizontal line, otherwise it would have more than two nodal domains. So we can reproduce the first argument.

Notice the following:

- $\overline{U_{S_0}} = 0$  on  $\partial\mathbb{P}_n \cap \{y \geq 0\}$
- $\partial_n \overline{U_{S_0}} = 0$  on  $\partial\mathbb{P}_+ \cap \{(x, 0) : x \leq 0\}$
- $\partial_n \overline{U_{S_0}} = \partial_y \overline{U_{S_0}} = \frac{1}{2} \partial_x u_1$  on  $\partial\mathbb{P}_+ \cap \{(x, 0) : x \geq 0\}$ .

Since  $\partial_x u_1(0, 0) = \partial_x u_1(1, 0) = 0$  and  $u_1 \in H^2(\mathbb{P}_+)$  it follows that  $\partial_x u_1 \in H^{1/2}(\partial\mathbb{P}_+ \cap \{y = 0\})$ . It follows that  $\partial_n \overline{U_{S_0}} \in H^{1/2}(\mathbb{P}_+ \cap \{y = 0\})$ . Therefore  $\overline{U_{S_0}} \in H^2(\mathbb{P}_+)$ . As a direct consequence, the solution  $\mathbf{U} = (U^1, U^2)$  of (21) is locally in  $H^2$  when restricted to  $T_+, T_-$  or  $\mathbb{P}_n \setminus (T_+ \cup T_-)$ .

As a first step, let us get an *a priori* estimate on  $\|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}$ . The more problematic term is the mixed term involving mixed derivatives of the form  $\partial_{xy}$ . Below, an approach similar to [20, Lemma 4.3.1.3] is used.

**Lemma 3.1.** *We have:*

$$\int_{\mathbb{P}_+} \partial_{xx} \overline{U_{S_0}} \partial_{yy} \overline{U_{S_0}} - \int_{\mathbb{P}_+} \partial_{xy} \overline{U_{S_0}} \partial_{xy} \overline{U_{S_0}} = \frac{1}{2} \langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})}$$

where  $\tilde{w}$  is the extension by 0 of  $\partial_x u_1|_{S_0}$  on  $\mathbb{R} \setminus S_0$  and  $\tilde{v}$  is any extension in  $H^{\frac{1}{2}+\delta}(\mathbb{R})$  of  $\partial_x \overline{U_{S_0}}|_{S_0}$ .

*Proof:* Following [20, Section 4.3], when  $\Omega$  is a convex polygon  $\mathbf{a}_0, \dots, \mathbf{a}_{m-1}$  with edges  $\Gamma_i = [\mathbf{a}_i, \mathbf{a}_{i+1}]$ ,  $0 \leq i \leq m-1$  (indices considered modulo  $n$  when necessary), considering an orientation of  $\partial\Omega$  in the trigonometric sense, for  $\sigma > 1$  define the space

$$G^\sigma(\Omega) = \{(v, w) \in H^\sigma(\Omega) \times H^\sigma(\Omega) : (\eta_i, \mu_i) \cdot (v, w) = 0 \text{ on } \Gamma_j, 0 \leq j \leq m-2, w(\mathbf{a}_0) = w(\mathbf{a}_{m-1}) = 0\},$$

where  $\eta_i^2 + \mu_i^2 \neq 0$ ,  $0 \leq i \leq m-2$ . Assume furthermore that  $\mu_0 \mu_{m-2} \neq 0$ . Note that compared with [20, Section 4.3], one side is omitted in the definition above and pointwise values are imposed for  $w$  at  $\mathbf{a}_0$  and  $\mathbf{a}_{m-1}$ . Note that the definition of  $G^\sigma$  is meaningful, since pointwise values can be defined for  $H^\sigma$ , when  $\sigma > 1$ .

Consider  $(v, w) \in G^2(\Omega)$  and apply Green's formula:

$$(28) \quad \int_{\Omega} \partial_x v \partial_y w - \int_{\Omega} \partial_y v \partial_x w = \int_{\partial\Omega} v \partial_\tau w,$$

where  $\partial_\tau$  indicates the tangential derivative. On sides  $\Gamma_i$  where  $\mu_i \neq 0$  we have  $w = -\frac{\eta_i}{\mu_i} v$  which implies

$$(29) \quad \int_{\Gamma_i} v \partial_\tau w = -\frac{\eta_i}{\mu_i} \int_{\Gamma_i} v \partial_\tau v = -\frac{\eta_i}{2\mu_i} (v^2(\mathbf{a}_{i+1}) - v^2(\mathbf{a}_i)).$$

Note that this identity is meaningful for  $v \in H^2(\Omega)$ , since  $v \in C(\overline{\Omega})$ . If  $\mu_i = 0$  then by definition we have  $v = 0$  on  $\Gamma_i$ , and by continuity  $v(\mathbf{a}_i) = v(\mathbf{a}_{i+1}) = 0$ . Therefore when summing contributions (29) only vertices  $\mathbf{a}_0, \mathbf{a}_{m-1}$  and those for which the neighboring segments  $\Gamma_{i-1}, \Gamma_i$  have the coefficients  $\mu_{i-1}, \mu_i$  different from zero remain. It is immediate to notice that

$$\int_{\partial\Omega} v \partial_\tau w = \int_{\Gamma_{m-1}} v \partial_\tau w + \frac{1}{2} \sum_{\mu_i \mu_{i-1} \neq 0} \left( \frac{\eta_i}{\mu_i} - \frac{\eta_{i-1}}{\mu_{i-1}} \right) v^2(\mathbf{a}_i) + \frac{\eta_0}{\mu_0} v^2(\mathbf{a}_0) - \frac{\eta_{m-1}}{\mu_{m-1}} v^2(\mathbf{a}_{m-1}).$$

Note that if  $\left( \frac{\eta_i}{\mu_i} - \frac{\eta_{i-1}}{\mu_{i-1}} \right) \neq 0$  then  $(v, w)$  is orthogonal to two non-colinear vectors at  $\mathbf{a}_i$ , giving  $v(\mathbf{a}_i) = 0$ . Therefore

$$(30) \quad \int_{\Omega} \partial_x v \partial_y w - \int_{\Omega} \partial_y v \partial_x w = \int_{\Gamma_{m-1}} v \partial_\tau w + \frac{\eta_0}{\mu_0} v^2(\mathbf{a}_0) - \frac{\eta_{m-1}}{\mu_{m-1}} v^2(\mathbf{a}_{m-1})$$

Let now  $\delta > 0$  (and small) and consider  $(v, w) \in G^{1+\delta}(\Omega)$ . Since  $H^{1+\delta}(\Omega) \subset C(\overline{\Omega})$  continuously, a variant of (30) can be written, as follows. From the proof in [20, Lemma 4.3.1.3] it follows that  $G^2(\Omega)$  is dense in  $G^{1+\delta}(\Omega)$ . Consequently, we can approach  $(v, w)$  strongly in  $H^{1+\delta}(\Omega)$  by

sequences of functions  $(v_\varepsilon, w_\varepsilon)_\varepsilon \subseteq G^2(\Omega)$  and hence write (30) for  $(v_\varepsilon, w_\varepsilon)$ . We also have

$$\begin{aligned} \int_{\Omega} \partial_x v_\varepsilon \partial_y w_\varepsilon - \int_{\Omega} \partial_y v_\varepsilon \partial_x w_\varepsilon &\rightarrow \int_{\Omega} \partial_x v \partial_y w - \int_{\Omega} \partial_y v \partial_x w, \\ \frac{\eta_0}{\mu_0} v_\varepsilon^2(\mathbf{a}_0) - \frac{\eta_{m-1}}{\mu_{m-1}} v_\varepsilon^2(\mathbf{a}_{m-1}) &\rightarrow \frac{\eta_0}{\mu_0} v^2(\mathbf{a}_0) - \frac{\eta_{m-1}}{\mu_{m-1}} v^2(\mathbf{a}_{m-1}). \end{aligned}$$

In order to pass to the limit the term  $\int_{\Gamma_{m-1}} v_\varepsilon \partial_\tau w_\varepsilon$ , some observations are in order. Since  $w_\varepsilon$  vanishes at  $\mathbf{a}_0, \mathbf{a}_{m-1}$ , integration by parts gives  $\int_{\Gamma_{m-1}} v_\varepsilon \partial_\tau w_\varepsilon = -\int_{\Gamma_{m-1}} \partial_\tau v_\varepsilon w_\varepsilon$ . Moreover, denoting  $\tilde{w}_\varepsilon \in H^1(\mathbb{R})$  the extension by 0 of  $w_\varepsilon$  on  $\mathbb{R} \setminus \Gamma_{m-1}$  (here we assume  $\Gamma_{m-1} \subseteq \mathbb{R} \times \{0\}$ ) and identify  $\mathbb{R} \times \{0\}$  with  $\mathbb{R}$  and  $\tilde{v}_\varepsilon$  any extension in  $H^1(\mathbb{R})$  of  $v_\varepsilon$ , we get

$$-\int_{\Gamma_{m-1}} \partial_\tau v_\varepsilon w_\varepsilon = -\int_{\mathbb{R}} \partial_\tau \tilde{v}_\varepsilon \tilde{w}_\varepsilon = \int_{\mathbb{R}} \tilde{v}_\varepsilon \tilde{\partial}_\tau w_\varepsilon = \langle \partial_\tau \tilde{w}_\varepsilon, \tilde{v}_\varepsilon \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})}.$$

For instance, we shall take for  $\tilde{v}_\varepsilon$  any  $H^2$  extension of  $v_\varepsilon$  outside  $\Omega$  which is uniformly bounded in  $H^{1+\delta}(\mathbb{R}^2)$ . Then we have that  $\tilde{v}_\varepsilon \rightarrow \tilde{v}$  and  $\tilde{w}_\varepsilon \rightarrow \tilde{w}$  strongly in  $H^{\frac{1}{2}+\delta}(\mathbb{R})$ , where  $\tilde{v}$  is an extension on  $\mathbb{R} \setminus \Gamma_{m-1}$  of  $v$  and  $\tilde{w}$  is the extension by 0 of  $w$  on  $\mathbb{R} \setminus \Gamma_{m-1}$ . Consequently,

$$\langle \partial_\tau \tilde{w}_\varepsilon, \tilde{v}_\varepsilon \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})} \rightarrow \langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})},$$

hence

$$(31) \quad \int_{\Omega} \partial_x v \partial_y w - \int_{\Omega} \partial_y v \partial_x w = \langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})} + \frac{\eta_0}{\mu_0} v^2(\mathbf{a}_0) - \frac{\eta_{m-1}}{\mu_{m-1}} v^2(\mathbf{a}_{m-1})$$

Apply (31) for  $\Omega = \mathbb{P}_+$ . Let  $k = \lfloor n/2 \rfloor$  and denote  $\mathbf{b} = \mathbf{a}_k$  if  $n$  is even or the midpoint of  $[\mathbf{a}_{k-1}, \mathbf{a}_k]$  if  $n$  is odd. Then the vertices of  $\mathbb{P}_+$  are  $\mathbf{a}_0, \dots, \mathbf{a}_{k-1}, \mathbf{b}$  and the origin  $\mathbf{o}$ .

Furthermore,  $\Gamma_{m-1} = [\mathbf{o}\mathbf{a}_0]$  and taking  $v = \partial_x \overline{U_{S_0}} \in H^{1+\delta}(\mathbb{P}_+)$ ,  $w = \partial_y \overline{U_{S_0}} \in H^{1+\delta}(\mathbb{P}_+)$ . On the segment  $\Gamma_{m-1}$  we have  $w = \frac{1}{2} \partial_x u_1$ , therefore  $w(\mathbf{o}) = w(\mathbf{a}_0) = 0$ . Note that  $w = \partial_y \overline{U_{S_0}}$  verifies  $w = 0$  on  $\Gamma_{m-2}$ , i.e.  $(0, 1) \cdot (v, w) = 0$ . On  $\Gamma_0$  the function  $\overline{U_{S_0}}$  verifies Dirichlet boundary conditions  $\overline{U_{S_0}} = 0$  thus  $\nabla \overline{U_{S_0}} \cdot \overrightarrow{\mathbf{a}_0 \mathbf{a}_1} = 0$ . Moreover, we have  $\partial_y \overline{U_{S_0}}(\mathbf{a}_0) = \partial_x u_1(\mathbf{a}_0) = 0$ . Therefore (31) holds with  $v(\mathbf{a}_0) = 0$  and  $\eta_{m-1} = 0$  and the result follows.  $\square$

**Lemma 3.2** (Extension operator with computed norm). *Let  $\mathbb{P}_n$  be the regular polygon with  $n$  sides and circumcircle of unit radius. There exists a linear extension operator*

$$\mathcal{E} : H^1(\mathbb{P}_n) \rightarrow H^1(\mathbb{R}^2)$$

such that  $\|\mathcal{E}\| \leq C_n$ , where  $C_5 = 4$  and for  $n \geq 6$

$$C_n = \left[ 4 + 24 \cos^2 \frac{2\pi}{n} \right]^{1/2}.$$

**Remark 3.3.** *Note that if such an extension operator exists then the same upper bound  $C_n$  can be used for an extension operator on the half plane  $\mathbb{R}_+$ .*

*Proof:* Given a function in  $H^1(\mathbb{P}_n)$ , an extension to  $H^1(\mathbb{R}^2)$  having controlled norm is constructed using reflections and a cutoff function.

**Reflection.** A distinction needs to be made between the case of the pentagon and  $n \geq 6$ .

Consider the regular pentagon inscribed in the unit circle with center  $\mathbf{o}$ , denoted  $\mathbb{P}_5 = \mathbf{a}_0 \mathbf{a}_1 \mathbf{a}_2 \mathbf{a}_3 \mathbf{a}_4$ . Consider  $\mathbf{b}_1$  the reflection of  $\mathbf{o}$  across the line  $\mathbf{a}_0 \mathbf{a}_1$  and  $\mathbf{b}_{n-1}$  the reflection of  $\mathbf{o}$  across the line  $\mathbf{a}_0 \mathbf{a}_{n-1}$ . The region in the angle  $\angle \mathbf{b}_1 \mathbf{a}_0 \mathbf{b}_{n-1}$  needs to be covered with additional copies of the current figure.

To this end construct the isosceles trapeze  $\mathbf{T} = \mathbf{o}\mathbf{a}_0 \mathbf{c}_1$  with  $\mathbf{c}_1 \in [\mathbf{a}_1 \mathbf{b}_1]$ . Then place two copies of  $\mathbf{T}$ ,  $\mathbf{b}_1 \mathbf{a}_0 \mathbf{x}_0 \mathbf{x}_1$  and  $\mathbf{b}_{n-1} \mathbf{a}_0 \mathbf{x}_0 \mathbf{x}_1$ , in the angle  $\mathbf{b}_1 \mathbf{o} \mathbf{b}_{n-1}$  such that the side congruent to  $\mathbf{a}_0 \mathbf{o}$  in  $\mathbf{T}$  aligns with  $\mathbf{a}_0 \mathbf{b}_1$  and  $\mathbf{a}_0 \mathbf{b}_{n-1}$ . See Figure 3 for the illustration. This reflection is repeated for every vertex of the pentagon. The first stage of the reflection produces a copy of any point in  $\mathbb{P}_n$  and the second stage produces at most 4 extra copies. Thus, globally, the reflection process produces at most 6 extra copies of any point in  $\mathbb{P}_5$ .



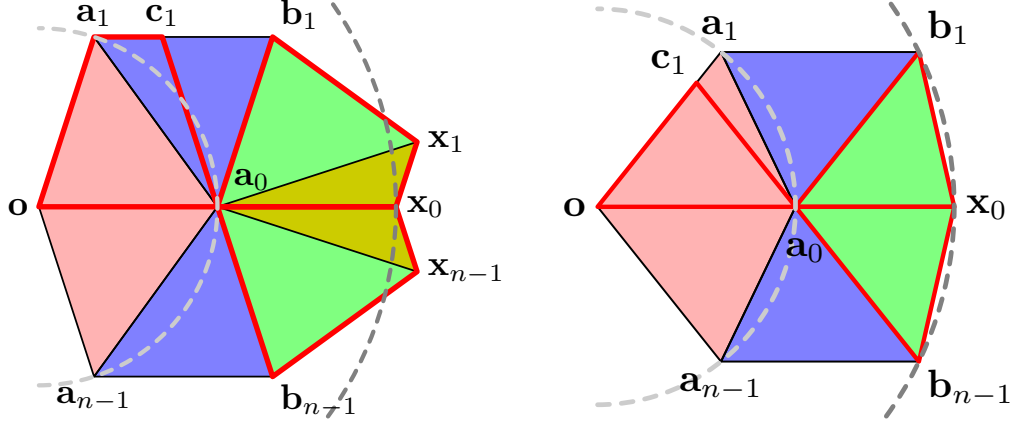


FIGURE 3. Reflection procedure: for the pentagon at least four reflections are needed (left), for  $n \geq 6$  three reflections suffice (right). The thresholds for the application of the cutoff function are also illustrated with dashed lines. The cutoff function equals 1 inside the small disk, zero outside the big disk and is affine in the radial direction in-between.

For  $n \geq 6$  a general extension procedure producing at most 4 extra copies of any point in  $\mathbb{P}_n$  is available. Consider  $\mathbf{a}_0, \dots, \mathbf{a}_{n-1}$  the vertices of  $\mathbb{P}_n$  and construct  $\mathbf{b}_1, \mathbf{b}_{n-1}$  reflections of the origin  $\mathbf{o}$  across  $\mathbf{a}_0\mathbf{a}_1$  and  $\mathbf{a}_0\mathbf{a}_{n-1}$ , respectively.

Since  $\mathbf{o}\mathbf{a}_1 \geq \mathbf{a}_0\mathbf{a}_1$  there exists  $\mathbf{c}_1$  on  $\mathbf{o}\mathbf{a}_1$  such that the triangle  $\mathbf{T} = \mathbf{o}\mathbf{a}_0\mathbf{c}_1$  is isosceles. Consider two copies of  $\mathbf{T}$ , denoted  $\mathbf{b}_1\mathbf{a}_0\mathbf{x}_0$ ,  $\mathbf{b}_{n-1}\mathbf{a}_0\mathbf{x}_0$ , covering the angle  $\angle \mathbf{b}_1\mathbf{a}_0\mathbf{b}_{n-1}$  such that the side  $\mathbf{a}_0\mathbf{o}$  of  $\mathbf{T}$  overlaps with  $\mathbf{a}_0\mathbf{b}_1$ ,  $\mathbf{a}_0\mathbf{b}_{n-1}$ , respectively. This reflection contains at most four copies of any point in  $\mathbb{P}_n$ . See Figure 3 for an illustration.

**Cutoff.** Consider now for any function  $u \in H^1(\mathbb{P}_n)$  the reflected function  $\bar{u}$  following the procedure above and extended arbitrarily afterwards in  $H^1(\mathbb{R}^2)$ . Consider a cutoff function  $\Phi_n$  verifying:

$$\Phi_n = \begin{cases} 1 & |x| \leq 1 \\ 0 & |x| \geq 1 + \ell_n \\ (\ell_n + 1 - |x|)/\ell_n & |x| \in [1, 1 + \ell_n], \end{cases}$$

where  $\ell_n$  is chosen maximal such that the ball of radius  $1 + \ell_n$  is included in the polygon obtained from  $\mathbb{P}_n$  after the reflection process described in the first step.

With these considerations, an explicit extension operator is defined for any  $\mathbb{P}_n$  by

$$\mathcal{E}(u) = \Phi_n \bar{u}.$$

Let us evaluate the  $H^1(\mathbb{R}^2)$  norm of  $\mathcal{E}(u)$  and compare it with the norm of  $u$  in  $H^1(\mathbb{P}_n)$ . Let  $k_n$  be the minimum number of copies for points in  $\mathbb{P}_n$  involved in the construction of the reflected polygon. Then

$$\int_{\mathbb{R}^2} \mathcal{E}(u)^2 = \int_{\mathbb{R}^2} (\Phi_n \bar{u})^2 \leq \int_{B(0,1+\ell_n)} \bar{u}^2 \leq k_n \int_{\mathbb{P}_n} u^2.$$

The similar estimation for the gradient gives

$$\begin{aligned} \int_{\mathbb{R}^2} |\nabla \mathcal{E}(u)|^2 &= \int_{\mathbb{R}^2} |\nabla \Phi_n \bar{u} + \nabla \bar{u} \Phi_n|^2 \\ &\leq 2 \int_{\mathbb{R}^2} |\nabla \Phi_n|^2 \bar{u}^2 + 2 \int_{\mathbb{R}^2} |\nabla \bar{u}|^2 \Phi_n^2 \\ &\leq \frac{2}{\ell_n^2} \int_{B(0,1+\ell_n) \setminus B(0,1)} \bar{u}^2 + 2 \int_{B(0,1+\ell_n)} |\nabla \bar{u}|^2 \\ &\leq 2k_n \int_{\mathbb{P}_n} |\nabla u|^2 + \frac{2(k_n - 1)}{\ell_n^2} \int_{\mathbb{P}_n} \bar{u}^2. \end{aligned}$$



Combining the two estimates gives

$$\|\mathcal{E}(u)\|_{H^1(\mathbb{R}^2)}^2 \leq \max \left\{ 2k_n, \frac{2(k_n - 1)}{\ell_n^2} + k_n \right\}.$$

For the pentagon we have  $k_n = 6$ ,  $\ell_n = 1$  giving  $C_5 = 4$ . For  $n \geq 6$  we have  $k_n = 4$ ,  $\ell_n = \frac{1}{2 \cos \frac{2\pi}{n}}$ , giving

$$C_n = \sqrt{4 + 24 \cos^2 \frac{2\pi}{n}}.$$

□

Combining the previous results we may now find an explicit estimate for  $\|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}$ .

**Theorem 3.4.** *There exists an explicit upper bound for  $\|D^2 \overline{U_{S_0}}\|_{L^2}$  in terms of  $\lambda_1$ ,  $u_1$  and the geometry of  $\mathbb{P}_+$ .*

*The explicit finite element error estimate  $\|\nabla U_{S_0} - \nabla \Pi_{1,h} U_{S_0}\| \leq C_1 h \sqrt{2} \|D^2 \overline{U_{S_0}}\|$  holds for the symmetrized  $U_{S_0}$ , provided the segment  $S_0$  is meshed exactly.*

*Proof:* Consider  $\delta > 0$ ,  $S = [0, 1] \times \{0\}$  and  $f, g \in H^{1/2+\delta}(S)$  such that  $f(0) = f(1) = 0$ . Consequently, we can extend  $f$  by 0 on  $\mathbb{R} \setminus (0, 1)$  to get that  $f \in H^{1/2+\delta}(\mathbb{R})$  and so  $f' \in H^{-\frac{1}{2}+\delta}(\mathbb{R})$ . To simplify notations, assume  $g$  is an arbitrary extension of  $g$  to  $H^{1/2+\delta}(\mathbb{R})$  we have

$$(32) \quad \langle f', g \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})} \leq \|f\|_{H^{1/2}(\mathbb{R})} \|g\|_{H^{1/2}(\mathbb{R})}.$$

Indeed, taking the Fourier transform and noting that  $f$  extends with zero from  $S$  to  $\mathbb{R}$  we have

$$\langle f', g \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})} = \int_{\mathbb{R}} \hat{f} \hat{g}' = \int_{\mathbb{R}} i x \hat{f} \hat{g}.$$

Therefore

$$\begin{aligned} |\langle f', \tilde{g} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})}| &\leq \int_{\mathbb{R}} (1 + |x|^2)^{1/4} |\hat{f}| (1 + |x|^2)^{1/4} |\hat{g}| \\ &\leq \left[ \int_{\mathbb{R}} (1 + |x|^2)^{1/2} |\hat{f}|^2 \right]^{1/2} \left[ \int_{\mathbb{R}} (1 + |x|^2)^{1/2} |\hat{g}|^2 \right]^{1/2} \\ &= \|f\|_{H^{1/2}(\mathbb{R})} \|g\|_{H^{1/2}(\mathbb{R})}. \end{aligned}$$

Next, according to [40, Equation (2.3)] we have

$$(33) \quad \|u\|_{H^{1/2}(\mathbb{R})}^2 \leq C_{1,1} \|u\|_{H^1(\mathbb{R}^2)}^2$$

with  $C_{1,1} = \frac{\Gamma(1/2)}{(4\pi)^{1/2} \Gamma(1)} = \frac{1}{2}$ .

Turning to the estimation regarding  $D^2 \overline{U_{S_0}}$ , since  $-\Delta \overline{U_{S_0}} = -\lambda_1 \overline{U_{S_0}} - c_0 u_1$  we have, using Lemma 3.1 (and its notations for the extensions of  $\partial_x u_1, \partial_x \overline{U_{S_0}}$ ) that

$$\begin{aligned} \|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 &= \int_{\mathbb{P}_+} ((\partial_{xx} \overline{U_{S_0}})^2 + (\partial_{yy} \overline{U_{S_0}})^2 + 2\partial_{xy} \overline{U_{S_0}} \partial_{xy} \overline{U_{S_0}}) \\ &= \int_{\mathbb{P}_+} ((\partial_{xx} \overline{U_{S_0}})^2 + (\partial_{yy} \overline{U_{S_0}})^2 + 2\partial_{xx} \overline{U_{S_0}} \partial_{yy} \overline{U_{S_0}}) - \langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})} \\ &= \|\Delta \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 - \langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})} \\ (34) \quad &= \|\lambda_1 \overline{U_{S_0}} + c_0 u_1\|_{L^2(\mathbb{P}_+)}^2 - \langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})}. \end{aligned}$$

To estimate the first term above start from

$$\|\lambda_1 \overline{U_{S_0}} + c_0 u_1\|_{L^2(\mathbb{P}_+)}^2 = \lambda_1^2 \int_{\mathbb{P}_+} \overline{U_{S_0}}^2 + \frac{c_0^2}{2}.$$

Using  $\overline{U_{S_0}}$  as test function in the variational formulation for  $\overline{U_{S_0}}$  gives

$$\int_{\mathbb{P}_+} |\nabla \overline{U_{S_0}}|^2 - \lambda_1 \int_{\Omega} \overline{U_{S_0}}^2 = \int_{S_0} \partial_r u_1 \overline{U_{S_0}}.$$

and the trace theorem implies

$$\int_S \overline{U_{S_0}}^{-2} \leq \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2.$$

The inequality above is valid for any function  $U$  which is zero on  $\partial\mathbb{P}_n$  through direct application of the mean value theorem on vertical slices. The constant in the trace theorem can be taken equal to 1 since the polygon  $\mathbb{P}_n$  is included in the band  $\{|y| \leq 1\}$ .

Therefore

$$\int_{S_0} \partial_r u_1 \overline{U_{S_0}} \leq \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)} \|\partial_r u_1\|_{L^2(S_0)}.$$

Poincaré's inequality on the orthogonal of  $u_1$  in  $H^1(\mathbb{P}_+) \cap H_0^1(\mathbb{P}_n)$  gives

$$\left(1 - \frac{\lambda_1}{\lambda_2}\right) \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 \leq \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)} \|\partial_r u_1\|_{L^2(S)}.$$

Therefore, again using Poincaré's inequality

$$\sqrt{\lambda_2} \|\overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)} \leq \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)} \leq \frac{\lambda_2}{\lambda_2 - \lambda_1} \|\partial_r u_1\|_{L^2(S)},$$

giving the *a priori* bound

$$\|\overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 \leq \frac{\lambda_2}{(\lambda_2 - \lambda_1)^2} \|\partial_r u_1\|_{L^2(S)}^2.$$

For the second term in (34), using estimate (32) gives

$$|\langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})}| \leq \|\partial_x u_1\|_{H^{1/2}(\mathbb{R})} \|\partial_x \overline{U_{S_0}}\|_{H^{1/2}(\mathbb{R})},$$

where  $\partial_x u_1$  is extended with zero and  $\partial_x \overline{U_{S_0}}$  is extended arbitrarily from  $\mathbb{P}_+ \cap \{y = 0\}$  to  $\mathbb{R}$ . Using the same extension operator defined in Lemma 3.2 and (33) gives

$$\begin{aligned} |\langle \partial_\tau \tilde{w}, \tilde{v} \rangle_{H^{-1/2}(\mathbb{R}), H^{1/2}(\mathbb{R})}| &\leq \frac{C_n^2}{2} \|\partial_x u_1\|_{H^1(\mathbb{P}_n^+)} \|\partial_x \overline{U_{S_0}}\|_{H^1(\mathbb{P}_n^+)} \leq \\ &\leq \frac{C_n^2}{2} \sqrt{(\lambda_1^2 + \lambda_1)/2} \sqrt{\|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 + \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2}, \end{aligned}$$

where the equality  $\|D^2 u_1\|_{L^2(\mathbb{P}_n)}^2 = \|\Delta u_1\|_{L^2(\mathbb{P}_n)}^2 = \lambda_1^2$  was used.

Combining the previous estimates we obtain

$$\|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 \leq \lambda_1^2 \|\overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 + \frac{c_0^2}{2} + \frac{C_n^2}{2} \sqrt{\frac{\lambda_1^2 + \lambda_1}{2}} \sqrt{\|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 + \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2}$$

Denoting  $X = \|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}$  and adding an upper bound for  $\|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2$  to the previous inequality leads to a quadratic inequality of the form

$$X^2 + C \leq B + A\sqrt{X^2 + C},$$

which leads to  $X^2 \leq \left(\frac{A + \sqrt{A^2 + 4B}}{2}\right)^2 - C$ , where:

- $A$  is an upper bound for  $\frac{C_n^2}{2} \sqrt{\frac{\lambda_1^2 + \lambda_1}{2}}$
- $B$  is an upper bound for  $\lambda_1^2 \|\overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2 + \frac{c_0^2}{2} + \|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2$ ,  $c_0 = u_1(0)^2/4 = \|u_1\|_{L^\infty}^2/4 < \lambda_1/4$  (see Grebenkov [19, Formula (6.22)])
- $C$  is an upper bound for  $\|\nabla \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}^2$

This leads to an explicit upper bound for  $\|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}$ . The fact that such an upper bound can be found also shows implicitly that  $\overline{U_{S_0}}$  is indeed in  $H^2(\mathbb{P}_+)$ .

The second part of the statement of the theorem follows at once observing that  $\|D^2 U_{S_0}\|_{L^2(\mathbb{P}_n)} = \sqrt{2} \|D^2 \overline{U_{S_0}}\|_{L^2(\mathbb{P}_+)}$  and applying the interpolation estimate (10). The  $H^2$  norm of the symmetrized  $U_{S_0}$  is understood in a piece-wise sense, considering a crack on the segment  $S_0$ .  $\square$

**General singular distribution on rays.** Let the singular distribution be given by

$$(35) \quad (f, v) = \sum_{i=0}^{n-1} q_i \left( \int_{S_i} \partial_r u_1 v \right)$$

with  $\sum_{i=0}^{n-1} q_i = 0$ . Then any solution of

$$(36) \quad U \in H_0^1(\mathbb{P}_n), \quad a(U, v) = (f, v)_{H^{-1}, H^1}, \quad \int_{\mathbb{P}_n} U u_1 = 0$$

is a linear combination of rotations of the singular solution  $U_{S_0}$ . Considering the analogue equation for the Laplace operator, an improved estimate compared to Lemma 5.1 from [7] is found.

**Lemma 3.5.** *Let  $U \in H_0^1(\mathbb{P}_n)$  be the solution of*

$$-\Delta U = f \text{ in } H_0^1(\mathbb{P}_n)$$

*with  $f$  of the form (35). Consider  $P_h(U) \in \mathcal{V}_h \subset H_0^1(\mathbb{P}_n)$  the discrete solution corresponding to the same right hand side*

$$-\Delta P_h(U) = f \text{ in } \mathcal{V}_h.$$

*Then denoting  $C(\mathbf{q}) = \sum_{i=0}^{n-1} |q_i| C_1 \sqrt{2} \|D^2 \bar{U}_{S_0}\|_{L^2(\mathbb{P}_+)}$  gives*

$$(37) \quad \|\nabla U - \nabla P_h(U)\|_{L^2} \leq C(\mathbf{q})h$$

*and*

$$(38) \quad \|U - P_h(U)\|_{L^2} \leq C_1 C(\mathbf{q})h^2.$$

*Proof.* Simply use the estimates above and the Aubin-Nitsche lemma [9, p.136].  $\square$

This gives rise to the following estimate for solutions of problems similar to those involving the material derivatives.

**Theorem 3.6.** *Let  $U \in H_0^1(\mathbb{P}_n)$  be the solution of*

$$(39) \quad \begin{cases} -\Delta U - \lambda_1 U &= f & \text{in } \mathbb{P}_n \\ U &= 0 & \text{on } \partial \mathbb{P}_n \\ \int_{\mathbb{P}_n} u_1 U dx &= 0 \end{cases}$$

*where  $(f, u_1)_{H^{-1}, H_0^1} = 0$ ,  $f = f_{reg} + f_{sing}$  with  $f_{reg} \in L^2(\mathbb{P}_n)$  and  $f_{sing}$  given by (35). Assume  $f_h$  is a numerical approximation in  $H^{-1}(\mathbb{P}_n)$  of  $f$  which satisfies  $(f_h, u_{1,h})_{H^{-1}, H_0^1} = 0$  and  $(u_{1,h}, \lambda_{1,h})$  a numerical approximation of  $(u_1, \lambda_1)$  in  $H_0^1(\mathbb{P}_n) \times \mathbb{R}$ . Denote  $U_h$  the finite element solution in  $\mathcal{V}_h \subset H_0^1(\mathbb{P}_n)$  for*

$$(40) \quad \forall v \in \mathcal{V}_h, \quad \int_{\mathbb{P}_n} (\nabla U_h \cdot \nabla v - \lambda_{1,h} U_h v) dx = (f_h, v)_{H^{-1} \times H_0^1}$$

*together with the normalization*

$$(41) \quad \int_{\mathbb{P}_n} u_{1,h} U_h dx = 0,$$

*Then, denoting by  $V$  the solution of  $-\Delta V = \lambda_1 U + f$  in  $\mathcal{V}_h$ , we have*

$$\begin{aligned} \|\nabla U - \nabla U_h\|_{L^2} &\leq C_1 h \|\lambda_1 U + f_{reg}\|_{L^2} + C(\mathbf{q})h \\ &+ \lambda_{1,h}^{\frac{1}{2}} \left( (C_1 h)^2 \|\lambda_1 U + f_{reg}\|_{L^2} + C_1 C(\mathbf{q})h^2 \right) + \|V\|_{L^2} \|u_{1,h} - u\|_{L^2} \\ &+ \frac{\lambda_{2,h}^{\frac{1}{2}}}{\lambda_{2,h} - \lambda_{1,h}} \left( |\lambda_{1,h} - \lambda_1| \|U\|_{L^2} + \lambda_{1,h} C_1 C(\mathbf{q})h^2 + (1 + \lambda_{2,h})^{\frac{1}{2}} \|f - f_h\|_{H^{-1}} \right). \end{aligned}$$

The analogue  $L^2$  estimate follows

$$\begin{aligned} \|U - U_h\|_{L^2} &\leq C_1 h (C_1 h \|\lambda_1 U + f_{\text{reg}}\|_{L^2} + C(\mathbf{q})h) \\ &\quad + \left( (C_1 h)^2 \|\lambda_1 U + f_{\text{reg}}\|_{L^2} + C_1 C(\mathbf{q})h^2 \right) + \|V\|_{L^2} \|u_{1,h} - u\|_{L^2} \\ &\quad + \frac{1}{\lambda_{2,h} - \lambda_{1,h}} \left( |\lambda_{1,h} - \lambda_1| \|U\|_{L^2} + \lambda_{1,h} C_1 C(\mathbf{q})h^2 + (1 + \lambda_{2,h})^{\frac{1}{2}} \|f - f_h\|_{H^{-1}} \right). \end{aligned}$$

*Proof.* The proof follows the general estimate strategy: project onto  $\mathcal{V}^h$  using discrete problems with continuous right hand side, then consider discrete problems with different right hand sides. A slight complication arises from the normalization conditions.

**a) Interpolation error.** We denote  $U_{\text{reg}}, U_{\text{sing}}$  the solutions of

$$U_{\text{reg}} \in H_0^1(\mathbb{P}_n), \quad -\Delta U_{\text{reg}} = \lambda_1 U + f_{\text{reg}}, \quad U_{\text{sing}} \in H_0^1(\mathbb{P}_n), \quad -\Delta U_{\text{sing}} = f_{\text{sing}},$$

so that  $U = U_{\text{reg}} + U_{\text{sing}}$ . We introduce the auxiliary functions  $V_{\text{reg}}, V_{\text{sing}} \in \mathcal{V}^h$ ,  $V_{\text{reg}}, V_{\text{sing}}$  the finite element solutions of

$$V_{\text{reg}} \in \mathcal{V}^h, \quad -\Delta V_{\text{reg}} = \lambda_1 U + f_{\text{reg}} \quad V_{\text{sing}} \in \mathcal{V}^h, \quad -\Delta V_{\text{sing}} = f_{\text{sing}}.$$

For  $V_{\text{sing}}$ , the estimate (37) from Lemma 3.5 holds, and gives

$$\|\nabla U_{\text{sing}} - \nabla V_{\text{sing}}\|_{L^2} \leq C(\mathbf{q})h,$$

while for  $V_{\text{reg}}$  the classical estimate gives

$$\|\nabla U_{\text{reg}} - \nabla V_{\text{reg}}\|_{L^2} \leq C_1 h \|\lambda_1 U + f_{\text{reg}}\|_{L^2}.$$

**b) Normalization error.** Let us denote  $V = V_{\text{reg}} + V_{\text{sing}}$  and define  $\tilde{V} = V - (\int_{\mathbb{P}_n} V u_{1,h} dx) u_{1,h}$ . Then we have

$$(42) \quad \|\nabla \tilde{V} - \nabla V\|_{L^2} = \lambda_{1,h}^{\frac{1}{2}} \left| \int_{\mathbb{P}_n} (V u_{1,h} - U u_1) \right| \leq \lambda_{1,h}^{\frac{1}{2}} (\|U - V\|_{L^2} + \|V\|_{L^2} \|u_{1,h} - u\|_{L^2}).$$

The Aubin-Nitsche lemma gives  $\|U - V\|_{L^2} \leq C_1 h \|\nabla U - \nabla V\|_{L^2}$ .

**c) Discrete problems with different right hand sides.** We have that  $\tilde{V}$  is the finite element solution of

$$\tilde{V} \in \mathcal{V}^h, \quad -\Delta \tilde{V} - \lambda_{1,h} \tilde{V} = \lambda_1 U + f - \lambda_{1,h} V,$$

which gives

$$\int_{\mathbb{P}_n} |\nabla \tilde{V} - \nabla U_h|^2 dx - \lambda_{1,h} \int_{\mathbb{P}_n} |\tilde{V} - U_h|^2 dx = (\lambda_1 U + f - \lambda_{1,h} V - f_h, \tilde{V} - U_h)_{H^{-1} \times H_0^1}.$$

By the Poincaré inequality in the orthogonal of  $u_{1,h}$  in  $\mathcal{V}_h$  we get

$$\begin{aligned} \left(1 - \frac{\lambda_{1,h}}{\lambda_{2,h}}\right) \int_{\mathbb{P}_n} |\nabla \tilde{V} - \nabla U_h|^2 dx &\leq \|\lambda_1 U - \lambda_{1,h} V\|_{L^2} \|\tilde{V} - U_h\|_{L^2} + \|f - f_h\|_{H^{-1}} \|\tilde{V} - U_h\|_{H^1} \\ &\leq \|\lambda_1 U - \lambda_{1,h} V\|_{L^2} \lambda_{2,h}^{-\frac{1}{2}} \|\nabla \tilde{V} - \nabla U_h\|_{L^2} + \|f - f_h\|_{H^{-1}} \left(1 + \frac{1}{\lambda_{2,h}}\right)^{\frac{1}{2}} \|\nabla \tilde{V} - \nabla U_h\|_{L^2}. \end{aligned}$$

Finally,

$$(43) \quad \begin{aligned} \|\nabla \tilde{V} - \nabla U_h\|_{L^2(\mathcal{P}_n)} \\ \leq \frac{\lambda_{2,h}^{\frac{1}{2}}}{\lambda_{2,h} - \lambda_{1,h}} \left( |\lambda_{1,h} - \lambda_1| \|U\|_{L^2} + \lambda_{1,h} \|U - V\|_{L^2} + (1 + \lambda_{2,h})^{\frac{1}{2}} \|f - f_h\|_{H^{-1}} \right). \end{aligned}$$

The  $L^2$  estimate follows along the same lines.  $\square$

**Remark 3.7.** It can be observed that all estimates in the proof of Theorem 3.6 can be given in terms of  $\|\nabla U - \nabla V\|_{L^2}$ ,  $\|U - V\|_{L^2}$ ,  $\|u - u_{1,h}\|$ ,  $|\lambda_{1,h} - \lambda_1|$ ,  $\|f - f_h\|_{H^{-1}}$ . All these quantities can be evaluated explicitly in terms of the estimates for the eigenfunctions, eigenvalues and interpolation errors.

Using the notations of Theorem 3.6, consider two generic problems with solutions  $U^a, U^b$  corresponding to the right hand sides  $f^a, f^b$  (not necessarily those explicited in the previous section). Consider the associated notations  $V^a, V^b, \tilde{V}^a, \tilde{V}^b, U_h^a, U_h^b, f_h^a, f_h^b$  analogue to those introduced in the proof of Theorem 3.6. Consider the discrete and continuous the bilinear forms

$$a : H_0^1(\mathbb{P}_n) \times H_0^1(\mathbb{P}_n) \rightarrow \mathbb{R}, \quad a(u, v) = \int \nabla u \cdot \nabla v - \lambda_1 \int uv,$$

$$a_h : \mathcal{V}^h \times \mathcal{V}^h \rightarrow \mathbb{R}, \quad a_h(u, v) = \int \nabla u \cdot \nabla v - \lambda_{1,h} \int uv.$$

**3.1. Estimations for Hessian eigenvalues.** The goal is to estimate the difference between analytical eigenvalues of the Hessian matrix given in Theorem 2.1 and those computed using finite elements. This involves terms of the form

$$|a(U^a, U^b) - a_h(U_h^a, U_h^b)|.$$

Following [7] and the classical triangle inequality the following quantities are estimated, following the structure of the proof in Theorem 3.6:

**First term (interpolation errors).** Observing that  $U^a, U^b$  and  $V^a, V^b$  solve problems having the right hand sides  $f^a, f^b$  in  $H_0^1(\mathbb{P}_n)$  and  $\mathcal{V}^h$ , respectively, it follows that  $a(U^a, V^a) = a(V^a, V^b) = a(V^a, U^b)$ . Therefore

$$|a(U^a, U^b) - a(V^a, V^b)| = |a(U^a - V^a, U^b - V^b)| \leq \|\nabla(U^a - V^a)\|_{L^2} \|\nabla(U^b - V^b)\|_{L^2},$$

giving an estimate of order  $h^2$ . Note that the above follows from  $a(v, v) \leq \int_{\mathbb{P}_n} |\nabla v|^2$  and the Cauchy-Schwarz inequality for the positive bilinear form  $a(\cdot, \cdot)$ .

**Second term (normalization errors).**

$$|a(V^a, V^b) - a_h(\tilde{V}^a, \tilde{V}^b)| \leq \|\nabla V^a\|_{L^2} \|\nabla V^b - \nabla \tilde{V}^b\|_{L^2} + \|\nabla \tilde{V}^b\|_{L^2} \|\nabla V^a - \nabla \tilde{V}^a\|_{L^2} +$$

$$|\lambda_{1,h} - \lambda_1| \|\tilde{V}^a\|_{L^2} \|\tilde{V}^b\|_{L^2} + \lambda_1 \|\tilde{V}^b\|_{L^2} \|V^a - \tilde{V}^a\|_{L^2} + \lambda_1 \|V^a\|_{L^2} \|V^b - \tilde{V}^b\|_{L^2},$$

which, in view of inequality (42), leads to an approximation of order  $h^2$ , since the estimate depends on  $\|U - V\|_{L^2}$  and  $\|u_{1,h} - u_1\|_{L^2}$ .

**Third term (discrete errors).**

$$|a_h(\tilde{V}^a, \tilde{V}^b) - a_h(U_h^a, U_h^b)| \leq |a_h(\tilde{V}^a, \tilde{V}^b - U_h^b)| + |a_h(\tilde{V}^a - U_h^a, U_h^b)| \leq$$

$$\|\nabla \tilde{V}^a\|_{L^2} \|\nabla \tilde{V}^b - \nabla U_h^b\|_{L^2} + \|\nabla U_h^b\|_{L^2} \|\nabla \tilde{V}^a - \nabla U_h^a\|_{L^2}.$$

The last inequality is a consequence of the fact that  $a_h(\cdot, \cdot)$  is a scalar product on  $\{u_{1,h}\}^\perp$  in  $\mathcal{V}^h$  and of the Cauchy-Schwarz inequality together with the observation that  $a_h(v, v) \leq \int |\nabla v|^2$ . Using inequality (43) we get an approximation of order  $h$ , where the  $O(h)$  term is proportional to  $\|f^{a,b} - f_h^{a,b}\|_{H^{-1}}$ , which depends on  $\|\nabla u_1 - \nabla u_{1,h}\|_{L^2(\mathbb{P}_n)}$ .

For the sake of completeness, the required information for estimating the eigenvalues of the Hessian matrix is provided below. Full computations are made in [7]. The notations correspond to those in Theorem 2.1.

**Estimating  $A_k$ .** It can be seen that  $A_k = a(U_0^1, W^{A_k})$  where  $W^{A_k}$  solves

$$a(W, v) = (f_{\text{reg}}^{A_k}, v)_{H^{-1}, H_0^1} + (f_{\text{sing}}^{A_k}, v)_{H^{-1}, H_0^1}, \quad \forall v \in H_0^1(\mathbb{P}_n), \quad \int_{\mathbb{P}_n} W u_1 = 0$$

with

$$(f_{\text{reg}}^{A_k}, v)_{H^{-1} \times H_0^1} = \sum_{j=0}^{n-1} (\cos(j+1)k\theta + \cos jk\theta) \lambda_1 \int_{T_j} u_1 v$$

$$- \sum_{j=0}^{n-1} \frac{\cos(j+1)k\theta - \cos jk\theta}{\sin \theta} \int_{T_j} (-\sin(2j+1)\theta \partial_{xx}^2 u_1 + 2 \cos(2j+1)\theta \partial_{xy}^2 u_1 + \sin(2j+1)\theta \partial_{yy}^2 u_1) v$$

$$(f_{\text{sing}}^{A_k}, v)_{H^{-1} \times H_0^1} = - \sum_{j=0}^{n-1} \int_{S_j} \frac{\cos \theta}{\sin \theta} 2 \cos jk\theta (1 - \cos k\theta) \partial_r u_1 v.$$

An explicit *a priori* estimate for the error between  $a(U_0^1, W^{A_k})$  and the finite element counter part follows from the discussion above.

**Estimating  $B_k$ .** It can be seen that  $B_k = a(U_0^2, W^{B_k})$  where  $W^{B_k}$  solves

$$a(W, v) = (f_{\text{reg}}^{B_k}, v)_{H^{-1}, H_0^1} + (f_{\text{sing}}^{B_k}, v)_{H^{-1}, H_0^1}, \quad \forall v \in H_0^1(\mathbb{P}_n), \quad \int_{\mathbb{P}_n} W u_1 = 0$$

with

$$\begin{aligned} (f_{\text{reg}}^{B_k}, v)_{H^{-1} \times H_0^1} &= \frac{\cos \theta}{\sin \theta} \sum_{j=0}^{n-1} (\cos(j+1)k\theta - \cos jk\theta) \lambda_1 \int_{T_j} u_1 v \\ &- \sum_{j=0}^{n-1} \frac{\cos(j+1)k\theta - \cos jk\theta}{\sin \theta} \int_{T_j} (-\cos(2j+1)\theta \partial_{xx}^2 u_1 - 2 \sin(2j+1)\theta \partial_{xy}^2 u_1 + \cos(2j+1)\theta \partial_{yy}^2 u_1) v \\ (f_{\text{sing}}^{B_k}, v)_{H^{-1} \times H_0^1} &= \sum_{j=0}^{n-1} \int_{S_j} 2 \cos jk\theta (1 - \cos k\theta) \partial_r u_1 v. \end{aligned}$$

An explicit *a priori* estimate for the error between  $a(U_0^2, W^{B_k})$  and the finite element counter part follows from the discussion above. Note that the singular part is even with respect to  $y$ , while the function  $U_0^2$  is odd for both the continuous and discrete problems. Therefore the singular part cancels out and can be neglected in the computations in Theorem 3.6 leading to a stronger estimate.

**Estimating  $C_k$ .** It can be seen that  $C_k = a(U_0^1, W^{C_k})$  where  $W^{C_k}$  solves

$$a(W, v) = (f_{\text{reg}}^{C_k}, v)_{H^{-1}, H_0^1} + (f_{\text{sing}}^{C_k}, v)_{H^{-1}, H_0^1}, \quad \forall v \in H_0^1(\mathbb{P}_n), \quad \int_{\mathbb{P}_n} W u_1 = 0$$

with

$$\begin{aligned} (f_{\text{reg}}^{C_k}, v)_{H^{-1} \times H_0^1} &= \frac{\cos \theta}{\sin \theta} \sum_{j=0}^{n-1} (\sin(j+1)k\theta - \sin jk\theta) \lambda_1 \int_{T_j} u_1 v \\ &- \sum_{j=0}^{n-1} \frac{\sin(j+1)k\theta - \sin jk\theta}{\sin \theta} \int_{T_j} (-\cos(2j+1)\theta \partial_{xx}^2 u_1 - 2 \sin(2j+1)\theta \partial_{xy}^2 u_1 + \cos(2j+1)\theta \partial_{yy}^2 u_1) v \\ (f_{\text{sing}}^{C_k}, v)_{H^{-1} \times H_0^1} &= - \sum_{j=0}^{n-1} \int_{S_j} 2 \sin jk\theta (1 - \cos k\theta) \partial_r u_1 v. \end{aligned}$$

An explicit *a priori* estimate for the error between  $a(U_0^1, W^{C_k})$  and the finite element counter part follows from the discussion above. Note that the singular part is odd with respect to  $y$ , while the function  $U_0^1$  is even for both the continuous and discrete problems. Therefore the singular part cancels out and can be neglected in the computations in Theorem 3.6 leading to a stronger estimate.

The practical estimates for all quantities involved in the proof of Theorem 3.6 follows from the following:

- Theorem 3.4, Lemma 3.5
- Effective practical estimates computed in [7, Section 5]

#### 4. NUMERICAL VALIDATION

This section provides a series of results which are used in the validated computing process for the discrete eigenvalue problems and discrete linear systems. Although Intlab [42] has functions dedicated to these types of problems, applying the corresponding functions directly leads to inefficient computations overcoming, ultimately, the available computational power. To simplify computational aspects, various theoretical results are shown in the following.

Given a mesh size  $h$  and a numerical computation, all quantities in the *a priori* estimates are explicit. In all our computations the mesh considered is symmetric and corresponds to a regular polygon inscribed in the unit circle with one vertex at  $\mathbf{a}_0 = (1, 0)$ . See Figure 4 for an

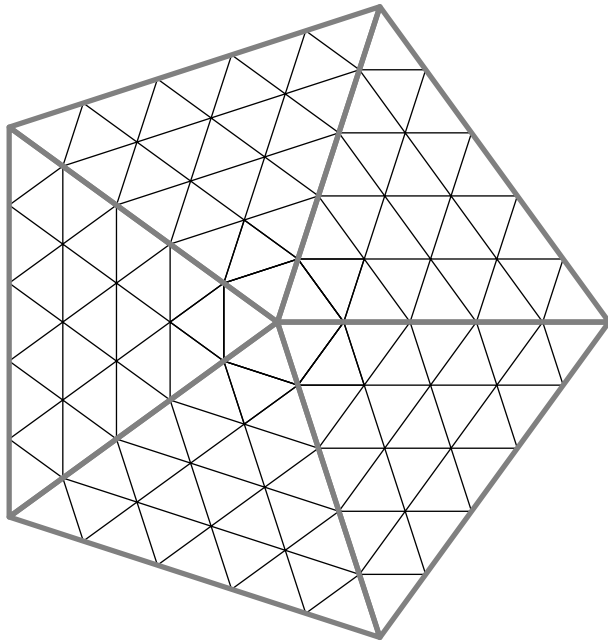


FIGURE 4. Symmetric mesh for a regular pentagon for  $m = 5$ : the ray  $[\mathbf{o}\mathbf{a}_0]$  is divided into  $m$  equal segments.

illustration. The mesh size  $h$  is given by  $1/m$  where  $m$  is the number of segments in which  $[\mathbf{o}\mathbf{a}_0]$  is divided. For a given parameter  $m$  the mesh has  $N = 1 + nm(m + 1)/2$  nodes.

First, the eigenvalues, eigenvectors and solutions to linear systems should be enclosed with certified intervals using interval arithmetic in Intlab [42]. Let  $K, M$  be the rigidity and mass matrices for the  $\mathbf{P}_1$  finite elements, defined by

$$K = \left( \int_{\mathbb{P}_n} \nabla \psi_i \cdot \nabla \psi_j \right)_{1 \leq i, j \leq N}, M = \left( \int_{\mathbb{P}_n} \psi_i \psi_j \right)_{1 \leq i, j \leq N},$$

where  $\psi_i$  denotes the piece-wise affine function on the mesh of  $\mathbb{P}_n$  taking the value 1 at the node  $1 \leq i \leq N$  and zero for all other nodes.

We are interested in solving the generalized eigenvalue problem  $Ku = \lambda Mu$ , imposing Dirichlet boundary conditions on nodes in  $\partial\mathbb{P}_n$ . This can be achieved as follows. Denote by  $K_0, M_0 \in \mathbb{R}^{d \times d}$  the submatrices of  $K$  and  $M$  corresponding to the inner nodes. Then the eigenvalue equation corresponding to the Dirichlet boundary condition is simply

$$(44) \quad K_0 u = \lambda M_0 u.$$

For  $h$  small enough, the eigenvalues of (44) having small magnitude approximate the eigenvalues of the Dirichlet Laplace operator on  $\mathbb{P}_n$ . One first objective is to solve (44) rigorously, for a given mesh size, including verified enclosures for the eigenvalues  $\lambda_{1,h}, \lambda_{2,h}$  and the first eigenvector  $u_{1,h}$ . In order to keep notations simple, since in this section we always deal with  $\lambda_i$  and  $u_i$  as solutions of linear systems, we drop the mention of  $h$  in the previous notations. The verification process can be achieved using Intlab [42] and every element in  $K_0, M_0$  should be provided using an explicit interval enclosure by floating point numbers.

The symmetry of the mesh of  $\mathbb{P}_n$  illustrated in Figure 4 indicates that all triangles in the mesh are similar to an isosceles triangle with the top angle equal to  $2\pi/n$ . In particular, the matrices  $K, M$  have precise expressions for all non-zero elements. For interior nodes, not lying on the boundary, we indicate the formulas corresponding to elements in  $K_0, M_0$ .

**Proposition 4.1.** *The elements of the rigidity matrix  $K_0$  corresponding to nodes not lying on the boundary of  $\mathbb{P}_n$  are the following:*

$$- i \text{ is the central node: } K_{ii} = n \tan \frac{\theta}{2}, K_{ij} = - \tan \frac{\theta}{2}$$



-  $i$  is not a central node:  $K_{ii} = \frac{2}{\sin\theta} + 2 \tan \frac{\theta}{2} = \frac{2}{\tan\theta} + 4 \tan \frac{\theta}{2}$ ,  $K_{ij} = -\frac{1}{\tan\theta}$  (2 times per row/column),  $K_{ij} = -\tan \frac{\theta}{2}$  (4 times per row/column)

If  $A_h = 0.5h^2 \sin\theta$  is the area of one triangle in the mesh, the elements of the mass matrix  $M_0$  corresponding to nodes not lying on the boundary of  $\mathbb{P}_n$  are

- $i$  is the central node:  $M_{ii} = nA_h/6$ ,  $M_{ij} = A_h/6$
- $i$  is not the central node:  $M_{ii} = A_h$ ,  $M_{ij} = A_h/6$ .

When dealing with equation (44) multiple theoretical aspects facilitate the validation process.

**Proposition 4.2.** a) The matrices  $K_0, M_0$  are non-singular and eigenpairs for (44) correspond to eigenpairs for the matrix  $M_0^{-1}K_0$ .

b)  $K_0$  is an  $M$ -matrix, i.e.  $K_0 = sI - B$  with  $B \geq 0$  (all elements of  $B$  are non-negative). Therefore  $K_0^{-1}$  has non-negative entries.

c) Furthermore,  $K_0^{-1}$  is positive (all entries are strictly positive). Therefore  $K_0^{-1}M_0$  is positive and the Perron-Frobenius theorem implies its largest eigenvalue is simple and corresponds to a positive eigenvector.

d) The smallest eigenvalue of  $M_0^{-1}K_0$  is simple and corresponds to a positive eigenvector. All other eigenvectors contain both positive and negative terms.

*Proof:* a) The result follows at once. Denote by  $I$  the set of inner nodes of the mesh and  $N_0$  the number of inner nodes. Consider  $x = (x_i)$  the basis coordinates of  $\psi = \sum_{i=1}^{N_0} x_i \psi_i$ , which is a  $\mathbf{P}_1$  finite element function verifying  $\psi = 0$  on  $\partial\mathbb{P}_n$ . Therefore, we have

$$x^T K_0 x = \|\nabla\psi\|_{L^2}^2 \geq 0, x^T M_0 x = \|\psi\|_{L^2}^2 \geq 0.$$

If one of the previous quantities is zero then obviously  $x = 0$ . Positive definiteness follows.

b) The fact that  $K_0$  is an  $M$ -matrix follows at once from Proposition 4.1. Furthermore, since each line or column of  $K_0$  has at least one non-diagonal entry which is strictly negative, each line and column of  $B$  has a non-diagonal entry which is strictly positive.

Let  $s > 0$  be the maximal element on the diagonal of  $K_0$ . Then  $\frac{1}{s}K_0 = I - \frac{1}{s}B$ . The spectral radius of  $\frac{1}{s}B$  is smaller than 1 (see Theorem 8.1.18 from [24]). Therefore  $(\frac{1}{s}K_0)^{-1} = \sum_{k=0}^{\infty} (\frac{1}{s}B)^k$ , which has non-negative entries since  $B \geq 0$ .

c) Let  $e_i$  be a vector of the canonical basis of  $\mathbb{R}^{N_0}$  and  $y = K_0^{-1}e_i \geq 0$ . Then

$$(sI - B)y = e_i \implies sy = By + e_i.$$

Since  $s > 0$ , if  $y = (y_j)_{j=1}^{N_0}$  it follows that  $y_i > 0$ . Next, if  $B_{jk} > 0$  then  $y_j > 0$  if and only if  $y_k > 0$ . To see this observe that

$$sy_j \geq B_{jk}y_k \text{ and } sy_k \geq B_{jk}y_j.$$

Inductively, since  $y_i > 0$  all successive neighbors of the node  $i$  verify  $y_j > 0$ . Finally, since the mesh is connected, we get  $y > 0$ . Since the result is valid for all elements of the canonical basis, we have  $K_0^{-1} > 0$ .

The matrix  $M_0$  is non negative by construction and has non-zero elements on every line and column. Therefore  $K_0^{-1}M_0 > 0$ .

The Perron-Frobenius theorem implies that the largest eigenvalue of  $K_0^{-1}M_0$  is simple and corresponds to an eigenvector whose elements are strictly positive. See [24, Section 8.2] for more details.

d) The largest eigenvalue of  $K_0^{-1}M_0$  corresponds to the inverse of the smallest eigenvalue of  $M_0^{-1}K_0$ . Therefore the smallest eigenvalue of (44) is simple and corresponds to a positive eigenvector. All other eigenvector verifies  $u_j M u_1 = 0$ , for all  $j > 1$ . Since  $M$  has positive entries on every line and all elements of  $u_1$  are positive, each  $u_j$ ,  $j > 1$  must have both positive and negative elements.  $\square$

The following result is useful for the numerical validation of the first three eigenvalues of the regular  $n$ -gons. Proposition 4.2 already shows that the first discrete eigenvalue  $\lambda_{1,h}$  is simple and corresponds to an eigenvector with constant sign. The following result shows that the second eigenvalue is multiple and its cluster is isolated from the rest by the fourth eigenvalue on the unit

disk  $B_1$ . Therefore, for  $h$  small enough, if two distinct discrete eigenvalues smaller than  $\lambda_4(B_1)$  are found for (15) then these correspond to  $\lambda_{1,h}$  and  $\lambda_{2,h}$ .

**Proposition 4.3.** *a) The second eigenvalue of the regular  $n$ -gon is at least double. If  $\mathcal{T}_h$  is a symmetric mesh of  $\mathbb{P}_n$  like in Figure 4 then the discrete eigenvalue  $\lambda_{2,h}$  for (15) is multiple.*

*b) Consider  $n \geq 5$ . For a mesh size  $h > 0$  small enough, quantifiable depending on  $n$ , there exist at most three discrete eigenvalues for (15) in the interval  $[0, \lambda_4(B_1)]$ , where  $B_1$  is the unit disk. For  $n \geq 5$ , these three eigenvalues correspond to:  $\lambda_{1,h}$ , which is simple, and the cluster of  $\lambda_{2,h}$  with multiplicity 2.*

Note that on the regular polygon the second eigenvalue is precisely double. This could be justified with some extra arguments, but this is not necessary for our further analysis.

*Proof:* a) The regular  $n$ -gon is connected, therefore the first eigenvalue  $\lambda_1$  is simple and corresponds to an eigenfunction  $u_1$  having constant sign. Consider the second eigenvalue  $\lambda_2$  and an associated eigenfunction  $u_2$ . It is known [1] that the nodal line corresponding to  $u_2$  touches the boundary at exactly two points.

Assume  $\lambda_2$  is simple. Then any rotation  $\tilde{u}_2$  of  $u_2$  with a multiple of  $2\pi/n$  around the origin gives an eigenfunction for  $\lambda_2$ . Since the nodal line of  $u_2$  touches the boundary at exactly two points, it is impossible that  $\tilde{u}_2$  is a multiple of  $u_2$ . Therefore the multiplicity of  $\lambda_2$  is at least 2.

Assume the regular  $n$ -gon is meshed with a symmetric mesh  $\mathcal{T}_h$  like in Figure 4. Consider  $\lambda_{2,h}$  a second eigenvalue for the discrete problem (15) together with a second eigenvector  $u_{2,h}$ . Since  $u_{2,h}$  is orthogonal on  $u_{1,h}$ , which has constant sign, there exist positive and negative elements in  $u_{2,h}$ . Consider a permutation of the elements of  $u_{2,h}$  corresponding to a rotation of angle multiple of  $2\pi/n$  on the mesh. The positivity and negativity of elements of  $u_{2,h}$  cannot remain invariant for all such rotations. Therefore the discrete eigenvalue  $\lambda_{2,h}$  is also multiple. Passing to the limit when  $h \rightarrow 0$  this also gives an alternative proof for the multiplicity of  $\lambda_2(\mathbb{P}_n)$ .

b) Since the regular  $n$ -gon  $\mathbb{P}_n$  is inscribed in the unit ball, we have  $\mathbb{P}_n \subset B_1$ , therefore  $\lambda_k(\mathbb{P}_n) \geq \lambda_k(B_1)$  for any  $k \geq 1$ . In [39] it is shown that  $\lambda_1(\mathbb{P}_n)$  is decreasing with respect to  $n$ . Such a result is not available for higher eigenvalues. Whenever one can show that  $\lambda_2(\mathbb{P}_n) = \lambda_3(\mathbb{P}_n) < \lambda_4(B_1)$ , the result follows. For  $n = 4$  the second eigenvalue of the square is  $\lambda_2(\mathbb{P}_4) = 2.5\pi^2 < \lambda_4(B) = j_{2,1}^2 \approx 5.13\dots^2 = 26.31\dots$ , where  $j_{2,1}$  is the first zero of the Bessel function of the second kind. Therefore, for all  $n \geq 4$  one has  $\lambda_2(\mathbb{P}_n) = \lambda_3(\mathbb{P}_n) < \lambda_4(B_1)$ . For  $h$  small enough the discrete eigenvalues verify the same inequality. This will be checked *a posteriori* in all our computations.  $\square$

In [36] multiple results regarding validated computations regarding eigenvalues and eigenvectors are presented. Theorem 16 in the previous reference shows that in the particular case of symmetric matrices it is possible to obtain bounds for all eigenvalues and eigenvectors associated to simple eigenvalues in terms of residuals. A modification of this result corresponding to generalized eigenvalues is given below, notably using the natural matrix norm associated to the mass matrix. In some situations, this allows to avoid the usage of eigenvalue and eigenvector verification routines (like `verifyeig` in Intlab [42]) which are costly for large matrices.

**Proposition 4.4.** *Let  $K, M$  be symmetric matrices in  $\mathbb{R}^{N \times N}$  such that  $K, M$  are positive definite. Let  $\tilde{x} \in \mathbb{R}^N$  and  $\tilde{\lambda} \in \mathbb{R}$  be an approximate eigenpair. Denote the residual vector  $r := K\tilde{x} - \tilde{\lambda}M\tilde{x}$  and  $|\tilde{x}|_M^2 = \tilde{x}^T M \tilde{x}$ . Then the following assertions hold:*

a) *The interval  $\left[ \tilde{\lambda} - \sqrt{\frac{r^T M^{-1} r}{\tilde{x}^T M \tilde{x}}}, \tilde{\lambda} + \sqrt{\frac{r^T M^{-1} r}{\tilde{x}^T M \tilde{x}}} \right]$  contains at least one generalized eigenvalue of  $(K, M)$ .*

b) *If  $a \geq r$ ,  $|\tilde{x}|_M = 1$  and if  $[\tilde{\lambda} - a, \tilde{\lambda} + a]$  contains a simple eigenvalue  $\lambda^*$  but no more eigenvalues of  $(K, M)$  then there is an eigenvector  $x^*$  with associated with  $\lambda^*$  verifying*

$$|\tilde{x} - x^*|_M^2 \leq \frac{r^T M^{-1} r}{a^2} + \max \left\{ (|\tilde{x}|_M - 1)^2, \left( 1 - \sqrt{\tilde{x}^T M \tilde{x} - \frac{r^T M^{-1} r}{a^2}} \right)^2 \right\},$$

*provided the square root exists.*

*Proof:* The proof could follow from [36, Theorem 16] by a change of variables, choosing  $A = M^{-1/2}KM^{-1/2}$ , but for the sake of completeness the following simple arguments are given below.

a) Since generalized eigenvalues of  $(K, M)$  correspond to eigenvalues of  $M^{-1}K$ , which is symmetric, positive definite, there exists a basis of  $\mathbb{R}^N$  made of generalized eigenvectors  $(v_k)$  of  $(K, M)$  associated to eigenvalues  $\lambda_k > 0$ ,  $k = 1, \dots, N$ . The vectors  $v_k$  are assumed orthonormal with respect to the matrix  $M$ , in other words,  $v_i^T M v_j = \delta_{ij}$ . Therefore, if  $\tilde{x} = \sum_{i=1}^N \alpha_i v_i$  then  $\tilde{x}^T M \tilde{x} = \sum_{k=1}^N \alpha_k^2$ . On the other hand

$$r = K\tilde{x} - \tilde{\lambda}M\tilde{x} = \sum_{k=1}^N (\lambda_k - \tilde{\lambda})\alpha_k M v_k.$$

This implies that

$$(45) \quad r^T M^{-1} r = \sum_{i=1}^N (\lambda_i - \tilde{\lambda})^2 \alpha_i^2 \geq \left( \min_{k=1, \dots, N} |\lambda_k - \tilde{\lambda}|^2 \right) \tilde{x}^T M \tilde{x}.$$

This proves the first assertion.

b) Assume  $\lambda^* = \lambda_k$  is simple with eigenvector  $v_k$  the normalization conventions described previously. Using the decomposition  $\tilde{x} = \sum_{i=1}^N \alpha_i v_i$  we find

$$\tilde{x}^T M \tilde{x} = \sum_{j=1}^N \alpha_j^2.$$

The hypothesis on the separation of  $\tilde{\lambda}$  with respect to  $\lambda_j$ ,  $j \neq k$  gives

$$r^T M^{-1} r = (\lambda_k - \tilde{\lambda})^2 \alpha_k^2 + \sum_{j \neq k} (\lambda_j - \tilde{\lambda})^2 \alpha_j^2 \geq a^2 (\tilde{x}^T M \tilde{x} - \alpha_k^2).$$

On the other hand, the previous inequality gives

$$|\tilde{x} - v_k|_M^2 = \tilde{x}^T M \tilde{x} - 2\alpha_k + 1 \leq \frac{r^T M^{-1} r}{a^2} + (\alpha_k - 1)^2.$$

Without loss of generality, assume  $\alpha_k \geq 0$ . If  $\alpha_k \geq 1$  then  $1 \leq \alpha_k \leq \sqrt{\tilde{x}^T M \tilde{x}}$ , therefore

$$|\tilde{x} - v_k|_M^2 \leq \frac{r^T M^{-1} r}{a^2} + (|\tilde{x}|_M - 1)^2$$

For  $0 \leq \alpha_k < 1$  we have  $\alpha_k \geq \sqrt{|\tilde{x}|_M^2 - \frac{r^T M^{-1} r}{a^2}}$ , provided  $|\tilde{x}|_M^2 - \frac{r^T M^{-1} r}{a^2} \geq 0$ . In this case

$$|\tilde{x} - v_k|_M^2 \leq \frac{r^T M^{-1} r}{a^2} + \left( 1 - \sqrt{|\tilde{x}|_M^2 - \frac{r^T M^{-1} r}{a^2}} \right)^2.$$

The conclusion follows.  $\square$

**Remark 4.5.** *It can be observed that the technique used to prove the second point in Proposition 4.4 gives an estimate concerning the distance to the eigenspace associated to a cluster in case the eigenvalue is multiple. In this case the residual together with the minimal distance from the cluster to the closest neighboring eigenvalues is enough to give the estimate.*

*The expressions involving the matrix norm  $|\cdot|_M$  can be evaluated either explicitly by solving a validated linear system or precise properties of the matrix  $M$  (which is very particular in our case: see Proposition 4.1) can help to turn these estimates into a more explicit form.*

*According to [26] the eigenvalues of the mass matrix verify*

$$\frac{1}{2} \min M_{ii} \leq \lambda_{\min}(M) \leq \lambda_{\max}(M) \leq 2 \max M_{ii}.$$

*The explicit formulas for  $M_{ii}$  give*

$$\frac{1}{12} \min\{n, 6\} A_h \leq \lambda_{\min}(M) \leq \lambda_{\max}(M) \leq \frac{1}{3} \max\{n, 6\} A_h,$$

where  $A_h = 0.5h^2 \sin \theta$  is the area of a small triangle in the mesh shown in Figure 4. Therefore

$$r^T M^{-1} r \leq \frac{1}{\lambda_{\min}(M)} |r|^2 \leq \frac{12}{\min\{n, 6\} A_h} |r|^2, \quad |x|^2 \leq \frac{1}{\lambda_{\min}(M)} |x|_M^2 \leq \frac{12}{\min\{n, 6\} A_h} |x|_M^2.$$

Similar inequalities remain valid for  $(K_0, M_0)$  which are submatrices of  $(K, M)$ . Classical interlacing inequalities show that eigenvalues of  $K_0, M_0$  verify the same bounds.

A second key point is the approximation of the material derivatives defined in (8) which imply solving the discrete systems associated to the finite element formulations. For simplicity, the index  $h$  is dropped in the following. It should be noted that in the remaining of this section, only the discrete problem is considered. In this case, given the (discrete) eigenfunction  $u_1$  associated to  $\lambda_1$  the system

$$(46) \quad \begin{pmatrix} A & b \\ b^T & 0 \end{pmatrix} \begin{pmatrix} U \\ 0 \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}$$

needs to be solved for

$$A = K_0 - \lambda M_0 \in \mathbb{R}^{d \times d}, b = M_0 u \in \mathbb{R}^d.$$

The compatibility condition  $f^T u_1 = 0$  needs to be verified by  $f \in \mathbb{R}^d$ , where  $u_1$  is the eigenvector associated to the smallest eigenvalue  $\lambda_1$ .

Recall that  $T_{\pm}$  are the two triangles delimited by  $y = 0$  and  $y = \pm x \tan \frac{2\pi}{n}$ . Next,  $\varphi$  is the  $\mathbf{P}_1$  function defined on the triangulation of  $\mathbb{P}_n$  with  $n$  triangles where one node is at the center (see Figure 1) taking value 1 only at  $(1, 0)$ . We have

$$\nabla \varphi = \left(1, -\frac{1}{\tan \theta}\right) \chi_{T_+} + \left(1, \frac{1}{\tan \theta}\right) \chi_{T_-},$$

where  $\theta = 2\pi/n$ . The variational formulation for (21) reads:  $\mathbf{U} \in H_0^1(\mathbb{P}_n)^2$  verifying

$$a(\mathbf{U}, v) = \int_{\mathbb{P}_n} \left[ (\nabla u_1 \cdot \nabla \varphi) \nabla v + (\nabla \varphi \cdot \nabla v) \nabla u_1 - \begin{pmatrix} 2\lambda_1/n \\ 0 \end{pmatrix} u_1 v. \right]$$

The explicit equations verified  $\mathbf{U} = (U_1, U_2)$  defined in (8) are as follows:  $U_i \in H_0^1(\mathbb{P}_n)$

$$(47) \quad \int_{\mathbb{P}_n} \nabla U_1 \cdot \nabla v - \lambda_1 U_1 v = \int_{T_+ \cup T_-} (2\partial_x u \partial_x v + \partial_y \varphi (\partial_x u \partial_y v + \partial_y u \partial_x v)) - 2 \frac{\lambda_1}{n} \int_{\mathbb{P}_n} uv$$

$$(48) \quad \int_{\mathbb{P}_n} \nabla U_2 \cdot \nabla v - \lambda_1 U_2 v = \int_{T_+ \cup T_-} (2\partial_y \varphi \partial_y u \partial_y v + (\partial_x u \partial_y v + \partial_y u \partial_x v))$$

for every  $v \in H_0^1(\mathbb{P}_n)$ , with constraints  $\int_{\mathbb{P}_n} U_i u_1 = 0$ .

Consider the following partial rigidity matrices:

$$K_{xx} = \left( \int_{T_+ \cup T_-} \partial_x \psi_i \partial_x \psi_j \right)_{i,j=1}^{N_0}, \quad K_{yy} = \left( \int_{T_+} \partial_y \psi_i \partial_y \psi_j - \int_{T_-} \partial_y \psi_i \partial_y \psi_j \right)_{i,j=1}^{N_0}$$

$$K_{xy}^{\pm} = \left( \int_{T_{\pm}} \partial_x \psi_i \partial_y \psi_j + \partial_y \psi_i \partial_x \psi_j \right)_{i,j=1}^{N_0}$$

Note that these matrices are only considered for interior vertices of the mesh, since Dirichlet boundary conditions are taken on  $\partial \mathbb{P}_n$ . For the triangulation illustrated in Figure 4 the elements of the matrices above are all explicit in terms of  $\theta = \frac{2\pi}{n}$  and the usual trigonometric functions. The terms can be found using the following result.

**Proposition 4.6.** *For a triangle which is homothetic to  $T_+$  with factor  $h$ , with vertices labeled  $i, j, k$  in the order of  $(0, 0), (1, 0), (\cos \theta, \sin \theta)$ ,  $\theta = \frac{2\pi}{n}$ , the gradients of the associated  $\mathbf{P}_1$  finite element functions  $\psi_i, \psi_j, \psi_k$  are given by*

$$\begin{aligned} \partial_x \psi_i &= -\frac{1}{h}, & \partial_x \psi_j &= \frac{1}{h}, & \partial_x \psi_k &= 0 \\ \partial_y \psi_i &= -\frac{1}{h} \tan \frac{\theta}{2}, & \partial_y \psi_j &= -\frac{1}{h \tan \theta}, & \partial_y \psi_k &= \frac{1}{h \sin \theta}. \end{aligned}$$

These formulas coupled with the fact that all elements in the symmetric mesh have area  $\frac{1}{2}h^2 \sin \theta$  show that the elements of  $K_{xx}, K_{yy}, K_{xy}, K_{yx}$  are completely determined in function of  $\theta$  and do not depend on  $h$ .

Then the discrete systems associated to (47), (48) when using  $\mathbf{P}_1$  finite elements have the form

$$(49) \quad (K_0 - \lambda_1 M_0)U_1 = 2K_{xx}u_1 - \frac{1}{\tan \theta}(K_{xy}^+ - K_{xy}^-)u_1 - \frac{2\lambda_1}{n}M_0u_1$$

$$(50) \quad (K_0 - \lambda_1 M_0)U_1 = -\frac{2}{\tan \theta}K_{yy}u_1 + (K_{xy}^+ + K_{xy}^-)u_1,$$

verifying the following orthogonality relations

$$(51) \quad u_1 M_0 U_1 = u_1 M_0 U_2 = 0.$$

More precisely, the system of the form (46) needs to be solved for

$$A = K_0 - \lambda_1 M_0, b = M_0 u_1,$$

where  $u_1$  is the first generalized eigenvector associated to  $(K_0, M_0)$  and the right hand sides are computed using (49), (50).

According to [27], in order to precondition the saddle point system it is enough to consider a perturbation of the form

$$(52) \quad A(w) = A + wbb^T$$

and

$$(53) \quad \mathcal{A}(w) = \begin{pmatrix} A(w) & 0 \\ 0 & b^T A(w)^{-1} b \end{pmatrix}.$$

Then according to [27] and [16] we have the following result:

**Lemma 4.7.** a) Denoting  $\mathcal{H} = \begin{pmatrix} A & b \\ b^T & 0 \end{pmatrix}$ , if  $w > 0$ ,  $A$  is positive semi-definite and  $\mathcal{H}$  is non-singular then  $A(w)$  given in (52) is positive definite.

b) The eigenvalues of  $\mathcal{A}(w)^{-1}\mathcal{H}$  belong to

$$\left[-1, \frac{1 - \sqrt{5}}{2}\right] \cup \left[1, \frac{1 + \sqrt{5}}{2}\right].$$

c) If  $U^*$  is the exact solution of  $\mathcal{H}U^* = \bar{f} = \begin{pmatrix} f \\ 0 \end{pmatrix}$  then

$$(54) \quad \|U^* - U\|_2 \leq \frac{2}{\sqrt{5} - 1} \max\{\|A(w)^{-1}\|_2, \|A(w)\|_2 \|b^T b\|_2^{-1}\} \|\mathcal{A}U - \bar{f}\|_2.$$

*Proof:* a) Let  $x \in \mathbb{R}^d$ . Then  $x^T A(w)x = x^T Ax + w(b^T x)^2 \geq 0$ , since  $A$  is positive semi-definite. Moreover,  $x^T A(w)x = 0$  implies  $x^T Ax = 0$  and  $b^T x = 0$ . Therefore If  $\bar{x} = (x, 0)$  then  $\mathcal{H}\bar{x} = 0$  and since  $\mathcal{H}$  is non-singular,  $x = 0$ . Therefore  $A(w)$  is positive definite.

b) This is proved in [16, Theorem 2.5].

c) This is proved in [27, Theorem 2.2]. □

In Lemma 4.7 we have are two degrees of freedom regarding the choice of parameters. The first one is the scalar parameter  $w$ . Secondly, notice that multiplying the vector  $b$  with a scalar does not change solutions to the saddle point system (46). In the following  $b$  is chosen of the form  $b = \gamma M_0 u_1$ .

In the following, the dimension  $d$  of the system (46) is  $N_0$ , the number of nodes in the mesh, not lying on the boundary  $\partial\mathbb{P}_n$ . Consider the set of generalized eigenpairs  $(\lambda_i, u_i)$ ,  $i = 1, \dots, d$  for (44) which forms a basis of  $\mathbb{R}^d$ , orthonormalized with respect to  $M_0$ . Assume  $\lambda_1 \leq \dots \leq \lambda_d$ . For a general  $x \in \mathbb{R}^d$  let  $x = \sum_{i=1}^d \alpha_i v_i$  its decomposition in this basis.

Then, for  $b = \gamma M_0 u_1$  we have

$$x^T A(w)x = x^T (K_0 - \lambda_1 M_0)x + \gamma^2 w (u_1^T M_0 x)^2 = \sum_{i=1}^d \alpha_i^2 (\lambda_i - \lambda_1) + \gamma^2 w \alpha_1^2,$$

since  $b^T x = \gamma \sum_{i=1}^n \alpha_i u_i^T M_0 u_1 = \gamma \alpha_1$ . Therefore  $x^T A(w)x \geq \min\{\lambda_2 - \lambda_1, \gamma^2 w\}$ . Since  $A(w)$  is symmetric,  $\lambda_1$  is simple (Proposition 4.2) and  $w > 0$  we find

$$\|A(w)^{-1}\|_2 \leq \max\left\{\frac{1}{\lambda_2 - \lambda_1}, \frac{1}{\gamma^2 w}\right\}.$$

Continuing the estimates, we have  $\|A(w)\|_2 \|b^T b\|_2^{-1} \leq (\|A\|_2 + w \|b\|^2) / \|b\|^2 \leq \|A\|_\infty / \|b\|_2^2 + w$ . Replacing  $b = \gamma M_0 u_1$  gives

$$\|A(w)\|_2 \|b^T b\|_2^{-1} \leq \frac{\|K_0 - \lambda_1 M_0\|_\infty}{\gamma^2 \|M_0 u_1\|_2^2} + w.$$

The matrix norm  $\|K_0 - \lambda_1 M_0\|_\infty$  is easily evaluated using Intlab. Thus, once  $w$  and  $\gamma$  are chosen, estimate (54) is explicit once the residual  $\|\mathcal{A}U - f\|_2$  is known.

The inequalities above show that:

- if  $\gamma^2 w \geq \lambda_2 - \lambda_1$  then  $\|A(w)^{-1}\|_2 \leq 1/(\lambda_2 - \lambda_1)$ .
- choosing  $w = (\lambda_2 - \lambda_1)/\gamma^2$  gives

$$\|A(w)\|_2 \|b^T b\|_2^{-1} \leq \frac{\|K_0 - \lambda_1 M_0\|_\infty}{\gamma^2 \|M_0 u_1\|_2^2} + \frac{\lambda_2 - \lambda_1}{\gamma^2},$$

therefore, when  $\gamma$  increases the estimate gets stronger.

- Nevertheless, taking  $\gamma$  large will increase the interval evaluating how well the floating point approximation verifies the normalization constraint  $u_1 M_0 U = 0$ . Choosing  $\gamma$  too large will deteriorate the global residual estimate. In practical computations, choosing  $\gamma = \gamma_0 / \|M_0 u_1\|_2$  for  $\gamma_0 \in [1, 10]$  gives satisfying results.

Given  $\mathcal{H}$  the explicit matrix from (46), an interval enclosure  $\overline{\mathcal{H}}$  is considered. If enclosures for the discrete eigenvector  $u_1$  are available, then an enclosure  $\overline{f}$  for the right hand side of (49), (50) are computed. Given  $U$  a floating point solution found for solving (46) for the system given by the midpoint of  $\overline{\mathcal{H}}$  (through an iterative approach, Conjugate Gradient for example) with right hand side the midpoint of  $\overline{f}$ , the residual  $\|\overline{\mathcal{H}}U - \overline{f}\|_2$  is evaluated using interval arithmetic. Thus, a validated upper bound for  $\|\mathcal{H}U - f\|_2$  is found and an explicit bound for the distance  $\|U - U^*\|_2$  is obtained.

## 5. VALIDATION STRATEGY, IMPLEMENTATION AND RESULTS

**5.1. Detailed description of the validation procedure.** In this section, the validation strategy is presented, based on results proved in previous sections. All computations are performed using interval arithmetic in Intlab [42].

### Inputs.

- $n$ , the number of vertices of the regular polygon
- $m$  the number of segments in which the segment  $[0, 1]$  is divided in the mesh shown in Figure 4. This gives rise to the mesh size parameter  $h = 1/m$ .

The regular polygon  $\mathbb{P}_n$  is meshed with a mesh  $\mathcal{T}_h$  like in Figure 4, in particular the rays  $S_i$  are meshed exactly. The mesh is constructed manually and the submesh  $\mathcal{T}_h^0$  corresponding to points in  $\{x \geq 0, y \in [0, x \tan \theta]\}$  is considered. An explicit mapping allows to extend finite element functions on  $\mathcal{T}_h^0$  to  $\mathcal{T}_h$  respecting the dihedral symmetry. See Figure 5 for an example.

### A. Finite element computations.

**a) Rigidity and Mass matrices.** Following Proposition 4.1 the elements of the rigidity and mass matrices  $K_0, M_0$  associated to the interior nodes have explicit formulas in terms of  $\theta = 2\pi/n$ . The matrices are constructed using these expressions in Intlab. Thus  $K_0, M_0$  are interval matrices containing the exact rigidity and mass matrices. Since  $K_0, M_0$  are submatrices associated to the interior nodes, this automatically imposes the Dirichlet boundary conditions.



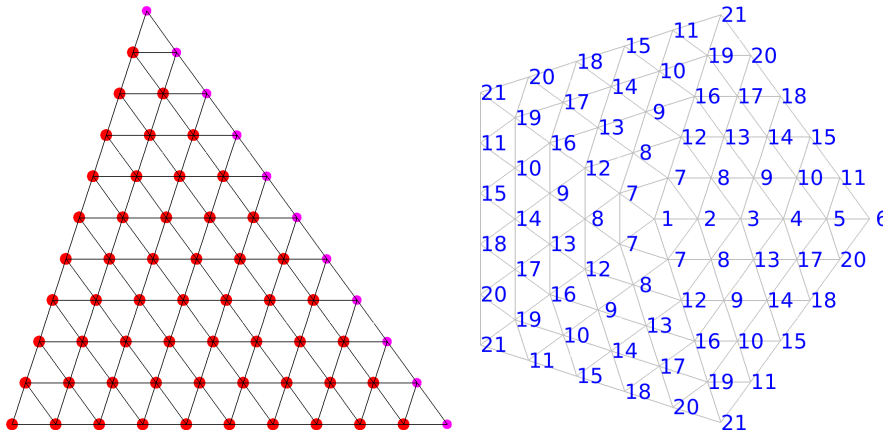


FIGURE 5. Mesh of a triangular slice  $\mathcal{T}_h^0$  of the regular  $n$ -gon (left). Mapping from the slice  $\mathcal{T}_h^0$  to the whole mesh used to extend a function by dihedral symmetry (right).

**b) The first discrete eigenvalue and eigenfunction.** The first two eigenvalues for the discrete eigenproblem (15),  $K_0 u_{i,h} = \lambda_{i,h} M_0 u_{i,h}$  are computed initially using floating point arithmetic. Proposition 4.4 gives two interval enclosures for  $\lambda_{1,h}$  and  $\lambda_{2,h}$  in terms of the residuals. Following Proposition 4.3 if these two intervals together with the *a priori* estimates (16) give upper bounds smaller than the fourth eigenvalue of the unit ball  $\lambda_4(B_1)$  then the two enclosures correspond indeed to the first two discrete eigenvalues of  $(K_0, M_0)$ .

For the first eigenvalue  $\lambda_{1,h}$ , which is simple (see Proposition 4.2) an enclosure is found for the first eigenvector  $u_{1,h}$ . This is achieved either using a modification of the routine `verifyeig` in Intlab [42] (where the matrix inversion is replaced with three validated linear systems solved with `verifylss`) applied on the slice  $\mathcal{T}_h^0$  or by using the direct bound using the residual, shown in Proposition 4.4. Whenever Intlab can be used for bounding the first eigenvector it gives a tighter enclosure than the residual estimate.

The second eigenvalue is always validated using the residual estimate given in Proposition 4.4. For the second eigenvalue the whole mesh of  $\mathbb{P}_n$  is used. For large meshes, the existence of symmetric eigenfunctions for  $\lambda_{2,h}$  allows to use only half the mesh for finding an interval enclosure for the second discrete eigenvalue. The associated matrices can be obtained taking submatrices corresponding to the full mesh and replacing the diagonal values corresponding to  $m/2$  nodes underlined in Figure 6.

**c) Approximation of the material derivatives.** Next, all information required for solving (49), (50) is computed using interval arithmetic. In particular the partial rigidity matrices described in Proposition 4.6 are computed using exact formulas in Intlab.

Systems (49), (50) are solved initially in floating point arithmetic using the Conjugate Gradient algorithm. An initialization with the zero vector guarantees that the orthogonality condition on  $u_{1,h}$  is preserved since the right hand sides are orthogonal to  $u_{1,h}$ .

Following Lemma 4.7 the residuals for the linear systems are computed using interval arithmetic for the floating point results obtained previously. All variables appearing in the estimate (54) are computed in floating point arithmetic. Notably, the interval enclosure of  $\lambda_{2,h} - \lambda_{1,h}$  obtained in the previous step is used.

In the end, interval enclosures for solutions of (49) and (50), verifying the orthogonality conditions (51), are obtained.

**B. Computation of the eigenvalues of the Hessian matrix.** Up to this point interval enclosures are available for  $\lambda_{1,h}, \lambda_{2,h}, u_{1,h}, U_h^1, U_h^2$ . These are used to compute the eigenvalues of the Hessian matrix of the first eigenvalue, given in Theorem 2.1.

The computation of the integrals of partial derivatives of  $u_{1,h}$  and  $U_h^i$  is achieved by assembling partial rigidity matrices for all triangles  $T_i$  corresponding to slices of the regular  $n$ -gon into  $n$  parts like in Figure 1. These matrices depend again explicitly on  $\theta = 2\pi/n$ .



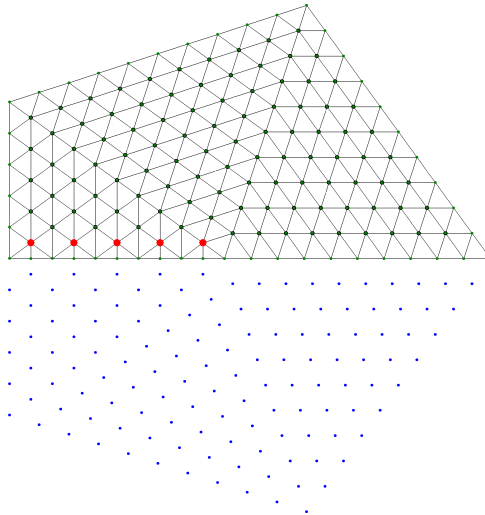


FIGURE 6. Using half the mesh of  $\mathbb{P}_n$  to solve the discrete problem associated to the second eigenvalue in (15). The associated mass and rigidity matrices are submatrices of those for the full polygon. The diagonal elements of the points underlined need to be modified slightly.

At the end of this step, the discrete approximations  $\mu_j^h$ ,  $j = 0, \dots, 2n - 1$  for the eigenvalues of the Hessian are available with guaranteed interval enclosures.

**C. A priori estimates for finite elements.** The final step is controlling the distances  $|\mu_j - \mu_j^h|$ ,  $j = 0, \dots, 2n - 1$  using the *a priori* error estimates for finite element computations:

- $|\lambda_k - \lambda_{k,h}|$  is bounded by (16) according to [33].
- $\|\nabla u_1 - \nabla u_{1,h}\|_{L^2}$  and  $\|u_1 - u_{1,h}\|_{L^2}$  are bounded by (18), (19).
- In some cases, the equality (20) may be used to iteratively replace the weakest of the three estimates above with a better one, therefore improving all estimates.

Finally, results in Section 3 are used to quantify the finite element errors for the solutions of the material derivatives (47), (48) and for equations described in Section 3.1 involved in the estimation of the errors for the Hessian eigenvalues.

**D. Certification of local minimality.** At the end of the previous step a guaranteed error estimate is found for each one of the Hessian eigenvalues in Theorem 2.1. According to [7, Section 4] four of the eigenvalues of the Hessian on the regular  $n$ -gon are identically equal to zero. Thus, if  $2n - 4$  of the interval enclosures found for the Hessian eigenvalues are included in  $(0, +\infty)$ , then the numerical certification of local minimality succeeds: the regular  $n$ -gon is a local minimizer for the first Dirichlet-Laplace eigenvalue.

**5.2. Simulations and Results.** Given  $n$  the number of vertices and  $m$  the number of segments on the ray  $[\mathbf{o}\mathbf{a}_0]$  in the mesh illustrated in Figure 4, the computation described in the previous section is performed. The validation procedure succeeds for  $n \in \{5, 6\}$ . The case  $n \geq 7$  is inaccessible for the moment. It is not a question of computational power: the *a priori* error estimates are not small enough compared to the size of the intervals associated to the variables appearing in the computation. Details are shown below.

**The regular pentagon.** The computations described in the previous section are performed for meshes corresponding to  $m \in \{200, \dots, 600\}$ . It can be seen in Figure 7 that for  $m$  large enough the  $2n - 4$  intervals containing the non-zero eigenvalues of the Hessian matrix (see Theorem 2.1) are contained in  $(0, \infty)$ , therefore these eigenvalues are guaranteed to be positive. This shows that the regular pentagon is a local minimizer for problem (2). The validation procedure succeeds already for  $m = 250$ . Multiple observations are given below:

(a) **Evolution of intervals enclosures with respect to  $h = 1/m$ .** The size of the intervals stabilizes, even though the mesh becomes finer and finer. This is due to the quick growth in the

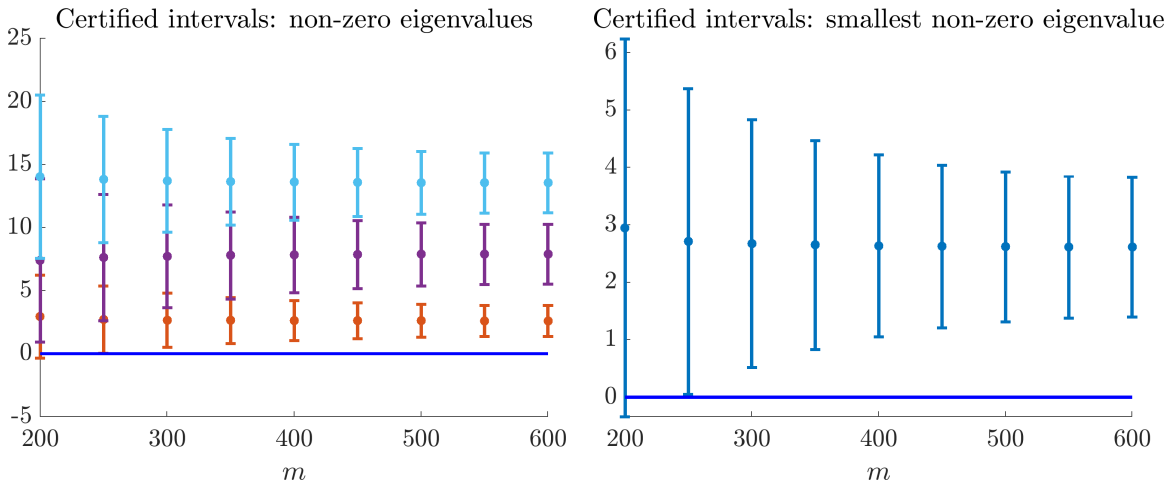


FIGURE 7. Computations for the **regular pentagon**. Left: certified intervals containing the non-zero eigenvalues of the Hessian (three pairs of double eigenvalues). Right: certified intervals containing the smallest non-zero eigenvalue.

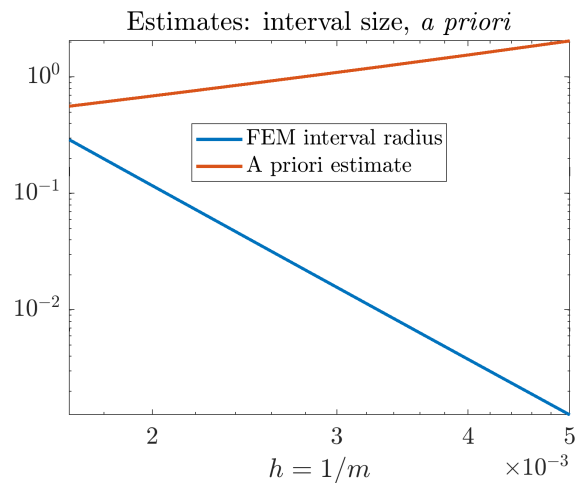


FIGURE 8. Computations for the **regular pentagon**. The radius of the intervals for the FEM computation of the Hessian eigenvalues is shown with respect to the mesh size  $h = 1/m$ , together with the corresponding *a priori* estimate, whose leading term is  $O(h)$ .

interval size for the eigenvalues  $\mu_j^h$  computed with finite elements (formulas from Theorem 2.1 with finite element variables). Multiple factors contribute to this:

- The condition number of the rigidity matrix behaves like  $O(h^{-2}) = O(m^2)$  (see [26]). Therefore, the performance of linear solvers becomes worse for finer meshes. This also affects the interval enclosures produced by Intlab [42].
- The linear systems (49), (50) have wider right hand sides for fine meshes, therefore produce wider results.
- The Hessian eigenvalues computations depend on integrals of the form  $\int_{T_i} \partial_{x,y} u_{1,h} \cdot \partial_{x,y} U_h^{1,2}$ . These evaluations produce intervals proportional to the number of triangles in the slice  $T_i$ , which is  $O(m^2) = O(h^{-2})$ .
- The combined effect of all these aspects is illustrated in Figure 8 where it can be seen that the interval size increases drastically when  $h \rightarrow 0$ . The *a priori* estimate verify the theoretical  $O(h)$  bound, while the interval sizes behave like  $O(h^{-5})$ .

(b) **Evaluation of finite element quadratures with intervals.** Let us discuss in more detail the evaluation of finite element terms of the form  $\int_{T_i} \partial_{x,y} u_{1,h} \cdot \partial_{x,y} U_h^{1,2}$  with intervals. It is

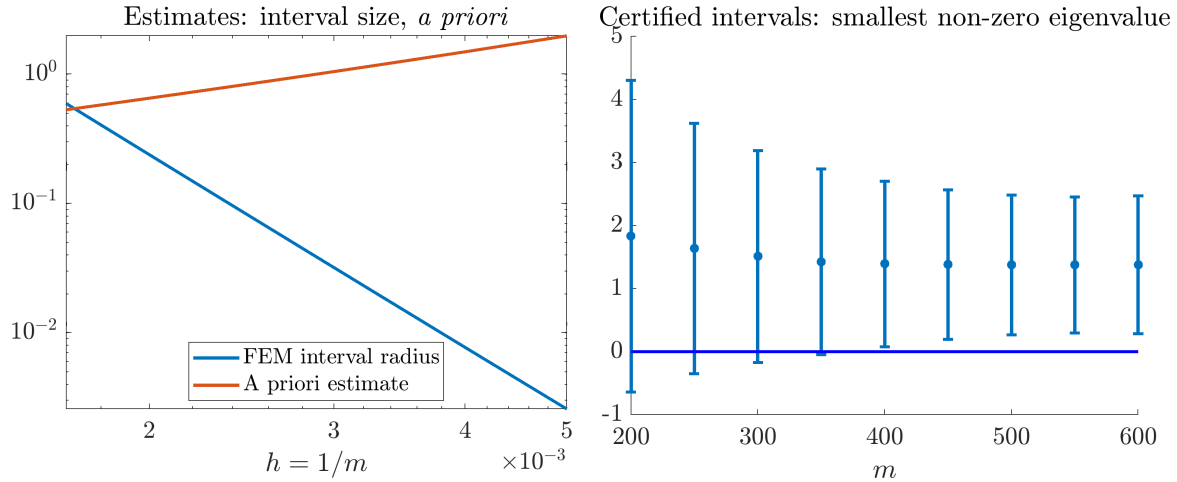


FIGURE 9. Computations for the **regular hexagon**. Left: evolution of the *a priori estimates* and the size of the intervals for  $\mu_j^h$  coming from the FEM computations. Right: certified intervals containing the smallest non-zero eigenvalue. The validation of local minimality succeeds in this case.

apparent from Proposition 4.6 that the elements of the rigidity matrices used to compute these quantities do not depend on the scaling of the triangle. The area of the triangle is canceled when computing the integral of a term of size  $1/h^2$ . Therefore, assuming  $u_h, U_{1,2}^h$  are computed as solutions of linear systems and eigenvalue problems, obtaining intervals of a given size  $\rho$ , assuming this size is not worsened when  $h \rightarrow 0$  (this is often not the case), the interval enclosure for  $\int_{T_i} \partial_{x,y} u \cdot \partial_{x,y} U^{1,2}$  will be proportional to the number of triangles in the mesh of  $T_i$  times  $\rho$ , giving  $O(m^2 \rho) = O(\rho h^{-2})$  in our case. Therefore, the sole evaluation of the quadrature of a triangulation produces a drastic increase in the resulting intervals.

Another interpretation for the increase in interval sizes when dealing with integrals of derivatives of  $\mathbf{P}_1$  functions given with intervals, can be given as follows. If the discrete functions are given as intervals of size  $\rho$  then an interval enclosure of the integral includes the worst case scenario when the function oscillates with amplitude  $\rho$  from one vertex to the next one. Thus, errors are proportional to the number of oscillations, i.e. the number of triangles.

For terms of the form  $\int_{T_i} \nabla u_{1,h} \cdot \nabla U_h^{1,2}$ , the fact that  $u_{1,h}$  is a discrete eigenvalue should be used, to obtain the equivalent formula  $\lambda_{1,h} \int_{T_i} u_{1,h} U_h^{1,2}$ . This latter integral is computed using a mass matrix, whose elements are proportional to  $h^2$ . Therefore, there is no explosion in the interval size as the mesh increases for integrals not containing gradients. This observation motivates the use of integrals without derivatives, whenever possible, to decrease the size of the interval enclosures.

The previous observation can be seen as the interval arithmetic analogue of the *a priori estimates*:

- The estimate  $\|\nabla u - \nabla u_h\|_L^2$  is of order  $h$ , while  $\|u - u_h\|_{L^2} = O(h^2)$ , assuming  $u \in H^2$ .
- Assuming triangles of uniform area, and enclosures for finite element variables  $u_h, v_h$  with uniform intervals with respect to  $h$ , the interval enclosure for  $\int_T \nabla u_h \cdot \nabla v_h$  has a radius  $O(h^{-2})$  bigger than the one for the enclosure for  $\int_T u_h v_h$ .

**The regular hexagon.** Computations for the regular hexagon for  $m \in \{200, 250, \dots, 600\}$  are shown in Figure 9. The certification of local minimality using the method proposed in the previous section succeeds again for  $m$  large enough: the certified lower bound for the smallest non-zero eigenvalue becomes positive. The validation procedure succeeds already for  $m = 380$ . On the other hand, compared to the regular pentagon, for  $m = 600$  the radius of the intervals for the Hessian eigenvalues computed using finite elements becomes larger than the *a priori estimate*. This shows that refining the mesh further will not improve the results.

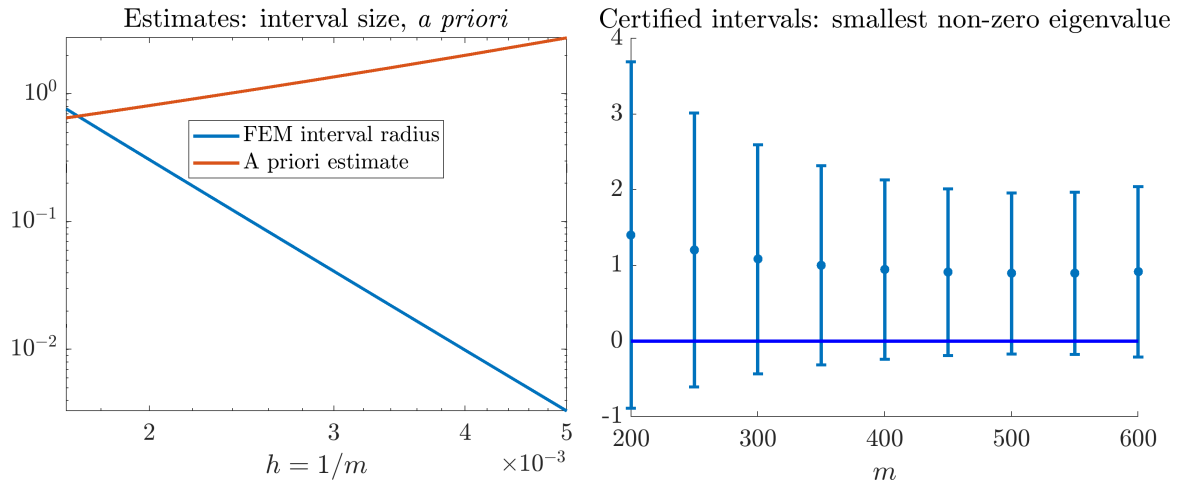


FIGURE 10. Computations for the **regular heptagon**. Left: evolution of the *a priori estimates* and the size of the intervals for  $\mu_j^h$  coming from the FEM computations. Right: certified intervals containing the smallest non-zero eigenvalue. The validation of local minimality does not succeed in this case.

$n$	Results from [7]			Results of this paper			
	$h$	DoF	Intervals	$h = 1/m$	m	DoF	Intervals
5	9.8e-4	2.5 million	✗	0.0040	250	156876	✓
6	4.2e-4	17 million	✗	0.0026	380	434341	✓
7	1.9e-4	97 million	✗	-	-	-	
8	1.35e-4	220 million	✗	-	-	-	

TABLE 1. Comparing the estimates in [7] with those of this paper. The improved error estimates allow us to use interval arithmetic in all computation, validating the local minimality for the regular pentagon and the regular hexagon.

**The regular heptagon.** Computations for the regular heptagon for  $m \in \{200, 250, \dots, 600\}$  are shown in Figure 10. The certification of local minimality fails in this case. The radius of the intervals for the Hessian eigenvalues computed using finite elements becomes larger than the *a priori* estimate for  $m = 600$ . Moreover, it can be observed that the radius of the enclosure of the smallest eigenvalue increases for  $m \geq 500$  even though the *a priori* estimate continues to decrease.

**Comparison with previous results.** Results provided in [7] used weaker *a priori* estimates. Meshes for which these estimates could validate the positivity of the non-zero eigenvalues of the Hessian matrix were too large to validate local minimality using intervals. The improvement of the constant in the interpolation constant (10) was also helpful to further decrease the computational cost. Table 1 shows a quick comparison of the results. Mesh sizes are chosen in each case such that  $h$  is as large as possible while *a priori* estimates guarantee the positivity of the Hessian eigenvalues.

**5.3. Implementation.** The Matlab code implementing the steps described above allowing to certify the local minimality for  $n \in \{5, 6\}$  can be found at the following repository:

<https://github.com/bbogo/PolyaSzego>

It requires a working installation of Intlab [42]. See Appendix B for more details.

**5.4. Conclusion, Perspectives.** The theoretical results and certified numerical simulations presented in this paper show the following result:

**Theorem 5.1.** *The regular  $n$ -gon is a local minimizer for the first Dirichlet-Laplace eigenvalue among  $n$ -gons with fixed area for  $n \in \{5, 6\}$ .*

Concerning perspectives for improving the results we mention the following:

- **A posteriori estimates.** Refining the error estimate for  $\|\nabla u_1 - \nabla u_{1,h}\|_{L^2(\mathbb{P}_n)}$  would directly influence the leading term for the error estimates discussed in Section 3.1. Having better estimates would allow the validation procedure to succeed for meshes with fewer triangles, where using interval arithmetics for finite element computations will lead to tighter interval enclosures. The results in [8] should be applied in this context.
- **Higher order elements.** The context of  $\mathbf{P}_2$  finite elements may be of interest since it gives improved error estimates. Results in [35] could help to adapt such techniques to our problem.
- **Spectral methods.** The method of particular solutions was used in [11], [10] to compute enclosures for the Dirichlet Laplacian eigenvalues and eigenfunctions on polygons. Similar results should be applied to problems (21), where the singular behavior should be taken into account in the choice of particular solutions. Methods described in [18] behave well in singular contexts. Investigating quantified error bounds for such methods is of interest concerning the results of this paper.

#### APPENDIX A. MORLEY ELEMENT AND INTERPOLATION CONSTANTS

Consider  $T$  a triangle in the plane with vertices  $x_1, x_2, x_3$  labeled in the positive trigonometric sense. Denote by  $\gamma_1, \gamma_2, \gamma_3$  the edges opposite to  $x_1, x_2, x_3$ , respectively, having positive orientation. By definition, the Lagrange  $\mathbf{P}_1$  interpolant  $\Pi(u)$  of a function  $u$  defined on a triangle  $T$ , is the affine function taking the same values as  $u$  at the vertices of  $T$ . Therefore  $\Pi(u) - u$  takes zero values at the vertices of  $T$ . Consider the space

$$V^2(T) = \{u \in H^2(T) : u(x_k) = 0, k = 1, 2, 3\}.$$

Note that this space is well defined since functions in  $H^2(T)$  are continuous, thus point values can be considered. In order to obtain the best constant in the inequality

$$(55) \quad \|\nabla(\Pi(u) - u)\|_{L^2} \leq C(T)|u|_{H^2},$$

where  $|u|_{H^2}^2 = \|\partial_{xx}u\|_{L^2(T)}^2 + 2\|\partial_{xy}u\|_{L^2(T)}^2 + \|\partial_{yy}u\|_{L^2(T)}^2$  the following Rayleigh quotient is considered:

$$C(T) = \sup_{u \in V^2(T) \setminus \{0\}} \frac{\|\nabla u\|_{L^2(T)}}{|u|_{H^2(T)}}.$$

This constant is well defined and an upper bound can be found through classical estimates involving the eigenvalues of the Laplacian and the Poincaré inequality. Indeed, assume for contradiction that for some sequence  $(u_n) \in H^2(T)$  we have  $\|\nabla u_n\|_{L^2(T)} = 1$  and  $|u_n|_{H^2(T)} \rightarrow 0$ . Let us denote  $\bar{u}_n = \frac{1}{|T|} \int_T u_n$ , so that we can apply the Poincaré-Wirtinger inequality in  $H^1(T)$  to the functions  $u_n - \bar{u}_n$ , hence their  $L^2$  norms are uniformly bounded. This means that the sequence  $(u_n - \bar{u}_n)_n$  is bounded in  $H^2$  and each function  $v_n := u_n - \bar{u}_n$  is constant on the vertices of  $T$ , while  $\|\nabla v_n\|_{L^2(T)} = 1$  and  $|v_n|_{H^2(T)} \rightarrow 0$ . We can assume that  $v_n$  converges weakly in  $H^2(T)$  to some function  $v$ . In particular, the convergence will also hold in  $C^{0,\alpha}(T)$ , for every  $\alpha \in (0, 1)$  and in  $H^1(T)$ -strongly. We get that  $|v|_{H^2(T)} = 0$ , so that  $v$  is affine. But,  $v$  is constant on the vertices of  $T$ , so that  $v$  is a constant function, in contradiction with the fact that  $\int_T \|\nabla v\|^2 = 1$ .

This constant is denoted  $C_4(T)$  in [29], where it is shown that

$$(56) \quad C(T) \leq \sqrt{\frac{n^2}{n^2 - 1}} C^{(n)}(T),$$

where  $C^{(n)}(T)$  is defined as follows:

- Given a triangle  $\tau$  with vertices  $p_k$  and edges  $\gamma_k$ ,  $i = 1, 2, 3$  and  $\varphi \in H^2(\tau)$  consider the interpolation operator  $\Pi_\tau^{(\beta)}$  (using the same notations as in [29]) such that:
  - a)  $\Pi_\tau^{(\beta)} \varphi \in \{a_1 x^2 + a_2 xy + a_3 y^2 + a_4 x + a_5 y + a_6 : a_i \in \mathbb{R}, i = 1, \dots, 6\}$ .
  - b)  $\Pi_\tau^{(\beta)} \varphi(p_k) = \varphi(p_k)$ ,  $k = 1, 2, 3$

$$c) \int_{\gamma_k} \nabla \Pi_{\tau}^{(\beta)} \varphi \cdot \mathbf{n} = \int_{\Gamma_k} \nabla \varphi \cdot \mathbf{n}, \quad k = 1, 2, 3.$$

In other words,  $\Pi_{\tau}^{(\beta)} \varphi$  is a second order polynomial in  $x, y$  taking the same values as  $\varphi$  at the vertices of  $\tau$  and having the same average normal flux across the edges  $\gamma_k$ ,  $k = 1, 2, 3$ .

This interpolation operator is also called the Morley interpolant. Since the gradient of a quadratic function is a linear function, the average of the normal derivative on an edge of  $\tau$  is proportional to the normal derivative at the midpoint of the side. Thus, one can take as degrees of freedom for the Morley interpolant the values at the vertices and the normal derivatives at the midpoints of the sides.

- Given  $n \geq 1$  define  $T' = \cup_{k=1}^{n^2} \tau_k$  the triangulation of  $T$  into  $n^2$  congruent triangles similar to  $T$  (splitting each side of  $T$  into  $n$  equal segments like in Figure 5). Given  $u \in H^2(T)$  consider the interpolant  $\Pi^{(\beta)} u$  such that

$$\Pi^{(\beta)} u|_{\tau_k} = \Pi_{\tau_k}^{(\beta)} u.$$

In other words, on each sub-triangle  $\tau_k$ ,  $\Pi^{(\beta)} u$  is a quadratic polynomial, taking the same values as  $u$  at the vertices of  $\tau_k$  and having the same normal flux as  $u$  across the sides of  $\tau_k$ .

- The constant  $C^{(n)}(T)$  is defined by

$$(57) \quad C^{(n)}(T) = \sup_{u \in V^2(T) \setminus 0} \frac{\|\nabla \Pi^{(\beta)} u\|_{L^2(T')}}{|\Pi^{(\beta)} u|_{H^2(T')}}.$$

The notations  $L^2(T')$  and  $H^2(T')$  simply mean that functions are in the corresponding spaces for each member of the triangulation  $T'$ . The functions may not be globally  $H^2$ , for example.

The estimate (56) states that the discrete eigenvalue obtained when using the Morley finite element on a triangulation of  $T$  using equal triangles gives an upper bound for the constant in the interpolation estimate (55). The proof of the estimate is a simple consequence of the identity

$$|\varphi|_{H^2(\tau)}^2 = |\Pi_{\tau}^{(\beta)} \varphi|_{H^2(\tau)}^2 + |\varphi - \Pi_{\tau}^{(\beta)} \varphi|_{H^2(\tau)}^2,$$

which follows from the definition of the Morley interpolant. In other words, the Morley interpolant decreases the  $H^2$  semi-norm on a triangle.

In order to have a tight estimate for the interpolation constant  $C(T)$  it is enough to have a certified inclusion for  $C^{(n)}(T)$ , obtained using interval arithmetic.

To achieve this, we show below an intuitive strategy for assembling the matrices associated to the Morley element, which can help compute  $C^{(n)}(T)$ . For alternative approaches one might consider [28]. The strategy is as follows:

- From the degrees of freedom for the Morley element (values at vertices, normal derivatives at midpoints of the edges) one can find the full gradient at the midpoints of the edges with an elementary observation regarding quadratic functions: if  $P$  is quadratic on  $T$  then  $P(p_{k+1}) - P(p_k) = 2|p_{k+1} - p_k| \partial_{\tau} P \left( \frac{p_{k+1} + p_k}{2} \right)$ .
- Gradient of a quadratic function on  $T$  is affine on  $T$ . Knowing gradients at midpoints gives the gradients at vertices.
- Knowing gradients at vertices as affine functions leads to easy assembly of the matrices involving first and second derivatives using classical ideas regarding  $\mathbf{P}_1$  finite elements.

In the following, the contribution of a triangle of the mesh to the assembly matrices is described. The procedure could be applied to a general mesh  $\Omega$  where the degrees of freedom corresponding to Morley finite elements correspond to values at vertices and values on edges (corresponding, for example, to half the normal derivatives across the edges).

Supposing that  $v_1, v_2, v_3$  are the values of an quadratic function  $p$  at the vertex of  $T$  and  $\psi_{12}, \psi_{31}, \psi_{23}$  are half the normal derivatives at the midpoints of the sides  $x_i x_j$  and  $\ell_{ij} = |x_i - x_j|$  then the components of the gradient of  $p$  at the vertices of  $T$  can be found through the matrix

product:

$$\begin{pmatrix} p_{1,x} \\ p_{1,y} \\ p_{2,x} \\ p_{2,y} \\ p_{3,x} \\ p_{3,y} \end{pmatrix} = \mathcal{M}_T \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \psi_{23} \\ \psi_{31} \\ \psi_{12} \end{pmatrix}$$

where

$$\mathcal{M}_T = \begin{pmatrix} -I_2 & I_2 & I_2 \\ I_2 & -I_2 & I_2 \\ I_2 & I_2 & -I_2 \end{pmatrix} \begin{pmatrix} \mathbf{M}_{23} & 0 & 0 \\ 0 & \mathbf{M}_{31} & 0 \\ 0 & 0 & \mathbf{M}_{12} \end{pmatrix} \begin{pmatrix} 0 & -\frac{1}{\ell_{23}} & \frac{1}{\ell_{23}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{23} & 0 & 0 \\ \frac{1}{\ell_{31}} & 0 & -\frac{1}{\ell_{23}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{31} & 0 \\ -\frac{1}{\ell_{12}} & \frac{1}{\ell_{12}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{12} \end{pmatrix}.$$

The unitary matrices  $\mathbf{M}_{ij} = \frac{1}{\ell_{ij}}[(x_j - x_i)(x_j - x_i)^\perp]$  rotate a coordinate system aligned with  $x_j - x_i$  to make it align with the one corresponding to the canonical basis. Tangential and normal derivatives of  $p$  at midpoints are transformed into partial derivatives  $\partial_x p, \partial_y p$ . The values  $\sigma_{ij} \in \{-1, 1\}$  are arbitrary (but fixed) sign choices for the flux on the edge  $[x_i, x_j]$  giving an orientation for the flux, opposite for the two triangles adjacent to this side.

Next, consider the permutation matrix  $P$  which regroups on the first three components the values of  $\partial_x p$  and on the last three the values of  $\partial_y p$ .

$$\begin{pmatrix} p_{1,x} \\ p_{2,x} \\ p_{3,x} \\ p_{1,y} \\ p_{2,y} \\ p_{3,y} \end{pmatrix} = P \mathcal{M}_T \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \psi_{23} \\ \psi_{31} \\ \psi_{12} \end{pmatrix}.$$

Next, consider the mass and rigidity matrices  $K_T, M_T$  for  $\mathbf{P}_1$  finite elements on the triangle  $T$ .

Denoting by  $\mathbf{v}$  a generic vector containing all degrees of freedom for the Morley elements (one per vertex and one per edge) we have the following. Assume the domain  $\Omega$  is triangulated with triangles  $\{T_i\}_{i=1}^N$ .

A) The matrix  $\mathcal{K}_{xx}$  which for a Morley function  $\tilde{u}$  with degrees of freedom  $\mathbf{v}$  computes

$$\mathbf{v}^T \mathcal{K}_{xx} \mathbf{v} = |\tilde{u}|_{H^2(T')}^2 = \sum_{i=1}^N \int_{T_i} (\partial_{xx} \tilde{u})^2 + 2(\partial_{xy} \tilde{u})^2 + (\partial_{yy} \tilde{u})^2 = \sum_{i=1}^N \|\nabla p_x\|_{L^2(T_i)}^2 + \|\nabla p_y\|_{L^2(T_i)}^2,$$

where  $p_x, p_y$  denote degree one polynomials corresponding to  $\partial_x \tilde{u}, \partial_y \tilde{u}$  on the triangles of the mesh, is given as follows. For each triangle in the mesh, the contribution corresponding to the degrees of freedom associated to the values and fluxes along the elements of the triangle is given by

$$\mathcal{M}_T^T P^T \begin{pmatrix} K_T & 0 \\ 0 & K_T \end{pmatrix} P \mathcal{M}_T,$$

where  $K_T$  is the rigidity matrix for  $\mathbf{P}_1$  elements on the triangle  $T$ .

B) The matrix  $\mathcal{M}_{xx}$  which for a Morley function  $\tilde{u}$  with degrees of freedom  $\mathbf{v}$  computes

$$\mathbf{v}^T \mathcal{M}_{xx} \mathbf{v} = \|\nabla \tilde{u}\|_{L^2(T')}^2 = \sum_{i=1}^N \int_{T_i} (\partial_x \tilde{u})^2 + (\partial_y \tilde{u})^2 = \sum_{i=1}^N \|p_x\|_{L^2(T_i)}^2 + \|p_y\|_{L^2(T_i)}^2$$

is given as follows. Then for each triangle in the mesh, the contribution corresponding to the degrees of freedom associated to the values and fluxes along the elements of the triangle is given by

$$\mathcal{M}_T^T P^T \begin{pmatrix} M_T & 0 \\ 0 & M_T \end{pmatrix} P \mathcal{M}_T,$$



$\theta$	$2\pi/5$	$2\pi/6$	$2\pi/7$	$2\pi/8$	$2\pi/9$	$2\pi/10$
$C(T)$	0.3697	0.3200	0.3146	0.3107	0.3104	0.3128

TABLE 2. Certified upper bounds for  $\mathbf{P}_1$  interpolation constants on isosceles triangles corresponding to a slice of a regular  $n$ -gon for  $n \in \{5, 6, 7, 8, 9, 10\}$ .

where  $M_T$  is the mass matrix for  $\mathbf{P}_1$  elements on the triangle  $T$ .

The assembly procedure described above may not be the most efficient, but it has the advantage of being clear and explicit. Moreover, for the triangulation of the original triangle  $T$  into  $n^2$  congruent triangles the assembly can be rendered explicit and evaluated with INTLAB with a tight interval enclosure. The constant  $C^{(n)}(T)$  defined in (57) is given by

$$C^{(n)}(T) = \sqrt{1/\rho_1}, \text{ where } \rho_1 = \inf_{\mathbf{v}, \mathbf{v}(x_k)=0} \frac{\mathbf{v}^T \mathcal{K}_{xx} \mathbf{v}}{\mathbf{v}^T \mathcal{M}_{xx} \mathbf{v}},$$

which means that  $\rho_1$  is the smallest generalized eigenvalue for the matrices  $(\mathcal{K}_{xx}, \mathcal{M}_{xx})$ , taking into account the Dirichlet boundary conditions  $\mathbf{v}(x_1) = \mathbf{v}(x_2) = \mathbf{v}(x_3) = 0$ . To impose these boundary condition, the corresponding lines and columns are eliminated from  $(\mathcal{K}_{xx}, \mathcal{M}_{xx})$  obtaining  $(\mathcal{K}_{xx}^0, \mathcal{M}_{xx}^0)$ .

Given  $\bar{\rho}$ , the smallest generalized eigenvalue of  $(\mathcal{K}_{xx}^0, \mathcal{M}_{xx}^0)$  approximated using a floating point computation, the routine `isspd` from INTLAB verifies if the matrix

$$\mathcal{K}_{xx}^0 - (\bar{\rho} - \varepsilon) \mathcal{M}_{xx}^0$$

is symmetric positive definite for some  $\varepsilon > 0$  (for example  $\varepsilon = 10^{-6}$ ). If the numerical validation succeeds, an upper bound for  $C^{(n)}(T)$  is obtained, which gives an upper bound for  $C(T)$  from (56).

Observing that the constant  $C(T)$  defined in (55) scales linearly with the size of  $T$ , it is enough to compute it for a triangle with vertices  $x_1 = (0, 0)$ ,  $x_2 = (1, 0)$ ,  $x_3 = (a, b)$ , verifying  $b > 0$ . In this case, denoting with  $\psi_1, \psi_2, \psi_3$  the  $\mathbf{P}_1$  finite element functions on  $T$  it follows immediately that

$$\begin{aligned} \partial_x \psi_1 &= -1 & \partial_x \psi_2 &= 1 & \partial_x \psi_3 &= 0 \\ \partial_y \psi_1 &= \frac{a-1}{b} & \partial_y \psi_2 &= -\frac{a}{b} & \partial_y \psi_3 &= \frac{1}{b}. \end{aligned}$$

The expression of the rigidity matrix  $K_T$  on the triangle  $T$  follows, observing that the area of  $T$  is equal to  $b/2$ . For the mass matrix one obtains

$$M_T = |T| \begin{pmatrix} 1/6 & 1/12 & 1/12 \\ 1/12 & 1/6 & 1/12 \\ 1/12 & 1/12 & 1/6 \end{pmatrix}.$$

The code which for a given pair  $(a, b)$ ,  $b > 0$  gives a certified upper bound for the optimal interpolation constant  $C(T)$  can be found at the following repository:

<https://github.com/bbogo/PolyaSzego/tree/main/Morley>

The code requires a working installation of INTLAB [42]. It receives as inputs the vertex coordinates  $(a, b)$  and the number of segments  $m$  in which every side is divided for constructing the mesh. It outputs a certified upper bound for  $C(T)$  if the validation method succeeds. Otherwise, it outputs a message indicating that a lower bound was not found.

All results described in [29] can be verified using the code proposed above. The estimates relevant to the contents of this paper correspond to isosceles triangles with edge length 1 and central angle  $2\pi/n$ , giving  $a = \cos(2\pi/n)$ ,  $b = \sin(2\pi/n)$ ,  $n \in \{5, 6\}$ . The corresponding constants are summarized in Table 2, where dependence on  $C(T)$  the central angle of the regular polygon is emphasized.

## APPENDIX B. DESCRIPTION OF THE CODE

The Matlab code proving the local minimality of the regular  $n$ -gon for problem (2) can be found at the following repository:

<https://github.com/bbogo/PolyaSzego>

It uses interval arithmetic and requires a working installation of INTLAB [42]. Details regarding the implementation can be found in the code comments. For the sake of completeness, a description of the functions and of the validation process is described below. Before running the code, the subfolder `./Tools/` should be added to the Matlab path.

**a) Validation for the first two eigenvalues and the first eigenfunction.** The implementation is based on the discussion in Section 5. The first main function is `PolyaHessInterval.m`. For running the code corresponding to  $n = 5$  and  $m = 250$  division points on the ray `[oa0]` type:

```
>> PolyaHessInterval(5,250)
```

The code constructs two meshes, one for the slice  $T_+$  (see Figure 1), where the first eigenfunction is computed and validated, exploiting the symmetry. The first eigenpair is either validated using `verifyeig` in INTLAB or with a residual estimate. A second mesh of the full polygon is used for computing and validating the second eigenvalue. The validation uses the residual estimate from Proposition 4.4. At the end of the validation process the eigenvalue and eigenfunction enclosures are saved to a file to for a later use.

The matrices needed in the computation are assembled exactly. Only the incidence relations in the mesh are used in the assembly, all other aspects are computed analytically.

**b) Validating the local minimality of the regular  $n$ -gon.** The remaining of the validation procedure can be achieved using the function `PolyaHessIntervalU.m`. For running the code corresponding to  $n = 5$  and  $m = 250$  division points on the ray `[oa0]` type:

```
>> PolyaHessIntervalU(5,250)
```

```
Number of positive eigenvalues = 6
Proof of local minimality succeeded!
Degrees of Freedom (full mesh) 156876
intval ans =
[ -2.8083,  2.5145]
[ -2.8083,  2.5145]
[ -0.0001,  0.0001]
[ -0.0001,  0.0001]
[  0.0547,  5.3775]
[  0.0547,  5.3775]
[  2.6266, 12.6594]
[  2.6266, 12.6594]
[  8.8146, 18.8474]
[  8.8146, 18.8474]
```

LISTING 1. Output of the numerical validation code for  $n = 5$ ,  $m = 250$ . Six out of the ten eigenvalues are positive, proving the local minimality of the regular pentagon.

The linear systems associated to the material derivatives (49), (50) are solved and validated and an interval enclosure for the solution is found. Conjugate gradient is used to obtain a floating point solution, which is validated using Lemma 4.7. Afterwards, the eigenvalues of the Hessian matrix are evaluated from the finite element results, following Theorem 2.1. Next, all elements involving the *a priori* estimates in Sections 2, 3 quantifying the error between continuous solutions of the PDEs involved and the finite element counterparts are computed.

The code outputs a sequence of interval enclosures for the eigenvalues of the Hessian matrix. If  $2n - 4$  of these intervals are included in  $[0, +\infty)$  then the validation procedure succeeds and local minimality is proved (see [7, Section 4]). The local minimality validation succeeds for  $n = 5$  (taking  $m = 250$ ) and  $n = 6$ , (taking  $m = 380$ ). An example of output for  $n = 5$ ,  $m = 250$  is shown in Listing 1.

## APPENDIX C. FAST VALIDATION OF EIGENVALUES FOR LARGE SPARSE MATRICES

Applying directly the routine `verifyeig` from Intlab [42] to generalized eigenvalue problems of the form

$$K_0 u = \lambda M_0 u$$

leads to significant computational costs when  $K_0, M_0$  are of large sizes. Looking at the code provided with Intlab, it seems that the large cost comes from the inversion of a matrix. While it is true that multiple linear systems need to be solved using the matrix in question with various right hand sides, for large problems it is more efficient to consider linear systems instead. The function `verifylss` in Intlab allows to solve linear sparse systems of large sizes, therefore, this routine is used whenever a linear system is solved. The modifications are underlined in the Listings 2, 3. For the generalized eigenvalue problem the matrix `B` and its midpoint `midB` appear in the same places as in the original `verifyeig` function.

```

.....
R = midA - lambda*speye(n);
R(:,v) = -xs;
R = inv( R ); % matrix inversion
C = A - intval(lambda)*speye(n);
Z = - R * ( C * xs );
C(:,v) = -xs;
C = speye(n) - R * C;
Y = Z;
Eps = 0.1*mag(Y)*hull(-1,1) + midrad(0,
    realmin);
m = 0;
mmax = min( 15 * ( sum(sum(mag(Z(v,:))
    >.1)) + 1 ) , 20 );
ready = 0;
while ( ~ready ) && ( m<mmax ) && ( ~any(
    isnan(Y(:)) ) )
    m = m+1;
    X = Y + Eps;
    XX = X;
    XX(v,:) = 0;
    Y = Z + C*X + R*(XX*X(v,:));
    ready = all(all(in0(Y,X)));
end
.....

```

LISTING 2. The original `verifyeig` function in Intlab.

```

.....
R = midA - lambda*speye(n);
R(:,v) = -xs;
% without matrix inversion
C = A - intval(lambda)*speye(n);
Z = -verifylss(R,(C*xs));
C(:,v) = -xs;
Y = Z;
Eps = 0.1*mag(Y)*hull(-1,1) + midrad(0,
    realmin);
m = 0;
mmax = min( 15 * ( sum(sum(mag(Z(v,:))
    >.1)) + 1 ) , 20 );
ready = 0;
while ( ~ready ) && ( m<mmax ) && ( ~any(
    isnan(Y(:)) ) )
    m = m+1;
    X = Y + Eps;
    XX = X;
    XX(v,:) = 0;
    Y = Z+verifylss(R,(R-C)*X+XX*X(v,:));
    ready = all(all(in0(Y,X)));
end
.....

```

LISTING 3. Replacing matrix inversion with linear systems.

The modified `verifyeig` is applied to randomly generated matrices with sizes up to  $1000 \times 1000$  and compared with the original version. The results produced are identical (intervals have the same center and same radius). For full matrices the proposed modification is not faster than the original version for the cases tested.

Nevertheless, the situation changes when the matrices involved are sparse and come from discretization of finite element problems like (15). The generalized eigenvalue problem for a triangular slice meshed with congruent triangles obtained by dividing each one of the sides into  $m$  segments ( $m^2$  triangles) is solved with the original version of `verifyeig` and the modified version. The computations are realized for  $m \in \{30, 40, \dots, 90, 100\}$  and the difference in computation times is shown in Figure 11. The speed gain from removing the matrix inverse is significant, since Intlab uses a sparse version of the linear system verification routine `verifylss`. The verification with the modified version was tested successfully for triangle division parameter  $m$  up to 450, corresponding to matrix size of roughly  $10^5 \times 10^5$ .

For even finer meshes, further modifications are provided to the `verifyeig` routine. An initial floating point system used is replaced with an iterative method. This modification was tested

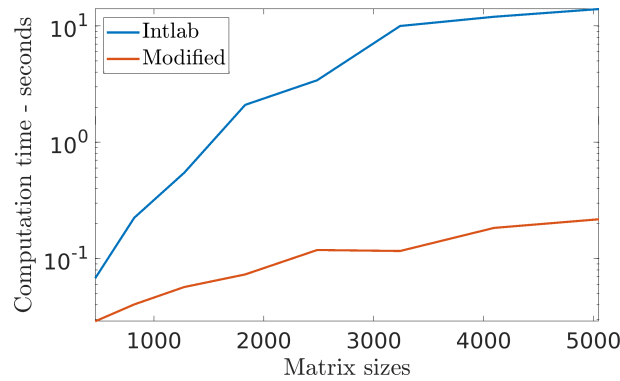


FIGURE 11. Comparison of computation times for the eigenvalue verification routine `verifyeig` in Intlab and the modified version, applied to sparse matrices coming from the discretization of a Laplace PDE. The significant time improvement allows to validate eigenvalues and eigenvectors for matrices of sizes up to  $10^5 \times 10^5$ .

for the mesh parameter  $m$  up to 600 corresponding to matrices of size  $180901 \times 180901$ . These modified versions of `verifyeig` could be of potential interest for other applications involving eigenvalues of large sparse matrices.

**Acknowledgements:** The authors thank Maxime Breden, Xuefeng Liu and Siegfried M. Rump for valuable discussions and advice regarding the validation procedure using interval arithmetic.

#### REFERENCES

- [1] G. Alessandrini. Nodal lines of eigenfunctions of the fixed membrane problem in general convex domains. *Comment. Math. Helv.*, 69(1):142–154, 1994.
- [2] P. Antunes and P. Freitas. New bounds for the principal Dirichlet eigenvalue of planar regions. *Experiment. Math.*, 15(3):333–342, 2006.
- [3] I. Babuška and V. Nistor. Interior numerical approximation of boundary value problems with a distributional data. *Numer. Methods Partial Differential Equations*, 22(1):79–113, 2006.
- [4] I. Babuška and V. Nistor. Boundary value problems in spaces of distributions on smooth and polygonal domains. *J. Comput. Appl. Math.*, 218(1):137–148, 2008.
- [5] A. H. Barnett, A. Hassell, and M. Tacy. Comparable upper and lower bounds for boundary values of Neumann eigenfunctions and tight inclusion of eigenvalues. *Duke Math. J.*, 167(16):3059–3114, 2018.
- [6] B. Bogosel. *Shape optimization and spectral problems*. PhD thesis, Universite Grenoble Alpes, 2015.
- [7] B. Bogosel and D. Bucur. On the polygonal Faber-Krahn inequality. *Journal de l'École polytechnique — Mathématiques*, 11:19–105, Dec. 2023.
- [8] E. Cancès, G. Dusson, Y. Maday, B. Stamm, and M. Vohralík. Guaranteed and robust a posteriori bounds for laplace eigenvalues and eigenvectors: Conforming approximations. *SIAM Journal on Numerical Analysis*, 55(5):2228–2254, Jan. 2017.
- [9] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].
- [10] J. Dahne, J. Gómez-Serrano, and K. Hou. A counterexample to Payne’s nodal line conjecture with few holes. *Commun. Nonlinear Sci. Numer. Simul.*, 103:Paper No. 105957, 13, 2021.
- [11] J. Dahne and B. Salvy. Computation of tight enclosures for Laplacian eigenvalues. *SIAM J. Sci. Comput.*, 42(5):A3210–A3232, 2020.
- [12] D. Daners. Krahn’s proof of the Rayleigh conjecture revisited. *Arch. Math. (Basel)*, 96(2):187–199, 2011.
- [13] M. C. Delfour and J. P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization, Second Edition*. Society for Industrial and Applied Mathematics, Jan. 2011.
- [14] S. Dominguez, N. Nigam, and B. Shahriari. A combined finite element and Bayesian optimization framework for shape optimization in spectral geometry. *Comput. Math. Appl.*, 74(11):2874–2896, 2017.
- [15] I. Fragalà and B. Velichkov. Serrin-type theorems for triangles. *Proc. Amer. Math. Soc.*, 147(4):1615–1626, 2019.
- [16] G. H. Golub, C. Greif, and J. M. Varah. An algebraic analysis of a block diagonal preconditioner for saddle point systems. *SIAM Journal on Matrix Analysis and Applications*, 27(3):779–792, Jan. 2005.
- [17] J. Gómez-Serrano and G. Orriols. Any three eigenvalues do not determine a triangle. *J. Differential Equations*, 275:920–938, 2021.

- [18] A. Gopal and L. N. Trefethen. Solving Laplace problems with corner singularities via rational functions. *SIAM J. Numer. Anal.*, 57(5):2074–2094, 2019.
- [19] D. S. Grebenkov and B.-T. Nguyen. Geometrical structure of Laplacian eigenfunctions. *SIAM Rev.*, 55(4):601–667, 2013.
- [20] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.
- [21] A. Henrot. *Extremum problems for eigenvalues of elliptic operators*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2006.
- [22] A. Henrot, editor. *Shape optimization and spectral theory*. De Gruyter Open, Warsaw, 2017.
- [23] A. Henrot and M. Pierre. *Shape variation and optimization*, volume 28 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2018. A geometrical analysis, English version of the French publication [MR2512810] with additions and updates.
- [24] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.
- [25] E. Indrei. On the first eigenvalue of the Laplacian for polygons. *Journal of Mathematical Physics*, 65(4), Apr. 2024.
- [26] L. Kamenski, W. Huang, and H. Xu. Conditioning of finite element equations with arbitrary anisotropic meshes. *Math. Comp.*, 83(289):2187–2211, 2014.
- [27] T. Kimura and X. Chen. Validated solutions of saddle point linear systems. *SIAM Journal on Matrix Analysis and Applications*, 30(4):1697–1708, 2009.
- [28] R. C. Kirby. A general approach to transforming finite elements. *SMAI J. Comput. Math.*, 4:197–224, 2018.
- [29] K. Kobayashi. On the interpolation constants over triangular elements. In *Applications of mathematics 2015*, pages 110–124. Czech. Acad. Sci., Prague, 2015.
- [30] J. Larsen-Scott and J. Clutterbuck. Local spectral optimisation for robin problems with negative boundary parameter on quadrilaterals. *Journal of Mathematical Physics*, 65(3), Mar. 2024.
- [31] A. Laurain. Distributed and boundary expressions of first and second order shape derivatives in nonsmooth domains. *J. Math. Pures Appl. (9)*, 134:328–368, 2020.
- [32] Q. Li and X. Liu. Explicit finite element error estimates for nonhomogeneous Neumann problems. *Appl. Math.*, 63(3):367–379, 2018.
- [33] X. Liu and S. Oishi. Verified eigenvalue evaluation for the Laplacian over polygonal domains of arbitrary shape. *SIAM J. Numer. Anal.*, 51(3):1634–1654, 2013.
- [34] X. Liu and S. Oishi. Verified eigenvalue evaluation for the Laplacian over polygonal domains of arbitrary shape. *SIAM J. Numer. Anal.*, 51(3):1634–1654, 2013.
- [35] X. Liu and C. You. Explicit bound for quadratic Lagrange interpolation constant on triangular finite elements. *Appl. Math. Comput.*, 319:693–701, 2018.
- [36] G. Mayer. Result verification for eigenvectors and eigenvalues. In J. Herzberger, editor, *Topics in Validated Computations*, chapter 8, pages 209–276. Elsevier Science B.V., 1994.
- [37] Z. Mghazli. Regularity of an elliptic problem with mixed Dirichlet-Robin boundary conditions in a polygonal domain. *Calcolo*, 29(3-4):241–267 (1993), 1992.
- [38] N. Nigam, B. o. Siudeja, and B. Young. A proof via finite elements for Schiffer’s conjecture on a regular pentagon. *Found. Comput. Math.*, 20(6):1475–1504, 2020.
- [39] C. Nitsch. On the first Dirichlet Laplacian eigenvalue of regular polygons. *Kodai Math. J.*, 37(3):595–607, 2014.
- [40] H. C. Pak and Y. J. Park. Sharp trace inequalities on fractional Sobolev spaces. *Math. Nachr.*, 284(5-6):761–763, 2011.
- [41] G. Pólya and G. Szegő. *Isoperimetric Inequalities in Mathematical Physics*. Annals of Mathematics Studies, no. 27. Princeton University Press, Princeton, N. J., 1951.
- [42] S. M. Rump. Verification methods: rigorous results using floating-point arithmetic. *Acta Numer.*, 19:287–449, 2010.
- [43] J. Sokolowski and J.-P. Zolesio. *Introduction to Shape Optimization*. Springer Berlin Heidelberg, 1992.

(Beniamin Bogosel) CMAP, CNRS, ÉCOLE POLYTECHNIQUE, INSTITUT POLYTECHNIQUE DE PARIS, 91120 PALAISEAU, FRANCE

*Email address:* beniamin.bogosel@polytechnique.edu

(Dorin Bucur) LABORATOIRE DE MATHÉMATIQUES UMR CNRS 5127, UNIVERSITÉ DE SAVOIE, CAMPUS SCIENTIFIQUE, 73376 LE-BOURGET-DU-LAC (FRANCE)

*Email address:* dorin.bucur@univ-savoie.fr