

## TWO-SIDED BOUNDS FOR EIGENVALUES OF DIFFERENTIAL OPERATORS WITH APPLICATIONS TO FRIEDRICHS, POINCARÉ, TRACE, AND SIMILAR CONSTANTS\*

IVANA ŠEBESTOVÁ<sup>†</sup> AND TOMÁŠ VEJCHODSKÝ<sup>‡</sup>

**Abstract.** We present a general numerical method for computing guaranteed two-sided bounds for principal eigenvalues of symmetric linear elliptic differential operators. The approach is based on the Galerkin method, on the method of a priori–a posteriori inequalities, and on a complementarity technique. The two-sided bounds are formulated in a general Hilbert space setting and as a byproduct we prove an abstract inequality of Friedrichs–Poincaré type. The abstract results are then applied to Friedrichs, Poincaré, and trace inequalities and fully computable two-sided bounds on the optimal constants in these inequalities are obtained. Accuracy of the method is illustrated in numerical examples.

**Key words.** bounds on spectrum, a posteriori error estimate, optimal constant, Friedrichs inequality, Poincaré inequality, trace inequality, Hilbert space

**AMS subject classifications.** 35P15, 35J15, 65N25, 65N30

**DOI.** 10.1137/13091467X

**1. Introduction.** Eigenvalue problems for differential operators have attracted a lot of attention as they have many applications. These include the dynamic analysis of mechanical systems [3, 15, 26], linear stability of flows in fluid mechanics [24], and electronic band structure calculations [30]. In this paper, we concentrate on guaranteed two-sided bounds of the principal (smallest) eigenvalue of symmetric linear elliptic operators. The standard Galerkin method for solution of eigenproblems is efficient and its convergence and other properties are well analyzed [5, 6, 11]. It is also well known for providing upper bounds on eigenvalues. However, in many applications a reliable *lower bound* of the smallest eigenvalue is the key piece of information and, unfortunately, the Galerkin method cannot provide it.

The question of lower bounds on the smallest eigenvalue has been studied for several decades. For example, see [31], where the second-order elliptic eigenvalue problems with Dirichlet boundary conditions are considered. Another technique that gives the lower bounds not only for the first eigenvalue is the method of intermediate problems. It is based on finding a base problem and subsequently introducing intermediate problems that give lower bounds for eigenvalues of the original problem and at the same time can be resolved explicitly; see, for example, [8, 9, 41]. A survey of this technique can be found in [16].

Nonconforming methods have been used for computing lower bounds on eigenvalues; see, for example, [4, 27, 29, 32, 42]. However, these lower bounds are valid asymptotically only and hence these methods *do not guarantee* that the computed

---

\*Received by the editors March 28, 2013; accepted for publication (in revised form) November 4, 2013; published electronically February 6, 2014.

<http://www.siam.org/journals/sinum/52-1/91467.html>

<sup>†</sup>Department of Numerical Mathematics, Faculty of Mathematics and Physics, Charles University in Prague, CZ-186 75 Praha 8, Czech Republic (ivasebestova@seznam.cz). This author was supported by the project MathMAC–University Center for Mathematical Modeling, Applied Analysis and Computational Mathematics of the Charles University in Prague.

<sup>‡</sup>Institute of Mathematics, Academy of Sciences, CZ-115 67 Praha 1, Czech Republic (vejchod@math.cas.cz). This author was supported by RVO 67985840.

approximation is really below the exact value. Recently, sufficient conditions for producing lower bounds for eigenvalues of symmetric elliptic operators by nonconforming methods have been provided in [19]. The described technique is based on satisfying the saturation assumption and on the condition that the local approximation property of the underlying finite element space is better than its global continuity property. It is proved that the second condition is met by the most commonly used nonconforming methods such as the Wilson element, the linear nonconforming Crouzeix–Raviart element, the enriched nonconforming rotated  $Q_1$  element, the Morley element, the Adini element, and the Morley–Wang–Xu element. The saturation assumption is proved for the Morley–Wang–Xu element, the enriched Crouzeix–Raviart element, and the Wilson element. Furthermore, new nonconforming methods satisfying these properties are proposed. However, no numerical experiments are presented.

Further, let us point out a recent result [28], where two-sided a priori bounds for the discretization error of eigenvectors are given.

The method of *a priori–a posteriori inequalities* that can be used for computation of lower bounds on eigenvalues was described and published in [22, 23, 35]. However, in these original publications  $C^2$ -continuous test and trial functions have been used in order to compute the actual lower bound. These functions are difficult to work with and therefore we couple the original method of a priori–a posteriori inequalities with the complementarity technique, where a certain flux function has to be reconstructed; see, e.g., [2, 14, 18, 33, 38, 40]. This flux reconstruction can be done in many ways. We choose the most straightforward approach that can be handled by the standard Raviart–Thomas–Nédélec finite element method.

Moreover, we generalize the original method of a priori–a posteriori inequalities to the case of a compact operator between a pair of Hilbert spaces. This generalization is especially useful for computing two-sided bounds of the optimal constant in Friedrichs, Poincaré, trace, and similar inequalities. It is based on the fact that the optimal constant in these inequalities is inversely proportional to the square root of the smallest eigenvalue of the corresponding symmetric linear elliptic partial differential operator.

Further, the generalization we have made enables us to set up an abstract framework in the Hilbert space setting. The abstract results then easily apply to symmetric linear elliptic partial differential operators and consequently to the optimal constant in the inequalities of Friedrichs–Poincaré type. Furthermore, as a byproduct of the abstract setting, we obtain a simple proof of the validity of an abstract inequality of Friedrichs–Poincaré type in the Hilbert space setting. The particular choices of the pair of Hilbert spaces, corresponding scalar products, and the compact operator then naturally yield the validity of Friedrichs, Poincaré, trace, Korn’s, and other inequalities.

The main motivation for our interest in two-sided bounds of the constants in Friedrichs, Poincaré, trace, and similar inequalities stems from the need of these bounds in a posteriori error estimation for numerical solutions of partial differential equations. In particular, the existing guaranteed upper bounds on the energy norm of the error utilize a kind of complementarity technique; see, e.g., [2, 12, 20, 33, 38, 40]. Estimates of this kind contain constants from Friedrichs, Poincaré, trace, and similar inequalities. Optimal values of these constants are often unknown and therefore suitable approximations have to be used in the error estimates. These approximations have to provide upper bounds on these constants in order to guarantee that the total error estimator is an upper bound on the error. Moreover, they have to be accurate due to the accuracy and efficiency of the error estimates.

The method presented in this paper provides accurate upper bounds on these constants. In addition, this method naturally considers the dependence of the optimal constants on the equation coefficients and on the boundary conditions. This dependence can be strong [39] and its capturing might be crucial for accuracy and robustness of a posteriori error bounds.

The rest of the paper is organized as follows. Section 2 introduces a general variational eigenvalue problem in the Hilbert space setting. It uses the spectral theory of compact operators to prove several properties of this eigenvalue problem including the existence of the principal eigenvalue. In section 3 we naturally prove the abstract inequality of Friedrichs–Poincaré type and show the relation between the optimal constant and the principal eigenvalue. Further, we briefly describe the Galerkin method that yields an upper bound on the principal eigenvalue and concentrate on the method of a priori–a posteriori inequalities and on an abstract complementarity result leading to a lower bound on the principal eigenvalue. Sections 4–6 apply the abstract results to the case of Friedrichs, Poincaré, and trace inequalities, and fully computable two-sided bounds on the optimal constants in these inequalities are obtained. Presented numerical experiments illustrate accuracy of the method and dependence of the optimal constants on a nonhomogeneous diffusion parameter. Finally, section 7 draws the conclusions.

**2. Variational eigenvalue problem in the Hilbert space setting.** Let  $V$  and  $H$  be two real Hilbert spaces with scalar products  $(\cdot, \cdot)_V$  and  $(\cdot, \cdot)_H$ , respectively. The norms induced by these scalar products are denoted by  $\|\cdot\|_V$  and  $\|\cdot\|_H$ . Further, let  $\gamma : V \rightarrow H$  be a continuous, linear, and *compact* operator. The center of our interest is the following eigenvalue problem. Find  $\lambda_i \in \mathbb{R}$ ,  $u_i \in V$ ,  $u_i \neq 0$  such that

$$(2.1) \quad (u_i, v)_V = \lambda_i (\gamma u_i, \gamma v)_H \quad \forall v \in V.$$

First, let us show that eigenvalues  $\lambda_i$  of (2.1) are positive.

LEMMA 2.1. *If  $u_i \in V$  is an eigenvector corresponding to an eigenvalue  $\lambda_i$  of (2.1), then  $\gamma u_i \neq 0$  and  $\lambda_i > 0$ .*

*Proof.* Since  $u_i \neq 0$ , we have by (2.1) that  $0 \neq \|u_i\|_V^2 = \lambda_i \|\gamma u_i\|_H^2$ . Thus,  $\gamma u_i \neq 0$  and  $\lambda_i$  has to be positive.  $\square$

Below we show that eigenvalues  $\lambda_i$  and eigenvectors  $u_i$  of (2.1) correspond to eigenvalues and eigenvectors of a compact operator, respectively. Consequently, these eigenvalues form a countable sequence that can be ordered as  $\lambda_1 \leq \lambda_2 \leq \dots$ . To show this correspondence, we define a *solution operator*  $S : H \rightarrow V$ . If  $f \in H$  is arbitrary, then the mapping  $v \mapsto (f, \gamma v)_H$  is a continuous linear form on  $V$  and, hence, the Riesz representation theorem yields existence of a unique element  $Sf \in V$  such that

$$(2.2) \quad (Sf, v)_V = (f, \gamma v)_H \quad \forall v \in V.$$

Consequently, the solution operator  $S$  is linear and continuous.

The composition of operators  $S$  and  $\gamma$  is a linear, continuous, and compact operator  $S\gamma : V \rightarrow V$ ; see [34, Theorem 4.18(f)]. In addition, this operator is self-adjoint, because definition (2.2) yields

$$(2.3) \quad (S\gamma u, v)_V = (\gamma u, \gamma v)_H = (\gamma v, \gamma u)_H = (S\gamma v, u)_V = (u, S\gamma v)_V.$$

Therefore, we can use the Hilbert–Schmidt spectral theorem for  $S\gamma$  (see [17, Theorem 4, Chapter II, section 3]) and obtain that  $V$  can be decomposed into a direct sum of two subspaces,

$$(2.4) \quad V = \mathcal{M} \oplus \ker(S\gamma),$$

where  $\ker(S\gamma) = \{v \in V : S\gamma v = 0\}$  is the kernel of  $S\gamma$  and  $\mathcal{M}$  is generated by all eigenvectors of the operator  $S\gamma$  corresponding to nonzero eigenvalues. Let us recall that  $u_i \in V$ ,  $u_i \neq 0$  is an eigenvector of  $S\gamma$  corresponding to an eigenvalue  $\mu_i \in \mathbb{R}$  if

$$(2.5) \quad S\gamma u_i = \mu_i u_i.$$

Furthermore, the Hilbert–Schmidt spectral theorem implies that the system  $u_1, u_2, \dots$  of eigenvectors corresponding to nonzero eigenvalues of (2.5) is countable and orthogonal in  $V$ . A simple consequence of (2.2) and (2.5) is the orthogonality of images  $\gamma u_i$ ,  $i = 1, 2, \dots$ , in  $H$ . In this paper we consider the following normalization of these eigenvectors:

$$(2.6) \quad (\gamma u_i, \gamma u_j)_H = \delta_{ij} \quad \forall i, j = 1, 2, \dots,$$

where  $\delta_{ij}$  stands for the Kronecker’s delta.

Now, let us observe that eigenproblems (2.1) and (2.5) correspond to each other and, therefore, the spectral properties of the compact operator  $S\gamma$  translate to the properties of the eigenproblem (2.1).

LEMMA 2.2. *Considering the above setting, the following statements hold true:*

1. *Number  $\lambda_i \in \mathbb{R}$  is an eigenvalue corresponding to the eigenvector  $u_i \in V$  of (2.1) if and only if  $\mu_i = 1/\lambda_i$  is a nonzero eigenvalue corresponding to the eigenvector  $u_i$  of the operator  $S\gamma$ ; see (2.5).*
2. *The number of eigenvalues  $\lambda_i$  of (2.1) such that  $\lambda_i \leq M$  is finite for any  $M > 0$ .*
3. *The value  $\lambda_1 = \inf_{u \in V, u \neq 0} \|u\|_V^2 / \|\gamma u\|_H^2$  is the smallest eigenvalue of (2.1).*

*Proof.* 1. Definition (2.2) yields identity  $(\gamma u_i, \gamma v)_H = (S\gamma u_i, v)_V$  for all  $v \in V$ . Hence, the equality (2.1) can be rewritten as  $(u_i, v)_V = \lambda_i (S\gamma u_i, v)_V$ , which is equivalent to (2.5) with  $\mu_i = 1/\lambda_i$  provided that  $\lambda_i \neq 0$  and  $\mu_i \neq 0$ . Since Lemma 2.1 guarantees  $\lambda_i > 0$  for all  $i = 1, 2, \dots$ , the only condition is  $\mu_i \neq 0$ .

2. If we denote the spectrum of  $S\gamma$  by  $\sigma(S\gamma)$ , then the compactness of  $S\gamma$  implies that the set  $[\varepsilon, \infty) \cap \sigma(S\gamma)$  is finite for any  $\varepsilon > 0$ ; see [34, Theorem 4.24(b)]. The statement follows immediately from the fact that  $\lambda_i = 1/\mu_i$  for  $\mu_i \neq 0$ .

3. Since  $S\gamma$  is self-adjoint (see (2.3)), the Courant–Fischer–Weyl min-max principle (see, e.g., [36]) implies that

$$\mu_1 = \sup\{(S\gamma v, v)_V : \|v\|_V = 1\} = \sup_{v \in V, v \neq 0} \frac{(S\gamma v, v)_V}{\|v\|_V^2} = \sup_{v \in V, v \neq 0} \frac{\|\gamma v\|_H^2}{\|v\|_V^2}$$

is finite and it is the largest eigenvalue of the operator  $S\gamma$ . Consequently,

$$(2.7) \quad \lambda_1 = \mu_1^{-1} = \left( \sup_{v \in V, v \neq 0} \frac{\|\gamma v\|_H^2}{\|v\|_V^2} \right)^{-1} = \inf_{v \in V, v \neq 0} \frac{\|v\|_V^2}{\|\gamma v\|_H^2}$$

is the smallest eigenvalue of problem (2.1). □

### 3. Abstract inequality of Friedrichs–Poincaré type.

**3.1. The proof of the abstract inequality.** Properties of the eigenproblem (2.1) can be utilized in a simple way to derive an abstract inequality of Friedrichs–Poincaré type. The Hilbert space versions of the particular Friedrichs, Poincaré, trace, Korn’s, and similar inequalities easily follow from this abstract result. For examples see sections 4–6.

**THEOREM 3.1** (abstract inequality). *Let  $\gamma : V \rightarrow H$  be a continuous, linear, and compact operator between Hilbert spaces  $V$  and  $H$ . Let  $\lambda_1$  be the smallest eigenvalue of the problem (2.1). Then*

$$(3.1) \quad \|\gamma v\|_H \leq C_\gamma \|v\|_V \quad \forall v \in V$$

with  $C_\gamma = \lambda_1^{-1/2}$ . Moreover, this constant is optimal in the sense that it is the smallest possible constant such that (3.1) holds for all  $v \in V$ .

*Proof.* The validity of the abstract inequality follows immediately from (2.7):

$$\|\gamma v\|_H^2 \leq \lambda_1^{-1} \|v\|_V^2 \quad \forall v \in V.$$

This inequality holds as the equality for  $v = u_1$  and thus the constant  $C_\gamma = \lambda_1^{-1/2}$  is optimal.  $\square$

**3.2. Upper bound on the smallest eigenvalue.** The upper bound on  $\lambda_1$  can be computed by the standard Galerkin method, which is both accurate and efficient [5, 6, 11]. This method is based on the projection of the eigenproblem (2.1) into a finite dimensional subspace  $V^h \subset V$ . We seek eigenvectors  $u_i^h \in V^h$ ,  $u_i^h \neq 0$  and eigenvalues  $\lambda_i^h$  such that

$$(3.2) \quad (u_i^h, v^h)_V = \lambda_i^h (\gamma u_i^h, \gamma v^h)_H \quad \forall v^h \in V^h.$$

Let  $\{\varphi_j\}_{j=1,\dots,N}$  be a basis of the space  $V^h$ . Then, we can formulate problem (3.2) equivalently as a generalized eigenvalue problem

$$A y_i = \lambda_i^h M y_i$$

for matrices  $A$  and  $M$  with entries

$$A_{jk} = (\varphi_k, \varphi_j)_V \quad \text{and} \quad M_{jk} = (\gamma \varphi_k, \gamma \varphi_j)_H, \quad j, k = 1, 2, \dots, N.$$

The eigenvectors  $y_i \in \mathbb{R}^N$  and  $u_i^h \in V^h$  are linked by the relation  $u_i^h = \sum_{j=1}^N y_{i,j} \varphi_j$ . The generalized matrix eigenvalue problem can be solved by efficient methods of numerical linear algebra [7].

The Galerkin method for eigenvalue problems is very well understood. The convergence and the speed of convergence of this method are established, for example, in [5, 6, 11]. It is well known [11] that the Galerkin method approximates the exact eigenvalues from above, hence

$$\lambda_i \leq \lambda_i^h \quad \forall i = 1, 2, \dots$$

In particular, the upper bound on the smallest eigenvalue  $\lambda_1$  and the corresponding lower bound on the optimal constant  $C_\gamma$  read

$$(3.3) \quad \lambda_1 \leq \lambda_1^h \quad \text{and} \quad (\lambda_1^h)^{-1/2} \leq C_\gamma.$$

**3.3. Lower bound on the smallest eigenvalue.** In this part we concentrate on a computable lower bound on the smallest eigenvalue  $\lambda_1$  of the problem (2.1). First, we formulate an auxiliary result, which states that the images  $\gamma u_i, i = 1, 2, \dots$ , of the orthogonal system of eigenvectors  $u_i$  satisfy the Parseval's identity.

LEMMA 3.2. *Let  $u_i, i = 1, 2, \dots$ , be the above specified orthogonal system of eigenvectors of the operator  $S\gamma$  corresponding to nonzero eigenvalues. Let these eigenvectors be normalized as in (2.6). Let  $u_* \in V$  be arbitrary. Then*

$$\|\gamma u_*\|_H^2 = \sum_{i=1}^{\infty} |(\gamma u_*, \gamma u_i)_H|^2.$$

*Proof.* Due to the decomposition (2.4), there exist unique components  $u_*^{\mathcal{M}} \in \mathcal{M}$  and  $u_*^0 \in \ker(S\gamma)$  such that  $u_* = u_*^{\mathcal{M}} + u_*^0$ . Since  $S\gamma u_*^0 = 0$ , we have  $0 = (S\gamma u_*^0, u_*^0)_V = (\gamma u_*^0, \gamma u_*^0)_H$  by (2.2) and hence  $\gamma u_*^0 = 0$ . Consequently,  $\gamma u_* = \gamma u_*^{\mathcal{M}}$ .

System  $\gamma u_i, i = 1, 2, \dots$ , forms an orthonormal basis in  $\gamma\mathcal{M}$ . Thus, we can use the standard Parseval's identity in  $\gamma\mathcal{M}$  [43, Theorem 2, Chapter III, section 4] to obtain

$$\|\gamma u_*\|_H^2 = \|\gamma u_*^{\mathcal{M}}\|_H^2 = \sum_{i=1}^{\infty} |(\gamma u_*^{\mathcal{M}}, \gamma u_i)_H|^2 = \sum_{i=1}^{\infty} |(\gamma u_*, \gamma u_i)_H|^2. \quad \square$$

The derivation of the lower bound on  $\lambda_1$  is based on the method of a priori–a posteriori inequalities. This method relies on an abstract theorem proved in [22]. We formulate this theorem in the setting of this paper and for the readers' convenience we present its brief proof. Notice that in contrast to [22], Theorem 3.3 operates with a pair of Hilbert spaces and with a compact operator between them.

THEOREM 3.3. *Let  $\gamma : V \rightarrow H$  be a continuous, linear, and compact operator between Hilbert spaces  $V$  and  $H$ . Let  $u_* \in V$  and  $\lambda_* \in \mathbb{R}$  be arbitrary. Let  $\lambda_i, i = 1, 2, \dots$ , be eigenvalues of (2.1). Let us consider  $w \in V$  such that*

$$(3.4) \quad (w, v)_V = (u_*, v)_V - \lambda_*(\gamma u_*, \gamma v)_H \quad \forall v \in V.$$

If  $\gamma u_* \neq 0$ , then

$$(3.5) \quad \min_i \left| \frac{\lambda_i - \lambda_*}{\lambda_i} \right| \leq \frac{\|\gamma w\|_H}{\|\gamma u_*\|_H}.$$

*Proof.* Using Lemma 3.2 and definitions (2.1) and (3.4), we obtain

$$\begin{aligned} \min_i \left| \frac{\lambda_i - \lambda_*}{\lambda_i} \right|^2 \|\gamma u_*\|_H^2 &\leq \sum_{i=1}^{\infty} \left| \frac{\lambda_i - \lambda_*}{\lambda_i} (\gamma u_*, \gamma u_i)_H \right|^2 \\ &= \sum_{i=1}^{\infty} \left| \frac{(u_i, u_*)_V}{\lambda_i} - \frac{(u_* - w, u_i)_V}{\lambda_i} \right|^2 \\ &= \sum_{i=1}^{\infty} \left| \frac{(w, u_i)_V}{\lambda_i} \right|^2 = \sum_{i=1}^{\infty} |(\gamma w, \gamma u_i)_H|^2 = \|\gamma w\|_H^2. \quad \square \end{aligned}$$

In order to obtain a computable lower bound on  $\lambda_1$ , we combine the estimate (3.5) with a complementarity technique; see sections 4–6. The bounds derived by the complementarity technique depend on the particular choice of spaces  $V$  and  $H$ , but

they have a common general structure. The following theorem utilizes this general structure and presents an abstract lower bound on  $\lambda_1$ .

**THEOREM 3.4** (abstract complementarity estimate). *Let  $u_* \in V$ ,  $\lambda_* \in \mathbb{R}$  be arbitrary and let  $w \in V$  satisfy (3.4). Let  $\lambda_1$  be the smallest eigenvalue of (2.1) and let the relatively closest eigenvalue to  $\lambda_*$  be  $\lambda_i$ , i.e., let*

$$(3.6) \quad \left| \frac{\lambda_1 - \lambda_*}{\lambda_1} \right| \leq \left| \frac{\lambda_i - \lambda_*}{\lambda_i} \right| \quad \forall i = 1, 2, \dots$$

Further, let  $A \geq 0$  and  $B \geq 0$  be such that

$$(3.7) \quad B < \lambda_* \|\gamma u_*\|_H$$

and

$$(3.8) \quad \|w\|_V \leq A + C_\gamma B,$$

where  $C_\gamma$  is the optimal constant from (3.1). Then

$$(3.9) \quad X_2^2 \leq \lambda_1 \quad \text{and} \quad C_\gamma \leq 1/X_2,$$

where

$$(3.10) \quad X_2 = \frac{1}{2} \left( -\alpha + \sqrt{\alpha^2 + 4(\lambda_* - \beta)} \right), \quad \alpha = \frac{A}{\|\gamma u_*\|_H}, \quad \text{and} \quad \beta = \frac{B}{\|\gamma u_*\|_H}.$$

*Proof.* By using the fact that  $C_\gamma = \lambda_1^{-1/2}$ , the estimate (3.5), assumptions of the theorem, and the inequality (3.1), we obtain the validity of the following relation:

$$\lambda_* C_\gamma^2 - 1 = \frac{\lambda_* - \lambda_1}{\lambda_1} \leq \min_i \left| \frac{\lambda_i - \lambda_*}{\lambda_i} \right| \leq \frac{\|\gamma w\|_H}{\|\gamma u_*\|_H} \leq C_\gamma \frac{\|w\|_V}{\|\gamma u_*\|_H} \leq C_\gamma \alpha + C_\gamma^2 \beta.$$

This is equivalent to the quadratic inequality

$$0 \leq C_\gamma^2(\beta - \lambda_*) + C_\gamma \alpha + 1.$$

By solving it for  $C_\gamma$  under the assumption (3.7), we conclude that this inequality is not satisfied for  $C_\gamma > 1/X_2$ . Thus,  $C_\gamma$  has to be at most  $1/X_2$ .  $\square$

The particular complementarity estimates have the form (3.8), where the numbers  $A$  and  $B$  are obtained by an approximate minimization procedure; see sections 4–6. The better the approximation of the exact minimizer we compute, the tighter bound (3.8) and consequently (3.9) we obtain. The exact minimizer yields equality in (3.8) and  $B = 0$ . Therefore, the assumption (3.7) can always be satisfied by computing a sufficiently accurate minimizer.

The assumption (3.6) is crucial and cannot be guaranteed unless lower bounds on  $\lambda_1$  and  $\lambda_2$  are known. However, since the Galerkin method is known to converge [5, 6, 11] with a known speed, very accurate approximations of  $\lambda_1$  and  $\lambda_2$  can be computed. If these approximations are well separated, they can be used in (3.6) to verify its validity with a good confidence.

In order to increase this confidence, we propose a test. This test is based on the following observation. Let  $\lambda_1^{\text{low}} \leq \lambda_1 \leq \lambda_* \leq \lambda_2^{\text{low}} \leq \lambda_2 \leq \lambda_2^{\text{up}}$ ,  $D_1 = (\lambda_* - \lambda_1^{\text{low}})/\lambda_1^{\text{low}}$ , and  $D_2 = (\lambda_2^{\text{low}} - \lambda_*)/\lambda_2^{\text{up}}$ . Then inequality  $D_1 \leq D_2$  implies the assumption (3.6).

Indeed, if all these inequalities are satisfied, then, clearly,  $|(\lambda_1 - \lambda_*)/\lambda_1| \leq D_1 \leq D_2 \leq |(\lambda_2 - \lambda_*)/\lambda_2| \leq |(\lambda_i - \lambda_*)/\lambda_i|$  for all  $i = 2, 3, \dots$  and (3.6) holds true.

Thus, if we knew guaranteed lower bounds  $\lambda_1^{\text{low}}$  and  $\lambda_2^{\text{low}}$  on the first eigenvalue  $\lambda_1$  and the second eigenvalue  $\lambda_2$  of (2.1), respectively, then we could use the Galerkin method to compute upper bounds  $\lambda^*$  and  $\lambda_2^{\text{up}}$  and check if  $\lambda^* \leq \lambda_2^{\text{low}}$  and the inequality  $D_1 \leq D_2$  would then guarantee the validity of (3.6). However, the guaranteed lower bounds  $\lambda_1^{\text{low}}$  and  $\lambda_2^{\text{low}}$  are not available.

Therefore, in practice we propose to set  $\lambda_1^{\text{low}} = X_2^2$  (see (3.9)),  $\lambda_2^{\text{low}} = (\lambda_* + \lambda_2^{\text{up}})/2$  and compute  $D_1$  and  $D_2$  with these values. This yields the following diagnostics indicating the validity of the assumption (3.6). If  $D_2 \leq 0$  or  $D_1 > D_2$ , then it is highly probable that some of the assumptions are not satisfied and the results should not be trusted. On the other hand if  $D_2 > 0$  and  $D_1$  is several times smaller than  $D_2$ , then we can have good confidence in the validity of (3.6). This diagnostics performs very well in all numerical experiments presented below. At the early stages of the computations the approximations are not very precise and the diagnostics showed that the results are untrustworthy. However, at the final stages of the computations the value  $D_1$  is at least five times smaller than  $D_2$ , providing good confidence in the validity of (3.6).

**4. Application to Friedrichs inequality.** In this section we apply the above general theory to the case of Friedrichs inequality. We will consider the variant of Friedrichs inequality that is suitable for general symmetric second-order linear elliptic differential operators. First, we introduce differential operators of this general type (see subsection 4.1) and provide the corresponding Friedrichs inequality (see subsection 4.2). In subsection 4.3 we derive two-sided bounds on the optimal constant in Friedrichs inequality that are based on the general theory and on the complementarity technique. In subsection 4.4 certain computational issues are discussed, and finally, in subsection 4.5 we present numerical results.

**4.1. A general symmetric second-order linear elliptic operator.** Let us consider a domain  $\Omega \subset \mathbb{R}^d$  with Lipschitz boundary  $\partial\Omega$ . Let  $\partial\Omega$  consist of two relatively open parts  $\Gamma_D$  and  $\Gamma_N$  such that  $\partial\Omega = \overline{\Gamma_D} \cup \overline{\Gamma_N}$  and  $\Gamma_D \cap \Gamma_N = \emptyset$ . Note that we admit the case where either  $\Gamma_D$  or  $\Gamma_N$  is empty. Further, let us consider a matrix function  $\mathcal{A} \in [L^\infty(\Omega)]^{d \times d}$  and coefficients  $c \in L^\infty(\Omega)$  and  $\alpha \in L^\infty(\Gamma_N)$ . Matrix  $\mathcal{A}$  is assumed to be symmetric and uniformly positive definite, i.e., there exists a constant  $C > 0$  such that

$$\xi^T \mathcal{A}(x) \xi \geq C |\xi|^2 \quad \forall \xi \in \mathbb{R}^d \text{ and for almost all } x \in \Omega,$$

where  $|\cdot|$  stands for the Euclidean norm. Coefficients  $c$  and  $\alpha$  are considered to be nonnegative.

Further, we introduce a subspace

$$H_{\Gamma_D}^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D \text{ in the sense of traces}\}$$

of the Sobolev space  $H^1(\Omega)$  of  $L^2(\Omega)$  functions with square integrable distributional derivatives. In what follows, we use the notation  $(\cdot, \cdot)$  and  $(\cdot, \cdot)_{L^2(\Gamma_N)}$  for the  $L^2(\Omega)$  and  $L^2(\Gamma_N)$  scalar products, respectively. Using this notation, we define a bilinear form

$$(4.1) \quad a(u, v) = (\mathcal{A} \nabla u, \nabla v) + (cu, v) + (\alpha u, v)_{L^2(\Gamma_N)}.$$



**4.2. Friedrichs inequality and the optimal constant.** The bilinear form (4.1) is a scalar product in  $H_{\Gamma_D}^1(\Omega)$  under the conditions presented in the following lemma. Its proof follows, for instance, from [21, Theorem 5.11.2].

LEMMA 4.1. *The bilinear form defined in (4.1) is a scalar product in  $H_{\Gamma_D}^1(\Omega)$  provided that at least one of the following conditions is satisfied:*

- (a)  $\text{meas}_{d-1} \Gamma_D > 0$ ,
- (b) *there exists a nonempty ball  $B \subset \Omega$  such that  $c > 0$  on  $B$ ,*
- (c) *there exists a subset  $\Gamma_\alpha \subset \Gamma_N$  such that  $\text{meas}_{d-1} \Gamma_\alpha > 0$  and  $\alpha > 0$  on  $\Gamma_\alpha$ .*

In this section we assume that at least one of conditions (a)–(c) of Lemma 4.1 is satisfied, hence, that  $a(\cdot, \cdot)$  is a scalar product in  $H_{\Gamma_D}^1(\Omega)$ . The equivalence of the norm induced by  $a(\cdot, \cdot)$  and the standard  $H^1$  norm yields completeness and therefore the space  $H_{\Gamma_D}^1(\Omega)$  equipped with the scalar product  $a(\cdot, \cdot)$  is a Hilbert space. This enables us to use the theory from section 2. We set

$$(4.2) \quad V = H_{\Gamma_D}^1(\Omega), \quad (u, v)_V = a(u, v), \quad H = L^2(\Omega), \quad (u, v)_H = (u, v),$$

and we define  $\gamma : H_{\Gamma_D}^1(\Omega) \rightarrow L^2(\Omega)$  as the identity operator, which is compact due to the Rellich theorem [1, Theorem 6.3]. With this setting and with the notation  $\|\cdot\|_a$  for the norm induced by  $a(\cdot, \cdot)$ , we obtain the validity of Friedrichs inequality.

THEOREM 4.2. *Let the bilinear form  $a(\cdot, \cdot)$  given by (4.1) form a scalar product in  $H_{\Gamma_D}^1(\Omega)$ . Then there exists a constant  $C_F > 0$  such that*

$$(4.3) \quad \|v\|_{L^2(\Omega)} \leq C_F \|v\|_a \quad \forall v \in H_{\Gamma_D}^1(\Omega).$$

Moreover, the optimal value of this constant is  $C_F = \lambda_1^{-1/2}$ , where  $\lambda_1$  is the smallest eigenvalue of the following problem: find  $u_i \in H_{\Gamma_D}^1(\Omega)$ ,  $u_i \neq 0$ , and  $\lambda_i \in \mathbb{R}$  such that

$$(4.4) \quad a(u_i, v) = \lambda_i (u_i, v) \quad \forall v \in H_{\Gamma_D}^1(\Omega).$$

*Proof.* Lemma 4.1 together with the equivalence of the norms  $\|\cdot\|_a$  and  $\|\cdot\|_{H^1(\Omega)}$  guarantees that  $H_{\Gamma_D}^1(\Omega)$  with the scalar product  $a(\cdot, \cdot)$  given by (4.1) is a Hilbert space. The statement then follows immediately from Theorem 3.1.  $\square$

Let us note that the most common version of Friedrichs inequality,

$$\|v\|_{L^2(\Omega)} \leq C_F \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega),$$

follows from (4.3) with  $\Gamma_D = \partial\Omega$ ,  $\Gamma_N = \emptyset$ ,  $\mathcal{A}$  being an identity matrix,  $c = 0$ , and  $\alpha = 0$ . As usual, we denote by  $H_0^1(\Omega)$  the space  $H_{\Gamma_D}^1(\Omega)$  with  $\Gamma_D = \partial\Omega$ .

**4.3. Two-sided bounds on Friedrichs constant.** A lower bound on Friedrichs constant  $C_F$  can be efficiently computed by the Galerkin method. We use the setting (4.2) and proceed as described in section 3.2. The upper bound on  $C_F$  is obtained by the complementarity technique presented in the following theorem.

Let  $\mathbf{H}(\text{div}, \Omega)$  stand for the space of  $d$ -dimensional vector fields with square integrable weak divergences and let  $\mathbf{n}$  be the unit outward-facing normal vector to the boundary  $\partial\Omega$ . Further, let  $\|\mathbf{q}\|_{\mathcal{A}}^2 = (\mathcal{A}\mathbf{q}, \mathbf{q})$  be a norm in  $[L^2(\Omega)]^d$  induced by the matrix  $\mathcal{A}$ .

THEOREM 4.3. *Let  $V = H_{\Gamma_D}^1(\Omega)$ ,  $u_* \in V$ , and  $\lambda_* \in \mathbb{R}$ . Let the bilinear form  $a(\cdot, \cdot)$  given by (4.1) form a scalar product in  $V$ . Let  $w \in V$  satisfy*

$$(4.5) \quad a(w, v) = a(u_*, v) - \lambda_*(u_*, v) \quad \forall v \in V,$$

where we recall that  $(\cdot, \cdot)$  stands for the  $L^2(\Omega)$  scalar product. Then

$$(4.6) \quad \|w\|_a \leq \|\nabla u_* - \mathcal{A}^{-1}\mathbf{q}\|_{\mathcal{A}} + C_F \|\lambda_* u_* - cu_* + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)} \quad \forall \mathbf{q} \in W,$$

where  $W = \{\mathbf{q} \in \mathbf{H}(\operatorname{div}, \Omega) : \mathbf{q} \cdot \mathbf{n} = -\alpha u_* \text{ on } \Gamma_N\}$ .

*Proof.* Let us fix any  $\mathbf{q} \in W$ , test (4.5) by  $v = w$ , and use the divergence theorem to express

$$\begin{aligned} \|w\|_a^2 &= (\mathcal{A}\nabla u_*, \nabla w) + (cu_*, w) + (\alpha u_*, w)_{L^2(\Gamma_N)} - \lambda_*(u_*, w) - (\mathbf{q}, \nabla w) - (\operatorname{div} \mathbf{q}, w) \\ &\quad + (\mathbf{q} \cdot \mathbf{n}, w)_{L^2(\Gamma_N)} = (\mathcal{A}(\nabla u_* - \mathcal{A}^{-1}\mathbf{q}), \nabla w) - (\lambda_* u_* - cu_* + \operatorname{div} \mathbf{q}, w). \end{aligned}$$

The Cauchy–Schwarz inequality and Friedrichs inequality (4.3) yield

$$\|w\|_a^2 \leq \|\nabla u_* - \mathcal{A}^{-1}\mathbf{q}\|_{\mathcal{A}} \|\nabla w\|_{\mathcal{A}} + C_F \|\lambda_* u_* - cu_* + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)} \|w\|_a.$$

Inequality  $\|\nabla w\|_{\mathcal{A}} \leq \|w\|_a$  finishes the proof.  $\square$

The estimate (4.6) is of type (3.8) with

$$(4.7) \quad A = \|\nabla u_* - \mathcal{A}^{-1}\mathbf{q}\|_{\mathcal{A}}, \quad B = \|\lambda_* u_* - cu_* + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)},$$

and  $C_\gamma = C_F$ . The numbers  $A$  and  $B$  can be readily computed as soon as suitable approximations  $\lambda_*$ ,  $u_*$  and a suitable vector field  $\mathbf{q} \in W$  are in hand. Estimate (3.9) then gives a guaranteed upper bound on Friedrichs constant  $C_F$ .

The crucial part is the computation of suitable approximations  $\lambda_*$  and  $u_*$  of the smallest eigenvalue  $\lambda_1$  of problem (4.4) and its corresponding eigenvector  $u_1$  such that the inequality (3.6) is satisfied. Equally crucial is a suitable choice of the vector field  $\mathbf{q} \in W$  such that estimate (4.6) provides a tight upper bound on  $\|w\|_a$ . A possible approach and practical details about these issues are provided in the next section.

**4.4. Flux reconstruction.** In order to compute two-sided bounds on Friedrichs constant, we proceed as follows. First, we use the Galerkin method (see section 3.2) to compute an approximation  $\lambda_1^h$  and  $u_1^h$  of the eigenpair  $\lambda_1$  and  $u_1$  of (4.4), respectively. We set  $\lambda_* = \lambda_1^h$ ,  $u_* = u_1^h$ , compute a suitable vector field  $\mathbf{q} \in W$ , and evaluate the numbers  $A$  and  $B$  by (4.7). The estimate (3.9) then provides the upper bound and the Galerkin approximation  $\lambda_1^h$  yields the lower bound (3.3) on Friedrichs constant  $C_F$ .

The key point is the computation of the suitable vector field  $\mathbf{q}$ . For simplicity, we choose a straightforward approach of approximate minimization of the upper bound (4.6) with respect to a suitable subset of  $W$ . First, we exploit the affine structure of  $W$ . Let us choose an arbitrary but fixed  $\bar{\mathbf{q}} \in W$ . The practical construction of this  $\bar{\mathbf{q}}$  is a geometrical issue depending on  $\Omega$  and coefficient  $\alpha$ . It suffices to construct a vector field  $\bar{\mathbf{q}}_1 \in \mathbf{H}(\operatorname{div}, \Omega)$  such that  $\bar{\mathbf{q}}_1 \cdot \mathbf{n} = 1$  on  $\Gamma_N$  and a function  $\bar{\alpha} \in H^1(\Omega)$  such that  $\bar{\alpha} = \alpha$  on  $\Gamma_N$ . Let us note that usually Raviart–Thomas–Nédélec space is considered for construction of vector fields in  $\mathbf{H}(\operatorname{div}, \Omega)$ . Then we can simply set  $\bar{\mathbf{q}} = -\bar{\alpha} u_* \bar{\mathbf{q}}_1$ . This  $\bar{\mathbf{q}}$  obviously satisfies the boundary condition  $\bar{\mathbf{q}} \cdot \mathbf{n} = -\alpha u_*$  on  $\Gamma_N$ . To guarantee  $\bar{\mathbf{q}} \in \mathbf{H}(\operatorname{div}, \Omega)$  we need an extra smoothness of  $\bar{\mathbf{q}}_1$  and  $\bar{\alpha}$ . Since  $u_* = u_1^h$  is piecewise polynomial and thus in  $W^{1,\infty}(\Omega)$ , it suffices to have  $\bar{\mathbf{q}}_1 \in [H^1(\Omega)]^d$  and  $\bar{\alpha} \in W^{1,\infty}(\Omega)$ , or in the opposite way  $\bar{\mathbf{q}}_1 \in [W^{1,\infty}(\Omega)]^d$  and  $\bar{\alpha} \in H^1(\Omega)$ .

In any case, having a  $\bar{\mathbf{q}} \in W$ , we can express the affine space  $W$  as  $W = \bar{\mathbf{q}} + W_0$ , where

$$W_0 = \{\mathbf{q} \in \mathbf{H}(\operatorname{div}, \Omega) : \mathbf{q} \cdot \mathbf{n} = 0 \text{ on } \Gamma_N\}$$

is already a linear space. The idea is to minimize the upper bound (4.6) over the set  $\bar{\mathbf{q}} + W_0^h$ , where  $W_0^h \subset W_0$  is a finite dimensional subspace. However, the right-hand side of (4.6) is a nonlinear functional in  $\mathbf{q}$  and thus, in order to simplify the computation, we use the idea from [33] and approximate it by a quadratic functional. We rewrite inequality (4.6) using notation (4.7) and Young's inequality as

$$(4.8) \quad \|w\|_a^2 \leq (A + C_F B)^2 \leq (1 + \varrho^{-1})A^2 + (1 + \varrho)C_F^2 B^2 \quad \forall \varrho > 0.$$

The right-hand side of this inequality is already a quadratic functional for a fixed  $\varrho$ , but the exact value of  $C_F$  is unknown in general. However, it is sufficient to find an approximate minimizer only. Therefore we approximate  $C_F$  by the available value  $(\lambda_1^h)^{-1/2}$ . This leads us to the minimization of

$$(4.9) \quad (1 + \varrho^{-1})\|\nabla u_1^h - \mathcal{A}^{-1}\mathbf{q}\|_{\mathcal{A}}^2 + (1 + \varrho)(\lambda_1^h)^{-1}\|\lambda_1^h u_1^h - cu_1^h + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)}^2$$

over the affine set  $\bar{\mathbf{q}} + W_0^h$  with a fixed value of  $\varrho > 0$ . This minimization problem is equivalent to seeking  $\mathbf{q}_0^h \in W_0^h$  satisfying

$$(4.10) \quad \mathcal{B}(\mathbf{q}_0^h, \mathbf{w}_0^h) = \mathcal{F}(\mathbf{w}_0^h) - \mathcal{B}(\bar{\mathbf{q}}, \mathbf{w}_0^h) \quad \forall \mathbf{w}_0^h \in W_0^h,$$

where

$$\begin{aligned} \mathcal{B}(\mathbf{q}, \mathbf{w}) &= (\operatorname{div} \mathbf{q}, \operatorname{div} \mathbf{w}) + \frac{\lambda_1^h}{\varrho}(\mathcal{A}^{-1}\mathbf{q}, \mathbf{w}), \\ \mathcal{F}(\mathbf{w}) &= \frac{\lambda_1^h}{\varrho}(\nabla u_1^h, \mathbf{w}) - (\lambda_1^h u_1^h - cu_1^h, \operatorname{div} \mathbf{w}). \end{aligned}$$

The computed vector field  $\mathbf{q}^h = \bar{\mathbf{q}} + \mathbf{q}_0^h \in W$  is then used in (4.7) to evaluate  $A$  and  $B$  and consequently the two-sided bounds  $C_F^{\text{low}} \leq C_F \leq C_F^{\text{up}}$ , where

$$(4.11) \quad C_F^{\text{low}} = (\lambda_1^h)^{-1/2} \quad \text{and} \quad C_F^{\text{up}} = 1/X_2,$$

and  $X_2$  is given by (3.10) with  $\gamma u_* = u_1^h$ ; see (3.3) and (3.9).

Note that problem (4.10) can be naturally approached by standard Raviart–Thomas–Nédélec finite elements [13]. Further we note that a particular value of the constant  $\varrho$  can influence the accuracy of the final bound. However, this influence was minor in all cases we numerically tested and the natural value  $\varrho = 1$  yielded accurate results. If necessary a simultaneous minimization of (4.9) with respect to both  $\varrho > 0$  and  $\mathbf{q} \in \bar{\mathbf{q}} + W_0^h$  can be performed.

**4.5. Numerical experiment.** In order to illustrate the capabilities of the above described approach for computation of two-sided bounds on Friedrichs constant, we present numerical results showing the dependence of Friedrichs constant on piecewise constant values of  $\mathcal{A}$ . We consider the general setting from section 4.1 and in particular we set  $\Omega = (-1, 1)^2$ ,  $\Gamma_N = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 = 1 \text{ and } -1 < x_2 < 1\}$ ,  $\Gamma_D = \partial\Omega \setminus \bar{\Gamma}_N$ ,  $c = 0$ , and  $\alpha = 0$ . The matrix  $\mathcal{A}$  is piecewise constant, defined as  $\mathcal{A}(x_1, x_2) = I$  for  $x_1 x_2 \leq 0$  and  $\mathcal{A}(x_1, x_2) = \tilde{a}I$  for  $x_1 x_2 > 0$ , where  $I \in \mathbb{R}^{d \times d}$  stands for the identity matrix and the value of the constant  $\tilde{a}$  is specified below.

We employ the standard piecewise linear triangular finite elements to discretize eigenvalue problem (4.4). Thus, we consider a conforming triangular mesh  $\mathcal{T}_h$  and seek the Galerkin solution of problem (4.4) in space

$$(4.12) \quad V^h = \{v^h \in H_{\Gamma_D}^1(\Omega) : v^h|_K \in P^1(K) \forall K \in \mathcal{T}_h\},$$

where  $P^1(K)$  denotes the space of affine functions on  $K$ ; see section 3.2.

We point out that the discontinuity of  $\mathcal{A}$  causes a strong singularity of the eigenvectors at the origin. Therefore, we employ a standard adaptive algorithm (see Algorithm 1) and construct adaptively refined meshes in order to approximate the singularity well. We use the localized version of (4.9) to define the error indicators

$$(4.13) \quad \eta_K^2 = (1 + \varrho^{-1}) \|\nabla u_1^h - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A},K}^2 + (1 + \varrho)(\lambda_1^h)^{-1} \|\lambda_1^h u_1^h - cu_1^h + \operatorname{div} \mathbf{q}\|_{L^2(K)}^2$$

for all  $K \in \mathcal{T}_h$ , where  $\|\mathbf{w}\|_{\mathcal{A},K}^2 = (\mathcal{A}\mathbf{w}, \mathbf{w})_{L^2(K)}$ .

This error indicator is a natural choice, because we compute both  $u_1^h$  and an approximation of  $\mathbf{q}$  on the same mesh and hence we need an indicator that combines errors of both these solutions. Quantities (4.13) do it naturally. Indeed, if  $u_1^h$  is not accurate, then its residual  $w$  is large and  $\eta_K$  show high values, because the sum of their squares approximates the upper bound (4.8). Similarly, if  $\mathbf{q}^h$  is far from the exact minimizer  $\mathbf{q}$ , then  $\eta_K$  exhibit high values as well. In addition, quantities  $\eta_K$  are readily available, because both norms in (4.13) have to be computed anyway to get the upper bound  $C_F^{\text{up}}$ .

We note that all numerical experiments in this paper use our own MATLAB finite element implementation. Meshes are refined by routines of the MATLAB `PDEtoolbox` and generalized eigenvalue problems for resulting sparse matrices are solved by the MATLAB routine `eigs` that is based on the `ARPACK` package [25].

The unknown value of Friedrichs constant lies between bounds  $C_F^{\text{low}}$  and  $C_F^{\text{up}}$  computed in Step 4 of Algorithm 1. When Algorithm 1 stops, the relative error is guaranteed to be at most the given tolerance  $E_{\text{TOL}}$ .

In this particular numerical experiment we have chosen the initial mesh with eight triangles as shown in Figure 4.1. The marking parameter in Step 6 and the tolerance for the relative error in Step 4 were chosen as  $\theta = 0.75$  and  $E_{\text{TOL}} = 0.01$ , respectively. Further, we naturally set  $\varrho = 1$ . We made an attempt to find an optimal value for  $\varrho$ .

---

ALGORITHM 1. Mesh adaptation algorithm for two-sided bounds on Friedrichs constant.

---

**Step 1** Construct an initial mesh  $\mathcal{T}_h$ .

**Step 2** Use the space (4.12) and compute Galerkin approximations  $\lambda_1^h \in \mathbb{R}$  and  $u_1^h \in V^h$  of the smallest eigenvalue and the corresponding eigenvector of problem (4.4), see (3.2).

**Step 3** Use the finite element space  $W_0^h = \{\mathbf{w}_h \in W_0 : \mathbf{w}_h \in [P^2(K)]^2 \forall K \in \mathcal{T}_h\}$  and solve (4.10).

**Step 4** Evaluate two-sided bounds (4.11). Set  $C_F^{\text{avg}} = (C_F^{\text{up}} + C_F^{\text{low}})/2$ ,  $E_{\text{REL}} = (C_F^{\text{up}} - C_F^{\text{low}})/C_F^{\text{avg}}$ , and stop the algorithm as soon as  $E_{\text{REL}} \leq E_{\text{TOL}}$ .

**Step 5** Compute error indicators (4.13) for all  $K \in \mathcal{T}_h$  and sort them in descending order:  $\eta_{K_1} \geq \eta_{K_2} \geq \dots \geq \eta_{K_{N_{\text{el}}}}$ , where  $N_{\text{el}}$  is the number of elements in  $\mathcal{T}_h$ .

**Step 6** (Bulk criterion.) Find the smallest  $n$  such that  $\theta^2 \sum_{i=1}^{N_{\text{el}}} \eta_{K_i}^2 \leq \sum_{i=1}^n \eta_{K_i}^2$ , where  $\theta \in (0, 1)$  is a parameter.

**Step 7** Construct a new mesh  $\mathcal{T}_h$  by refining elements  $K_1, K_2, \dots, K_n$ .

**Step 8** Go to Step 2.

---

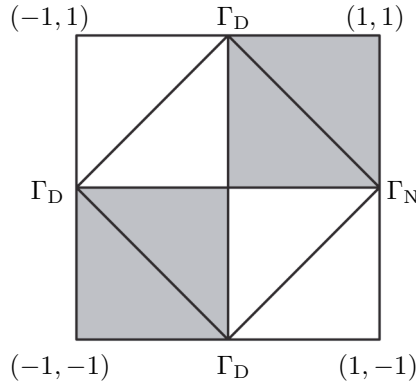


FIG. 4.1. The initial mesh for the adaptive algorithm. The Dirichlet and Neumann parts of the boundary are indicated as well as the piecewise constant matrix  $\mathcal{A}$ :  $\mathcal{A} = I$  in white elements and  $\mathcal{A} = \tilde{a}I$  in gray elements.

TABLE 4.1

Friedrichs constant. The lower bound  $C_F^{\text{low}}$ , upper bound  $C_F^{\text{up}}$ , relative error  $E_{\text{REL}}$ , and number of degrees of freedom  $N_{\text{DOF}} = \dim V^h$  for particular values of  $\tilde{a}$ .

$\tilde{a}$	$C_F^{\text{low}}$	$C_F^{\text{up}}$	$E_{\text{REL}}$	$N_{\text{DOF}}$
0.001	9.0086	9.0939	0.0094	4832
0.01	2.8697	2.8971	0.0095	5003
0.1	1.0035	1.0124	0.0088	7866
1	0.5693	0.5743	0.0086	4802
10	0.3173	0.3201	0.0088	7866
100	0.2870	0.2897	0.0095	5003
1000	0.2849	0.2876	0.0094	4832

However, this decreases the relative error by a factor of magnitude  $10^{-4}$ – $10^{-3}$  on a fixed mesh, which is below the target accuracy  $E_{\text{TOL}} = 0.01$ . Therefore, we use the natural value  $\varrho = 1$  that equilibrates both terms in (4.13). A similar statement can be made in all subsequent numerical experiments. On the other hand, certain examples might behave differently and optimization of  $\varrho$  could be important.

The results for a series of values of  $\tilde{a}$  are summarized in Table 4.1. For each particular value of  $\tilde{a}$ , we run the adaptive algorithm until the relative error drops below  $E_{\text{TOL}} = 0.01$ . This error level was reached in all cases using several thousand degrees of freedom. Notice the considerable dependence of the optimal value of Friedrichs constant on  $\tilde{a}$ . The values for  $\tilde{a} = 0.001$  and  $\tilde{a} = 1000$  differ more than 30 times. Further notice that the exact value of Friedrichs constant for  $\tilde{a} = 1$  is  $C_F = 4/(\pi\sqrt{5}) \approx 0.5694$ .

In order to illustrate the behavior of the adaptive process, we consider the case  $\tilde{a} = 0.001$ . The convergence of the bounds  $C_F^{\text{up}}$  and  $C_F^{\text{low}}$  is presented in Figure 4.2, left. The right panel of Figure 4.2 shows the convergence of the relative error  $E_{\text{REL}}$  together with the heuristic indicators  $D_1$  and  $D_2$  introduced at the end of section 3.3. The fact that  $D_1$  is several times smaller than  $D_2$  at the later stages of the adaptive process indicates good confidence in the validity of assumption (3.6). In addition, two adapted meshes are drawn in Figures 6.2–6.3 (left).

## 5. Application to the Poincaré inequality.

**5.1. Poincaré inequality and the optimal constant.** In this section we consider the case when none of conditions (a)–(c) of Lemma 4.1 is satisfied. Therefore, we

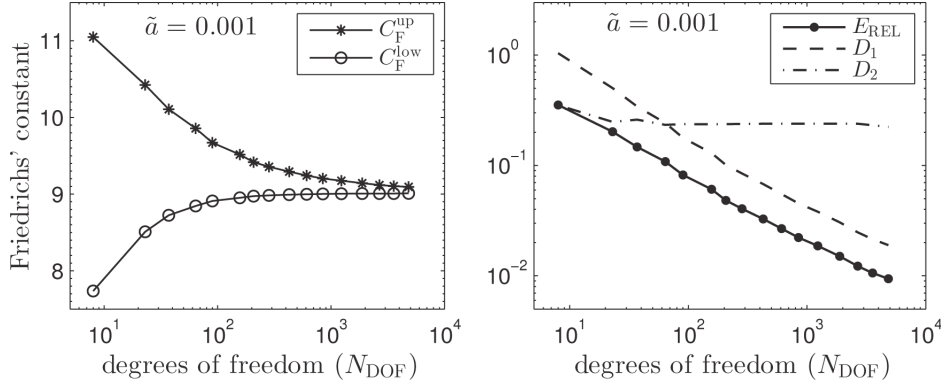


FIG. 4.2. Friedrichs constant for  $\tilde{a} = 0.001$ . Convergence of bounds  $C_F^{\text{up}}$  and  $C_F^{\text{low}}$  (left) and of the relative error  $E_{\text{REL}}$  and heuristic indicators  $D_1$  and  $D_2$  (right).

assume the general symmetric second-order elliptic operator as described in section 4.1 with  $c = 0$ ,  $\alpha = 0$ ,  $\Gamma_D = \emptyset$ , and  $\Gamma_N = \partial\Omega$ . We apply the general theory of section 3 with  $V = \tilde{H}^1(\Omega) = \{v \in H^1(\Omega) : (v, 1) = 0\}$ ,  $(u, v)_V = a(u, v) = (\mathcal{A}\nabla u, \nabla v)$ ,  $H = L^2(\Omega)$ , and  $(u, v)_H = (u, v)$ . It is an easy exercise to verify that  $a(\cdot, \cdot)$  forms a scalar product on  $\tilde{H}^1(\Omega)$  and that it induces a norm  $\|\cdot\|_a$  equivalent to the standard  $H^1$ -seminorm and  $H^1$ -norm. The operator  $\gamma : V \rightarrow H$  is set to be the identity mapping  $I : \tilde{H}^1(\Omega) \rightarrow L^2(\Omega)$ . This mapping is clearly compact, because the identity mapping from  $\tilde{H}^1(\Omega)$  to  $H^1(\Omega)$  is linear and continuous and the identity mapping from  $H^1(\Omega)$  to  $L^2(\Omega)$  is compact due to the Rellich theorem [1, Theorem 6.3]. This setting enables us to use the general conclusions of Theorem 3.1 and obtain the following result.

THEOREM 5.1. *There exists a constant  $C_P > 0$  such that*

$$(5.1) \quad \|v\|_{L^2(\Omega)} \leq C_P \|v\|_a \quad \forall v \in \tilde{H}^1(\Omega).$$

Moreover, the optimal value of this constant is  $C_P = \lambda_1^{-1/2}$ , where  $\lambda_1$  is the smallest positive eigenvalue of the following problem: find  $u_i \in H^1(\Omega)$ ,  $u_i \neq 0$ , and  $\lambda_i \in \mathbb{R}$  such that

$$(5.2) \quad a(u_i, v) = \lambda_i (u_i, v) \quad \forall v \in H^1(\Omega).$$

*Proof.* The inequality (5.1) follows immediately from Theorem 3.1. This theorem also implies that the optimal constant is  $C_P = \tilde{\lambda}_1^{-1/2}$ , where  $\tilde{\lambda}_1$  is the smallest eigenvalue of the following problem: find  $\tilde{u}_i \in \tilde{H}^1(\Omega)$ ,  $\tilde{u}_i \neq 0$ , and  $\tilde{\lambda}_i \in \mathbb{R}$  such that

$$(5.3) \quad a(\tilde{u}_i, v) = \tilde{\lambda}_i (\tilde{u}_i, v) \quad \forall v \in \tilde{H}^1(\Omega).$$

Notice that  $0 < \tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots$ . Similarly, the eigenvalues of (5.2) satisfy  $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots$  and the zero eigenvalue corresponds to a constant eigenvector  $u_0 = 1$ . It can be easily shown that  $\tilde{\lambda}_i = \lambda_i$  and  $\tilde{u}_i = u_i$  for all  $i = 1, 2, \dots$ . Thus, the smallest eigenvalue  $\tilde{\lambda}_1$  of (5.3) is equal to the smallest positive eigenvalue  $\lambda_1$  of (5.2) and the proof is finished.  $\square$

Let us note that the Poincaré inequality (5.1) can be equivalently formulated as

$$(5.4) \quad \|v - \tilde{v}\|_{L^2(\Omega)} \leq C_P \|v\|_a \quad \forall v \in H^1(\Omega), \quad \tilde{v} = (v, 1)/|\Omega|.$$

The common version of the Poincaré inequality

$$\|v - \tilde{v}\|_{L^2(\Omega)} \leq C_P \|\nabla v\|_{L^2(\Omega)} \quad \forall v \in H^1(\Omega), \quad \tilde{v} = (v, 1)/|\Omega|,$$

then follows from (5.4), or equivalently from (5.1), with  $\mathcal{A}$  being the identity matrix.

**5.2. Two-sided bounds on the Poincaré constant.** A lower bound on the Poincaré constant  $C_P$  can be computed in the standard way by the Galerkin method; see section 3.2. The only difference is that here we compute the approximation  $\lambda_1^h$  of the second smallest (the smallest positive) eigenvalue of (5.2), because the smallest eigenvalue is  $\lambda_0 = 0$ . In order to compute the upper bound, we employ the complementarity technique as follows.

**THEOREM 5.2.** *Let  $V = \tilde{H}^1(\Omega)$ ,  $u_* \in V$ , and  $\lambda_* \in \mathbb{R}$ . Let  $w \in V$  satisfy*

$$(5.5) \quad a(w, v) = a(u_*, v) - \lambda_*(u_*, v) \quad \forall v \in V.$$

Then

$$(5.6) \quad \|w\|_a \leq \|\nabla u_* - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}} + C_P \|\lambda_* u_* + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)} \quad \forall \mathbf{q} \in W_0,$$

where  $W_0 = \{\mathbf{q} \in \mathbf{H}(\operatorname{div}, \Omega) : \mathbf{q} \cdot \mathbf{n} = 0 \text{ on } \partial\Omega\}$ .

*Proof.* By fixing arbitrary  $\mathbf{q} \in W_0$ , testing (5.5) by  $v = w$ , and using the divergence theorem, we obtain

$$\begin{aligned} \|w\|_a^2 &= (\mathcal{A} \nabla u_*, \nabla w) - \lambda_*(u_*, w) - (\mathbf{q}, \nabla w) - (\operatorname{div} \mathbf{q}, w) \\ &= (\mathcal{A}(\nabla u_* - \mathcal{A}^{-1} \mathbf{q}), \nabla w) - (\lambda_* u_* + \operatorname{div} \mathbf{q}, w). \end{aligned}$$

The Cauchy–Schwarz inequality and Poincaré inequality (5.1) yield

$$\|w\|_a^2 \leq \|\nabla u_* - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}} \|\nabla w\|_{\mathcal{A}} + C_P \|\lambda_* u_* + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)} \|w\|_a.$$

Since  $\|\nabla w\|_{\mathcal{A}} = \|w\|_a$ , the proof is finished.  $\square$

We observe that complementarity estimate (5.6) is of type (3.8) with

$$(5.7) \quad A = \|\nabla u_* - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}}, \quad B = \|\lambda_* u_* + \operatorname{div} \mathbf{q}\|_{L^2(\Omega)},$$

and  $C_\gamma = C_P$ . As soon as suitable approximations  $\lambda_*$ ,  $u_*$ , and a suitable vector field  $\mathbf{q} \in W_0$  are available, the numbers  $A$  and  $B$  can be computed and used in (3.9)–(3.10) to obtain a guaranteed upper bound on the Poincaré constant  $C_P$ .

A straightforward approach for computation of  $\lambda_*$ ,  $u_*$ , and  $\mathbf{q}$  was described in section 4.4 for the case of Friedrichs constant. It can be directly used also for the Poincaré constant. It is even simpler, because  $c = 0$ ,  $\alpha = 0$ , and  $\Gamma_N = \partial\Omega$ . The only difference is that approximations  $\lambda_1^h$  and  $u_1^h$  of the second smallest (the first positive) eigenvalue of (5.2) and its corresponding eigenvector, respectively, have to be used. In particular, the approximation  $\mathbf{q}^h \in W_0$  is computed using (4.10). This vector field is then used in (5.7) to evaluate  $A$  and  $B$  and consequently the two-sided bounds  $C_P^{\text{low}} \leq C_P \leq C_P^{\text{up}}$ , where

$$(5.8) \quad C_P^{\text{low}} = (\lambda_1^h)^{-1/2} \quad \text{and} \quad C_P^{\text{up}} = 1/X_2,$$

and  $X_2$  is given by (3.10) with  $\gamma u_* = u_1^h$ ; see also (3.3) and (3.9).

TABLE 5.1

Poincaré constant. The lower bound  $C_P^{\text{low}}$ , upper bound  $C_P^{\text{up}}$ , relative error  $E_{\text{REL}}$ , and number of degrees of freedom  $N_{\text{DOF}} = \dim V^h$  for particular values of  $\tilde{a}$ .

$\tilde{a}$	$C_P^{\text{low}}$	$C_P^{\text{up}}$	$E_{\text{REL}}$	$N_{\text{DOF}}$
0.001	14.2390	14.3690	0.0091	3 400
0.01	4.5199	4.5623	0.0093	3 510
0.1	1.4849	1.4989	0.0094	4 382
1	0.6365	0.6424	0.0092	3 009
10	0.4696	0.4740	0.0094	4 382
100	0.4520	0.4562	0.0093	3 510
1000	0.4503	0.4544	0.0091	3 400

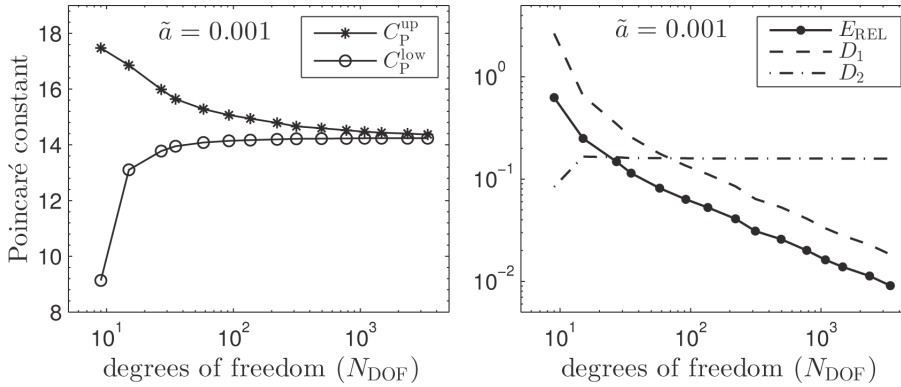


FIG. 5.1. Poincaré constant for  $\tilde{a} = 0.001$ . Convergence of bounds  $C_P^{\text{up}}$  and  $C_P^{\text{low}}$  (left) and of the relative error  $E_{\text{REL}}$  and heuristic indicators  $D_1$  and  $D_2$  (right).

**5.3. Numerical experiment.** We consider the same setting as in section 4.5. The only difference is that in the case of the Poincaré constant we assume  $\Gamma_N = \partial\Omega$  and  $\Gamma_D = \emptyset$ . We employ the adaptive algorithm as before (Algorithm 1) with clear modifications. We use error indicators (4.13), where  $\lambda_1^h$  and  $u_1^h$  are the Galerkin approximations of the second smallest (the smallest positive) eigenvalue  $\lambda_1$  of (5.2) and its corresponding eigenvector  $u_1$ , respectively. The relative error in Step 4 of Algorithm 1 is computed using two-sided bounds (5.8).

The obtained two-sided bounds on the Poincaré constant  $C_P$  for a series of values of  $\tilde{a}$  are presented in Table 5.1. The guaranteed 0.01 relative error tolerance was reached in all cases using several thousand degrees of freedom. As in the case of Friedrichs constant, we observe considerable dependence of the Poincaré constant  $C_P$  on  $\tilde{a}$ . Finally, we point out that the exact value of the Poincaré constant for  $\tilde{a} = 1$  is  $C_P = 2/\pi \approx 0.6366$ .

The progress of the adaptive algorithm is illustrated for the case  $\tilde{a} = 0.001$  in Figure 5.1, left, where the convergence of bounds  $C_P^{\text{up}}$  and  $C_P^{\text{low}}$  is shown. Figure 5.1, right, presents the relative error  $E_{\text{REL}}$  and heuristic indicators  $D_1$ ,  $D_2$ ; see the end of section 3.3. In later stages of the adaptive process the indicator  $D_1$  is considerably smaller than  $D_2$ , which indicates good confidence in the validity of assumption (3.6). Two adapted meshes are drawn in Figures 6.2–6.3 (middle).

## 6. Application to the trace inequality.

**6.1. Trace inequality and the optimal constant.** In order to apply the general theory from sections 2–3 to the case of the trace inequality, we consider the



same general symmetric second-order elliptic operator as in section 4.1. In addition we assume  $\text{meas}_{d-1} \Gamma_N > 0$  and that at least one of conditions (a)–(c) of Lemma 4.1 is satisfied. The general theory is applied with  $V = H_{\Gamma_D}^1(\Omega)$ ,  $(u, v)_V = a(u, v)$ ,  $H = L^2(\Gamma_N)$ , and  $(u, v)_H = (u, v)_{L^2(\Gamma_N)}$ . The operator  $\gamma : V \rightarrow H$  is the standard trace operator. Its compactness and other properties are provided in [21, Theorem 6.10.5]; see also [10]. The general result from Theorem 3.1 then translates as follows.

**THEOREM 6.1.** *Let the bilinear form  $a(\cdot, \cdot)$  given by (4.1) form a scalar product in  $H_{\Gamma_D}^1(\Omega)$ . Then there exists a constant  $C_T > 0$  such that*

$$(6.1) \quad \|v\|_{L^2(\Gamma_N)} \leq C_T \|v\|_a \quad \forall v \in H_{\Gamma_D}^1(\Omega).$$

Moreover, the optimal value of this constant is  $C_T = \lambda_1^{-1/2}$ , where  $\lambda_1$  is the smallest eigenvalue of the following problem: find  $u_i \in H_{\Gamma_D}^1(\Omega)$ ,  $u_i \neq 0$ , and  $\lambda_i \in \mathbb{R}$  such that

$$(6.2) \quad a(u_i, v) = \lambda_i (u_i, v)_{L^2(\Gamma_N)} \quad \forall v \in H_{\Gamma_D}^1(\Omega).$$

*Proof.* The statement follows immediately from Lemma 4.1 and Theorem 3.1.  $\square$

The common version of the trace inequality

$$\|v\|_{L^2(\partial\Omega)} \leq C_T \|v\|_{H^1(\Omega)} \quad \forall v \in H^1(\Omega)$$

then follows from (6.1) with  $\Gamma_D = \emptyset$ ,  $\Gamma_N = \partial\Omega$ ,  $\mathcal{A}$  being the identity matrix,  $c = 1$ , and  $\alpha = 0$ .

**6.2. Two-sided bounds on the trace constant.** As in the case of Friedrichs inequality, we compute the lower bound of the optimal value for the constant  $C_T$  by the Galerkin method; see section 3.2. In order to compute an upper bound on  $C_T$  we employ the complementarity technique as follows.

**THEOREM 6.2.** *Let  $V = H_{\Gamma_D}^1(\Omega)$ ,  $u_* \in V$ , and  $\lambda_* \in \mathbb{R}$ . Let the bilinear form  $a(\cdot, \cdot)$  given by (4.1) form a scalar product in  $V$ . Let  $w \in V$  satisfy*

$$(6.3) \quad a(w, v) = a(u_*, v) - \lambda_* (u_*, v)_{L^2(\Gamma_N)} \quad \forall v \in V.$$

Then, for any  $\mathbf{q} \in \mathbf{H}(\text{div}, \Omega)$

$$(6.4) \quad \|w\|_a \leq \|\nabla u_* - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}} + C_F \|cu_* - \text{div } \mathbf{q}\|_{L^2(\Omega)} + C_T \|\alpha u_* - \lambda_* u_* + \mathbf{q} \cdot \mathbf{n}\|_{L^2(\Gamma_N)}.$$

*Proof.* Let us fix any  $\mathbf{q} \in \mathbf{H}(\text{div}, \Omega)$ , test (6.3) by  $v = w$ , and use the divergence theorem to express

$$\begin{aligned} \|w\|_a^2 &= (\mathcal{A} \nabla u_*, \nabla w) + (cu_*, w) + (\alpha u_*, w)_{L^2(\Gamma_N)} - \lambda_* (u_*, w)_{L^2(\Gamma_N)} \\ &\quad - (\mathbf{q}, \nabla w) - (\text{div } \mathbf{q}, w) + (\mathbf{q} \cdot \mathbf{n}, w)_{L^2(\Gamma_N)} \\ &= (\mathcal{A}(\nabla u_* - \mathcal{A}^{-1} \mathbf{q}), \nabla w) + (cu_* - \text{div } \mathbf{q}, w) + (\alpha u_* - \lambda_* u_* + \mathbf{q} \cdot \mathbf{n}, w)_{L^2(\Gamma_N)}. \end{aligned}$$

The Cauchy–Schwarz inequality, Friedrichs inequality (4.3), and trace inequality (6.1) yield

$$\begin{aligned} \|w\|_a^2 &\leq \|\nabla u_* - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}} \|\nabla w\|_{\mathcal{A}} + C_F \|cu_* - \text{div } \mathbf{q}\|_0 \|w\|_a \\ &\quad + C_T \|\alpha u_* - \lambda_* u_* + \mathbf{q} \cdot \mathbf{n}\|_{L^2(\Gamma_N)} \|w\|_a. \end{aligned}$$

The inequality  $\|\nabla w\|_{\mathcal{A}} \leq \|w\|_a$  finishes the proof.  $\square$

As in the case of Friedrichs inequality, the bound (6.4) is of the type (3.8) with

$$(6.5) \quad A = \|\nabla u_* - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}} + C_F \|cu_* - \operatorname{div} \mathbf{q}\|_{L^2(\Omega)}, \quad B = \|\alpha u_* - \lambda_* u_* + \mathbf{q} \cdot \mathbf{n}\|_{L^2(\Gamma_N)},$$

and  $C_\gamma = C_T$ . Let us note that the complementarity estimate (6.4) is just one of several possibilities. This bound comes from [33] and contains Friedrichs constant  $C_F$ . Instead of its exact value an upper bound as computed in section 4 can be used here. However, there exist other variants of the complementarity technique that can be used to obtain a bound on  $\|w\|_a$ , avoiding the need of Friedrichs constant  $C_F$ . See, for example, [2, 12, 37, 40].

A suitable vector field  $\mathbf{q} \in \mathbf{H}(\operatorname{div}, \Omega)$  is computed by approximate minimization of the right-hand side of (6.4). In a similar way as we obtained the functional (4.9), we obtain the quadratic functional

$$(6.6) \quad \begin{aligned} & (1 + \varrho^{-1})(1 + \sigma^{-1}) \|\nabla u_1^h - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A}}^2 + (1 + \varrho)(1 + \sigma^{-1})(C_F^{\text{up}})^2 \|cu_1^h - \operatorname{div} \mathbf{q}\|_{L^2(\Omega)}^2 \\ & + (1 + \sigma)(\lambda_1^h)^{-1} \|\alpha u_1^h - \lambda_1^h u_1^h + \mathbf{q} \cdot \mathbf{n}\|_{L^2(\Gamma_N)}^2 \quad \forall \varrho > 0, \sigma > 0, \end{aligned}$$

where  $C_F^{\text{up}}$  is an upper bound of  $C_F$  computed as described in sections 4.3–4.4,  $\lambda_1^h$  is the approximation of the smallest eigenvalue of (6.2) obtained by the Galerkin method, and  $u_1^h$  is the corresponding approximate eigenvector. We look for the minimum of (6.6) over a finite dimensional subspace  $W^h$  of  $\mathbf{H}(\operatorname{div}, \Omega)$ . This minimization problem is equivalent to seeking  $\mathbf{q}^h \in W^h$  such that

$$\mathcal{B}(\mathbf{q}^h, \mathbf{w}) = \mathcal{F}(\mathbf{w}) \quad \forall \mathbf{w} \in W^h,$$

where

$$\begin{aligned} \mathcal{B}(\mathbf{q}, \mathbf{w}) &= \frac{1 + \varrho}{\sigma} \lambda_1^h (C_F^{\text{up}})^2 (\operatorname{div} \mathbf{q}, \operatorname{div} \mathbf{w}) + \frac{1 + \varrho}{\varrho \sigma} \lambda_1^h (\mathcal{A}^{-1} \mathbf{q}, \mathbf{w}) + (\mathbf{q} \cdot \mathbf{n}, \mathbf{w} \cdot \mathbf{n})_{L^2(\Gamma_N)}, \\ \mathcal{F}(\mathbf{w}) &= \frac{1 + \varrho}{\varrho \sigma} \lambda_1^h (\nabla u_1^h, \mathbf{w}) + \frac{1 + \varrho}{\sigma} \lambda_1^h (C_F^{\text{up}})^2 (cu_1^h, \operatorname{div} \mathbf{w}) \\ &+ (\lambda_1^h u_1^h - \alpha u_1^h, \mathbf{w} \cdot \mathbf{n})_{L^2(\Gamma_N)}. \end{aligned}$$

Practically, the classical Raviart–Thomas–Nédélec finite element method [13] can be used to solve this problem. Note that the natural values for  $\varrho$  and  $\sigma$  are  $\varrho = 1$  and  $\sigma = 2$ , because then  $(1 + \varrho^{-1})(1 + \sigma^{-1}) = (1 + \varrho)(1 + \sigma^{-1}) = (1 + \sigma) = 3$ ; see (6.6). Similarly to the case of Friedrichs constant, these natural values often yield accurate results. If not, a simultaneous minimization of (6.6) with respect to  $\varrho > 0$ ,  $\sigma > 0$ , and  $\mathbf{q} \in W^h$  can be performed.

The computed vector field  $\mathbf{q}^h \in W^h$  together with  $\lambda_* = \lambda_1^h$  and  $u_* = u_1^h$  are then substituted to (6.5) in order to evaluate  $A$  and  $B$ . Consequently, we obtain the two-sided bounds  $C_T^{\text{low}} \leq C_T \leq C_T^{\text{up}}$ , where

$$(6.7) \quad C_T^{\text{low}} = (\lambda_1^h)^{-1/2} \quad \text{and} \quad C_T^{\text{up}} = 1/X_2,$$

and  $X_2$  is given by (3.10) with  $\gamma u_* = \gamma u_1^h$ ; see (3.3) and (3.9).

**6.3. Numerical experiment.** In order to illustrate the numerical performance of the above described method, we consider the same example as in section 4.5. We proceed in the same way as in section 4.5 with clear modifications in order to compute

TABLE 6.1

Trace constant. The lower bound  $C_T^{\text{low}}$ , upper bound  $C_T^{\text{up}}$ , relative error  $E_{\text{REL}}$ , and number of degrees of freedom  $N_{\text{DOF}} = \dim V^h$  for particular values of  $\tilde{a}$ .

$\tilde{a}$	$C_T^{\text{low}}$	$C_T^{\text{up}}$	$E_{\text{REL}}$	$N_{\text{DOF}}$
0.001	17.8110	17.9760	0.0092	5 523
0.01	5.6490	5.7047	0.0098	5 418
0.1	1.8433	1.8593	0.0086	7 775
1	0.7963	0.8033	0.0088	5 499
10	0.5829	0.5880	0.0086	7 775
100	0.5649	0.5705	0.0098	5 421
1000	0.5632	0.5685	0.0092	5 523

two-sided bounds (6.7) on the trace constant. As before, the adaptive algorithm is steered by error indicators that are in this case defined by a localized version of (6.6):

$$\begin{aligned} \eta_K^2 &= (1 + \varrho^{-1})(1 + \sigma^{-1}) \|\nabla u_1^h - \mathcal{A}^{-1} \mathbf{q}\|_{\mathcal{A},K}^2 \\ &\quad + (1 + \varrho)(1 + \sigma^{-1})(C_F^{\text{up}})^2 \|cu_1^h - \text{div } \mathbf{q}\|_{L^2(K)}^2 \\ &\quad + (1 + \sigma)(\lambda_1^h)^{-1} \|\alpha u_1^h - \lambda_1^h u_1^h + \mathbf{q} \cdot \mathbf{n}\|_{L^2(\partial K \cap \Gamma_N)}^2 \quad \forall K \in \mathcal{T}_h. \end{aligned}$$

The parameters  $\varrho$  and  $\sigma$  are naturally chosen as  $\varrho = 1$  and  $\sigma = 2$ . The values of the upper bound  $C_F^{\text{up}}$  are taken from Table 4.1.

The obtained results are presented in Table 6.1. The method succeeded in obtaining guaranteed two-sided bounds on the trace constant with a relative error at most 0.01 in all cases using several thousand degrees of freedom. The particular value of the trace constant  $C_T$  depends considerably on  $\tilde{a}$ . The values for  $\tilde{a} = 0.001$  and  $\tilde{a} = 1000$  differ more than 30 times. Further, notice that the exact value of the trace constant for  $\tilde{a} = 1$  is  $C_T = (2/(\pi \coth \pi))^{1/2} \approx 0.7964$ .

As before, we illustrate the adaptive process for the case  $\tilde{a} = 0.001$ . Figure 6.1 presents the convergence of bounds  $C_T^{\text{up}}$  and  $C_T^{\text{low}}$  (left panel) and the relative error  $E_{\text{REL}}$  with heuristic indicators  $D_1, D_2$  (right panel). Good confidence in the validity of assumption (3.6) stems from the fact that  $D_1$  is several times smaller than  $D_2$  in later stages of the adaptive process. Figures 6.2–6.3 (right) show two of the adapted meshes.

Let us note that Tables 4.1, 5.1, and 6.1 show the highest numbers of degrees of freedom for  $\tilde{a} = 0.1$  and 10. This is probably a coincidence caused by the fact that

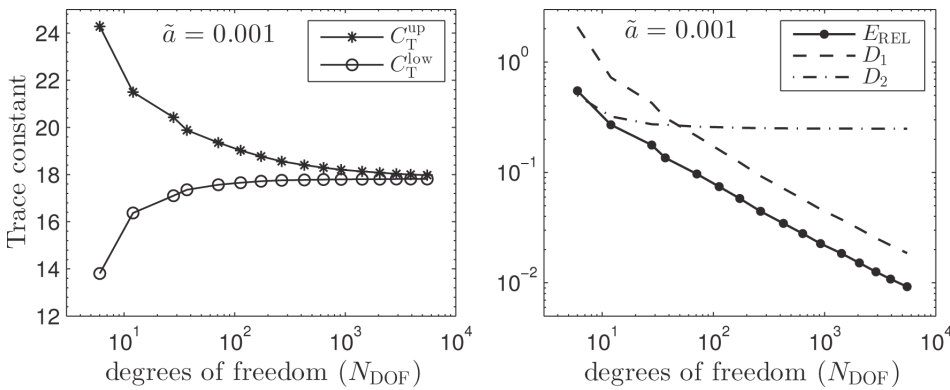


FIG. 6.1. Trace constant for  $\tilde{a} = 0.001$ . Convergence of bounds  $C_T^{\text{up}}$  and  $C_T^{\text{low}}$  (left) and of the relative error  $E_{\text{REL}}$  and heuristic indicators  $D_1$  and  $D_2$  (right).

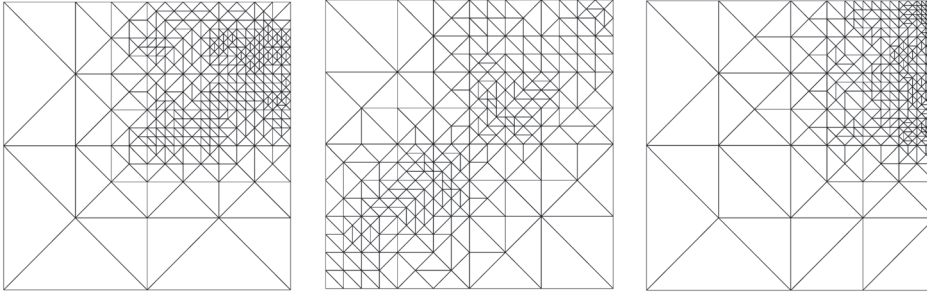


FIG. 6.2. Meshes in the middle of the adaptive process for  $\tilde{\alpha} = 0.001$ . From left to right: Friedrichs (adaptive step 8), Poincaré (adaptive step 7), trace constant (adaptive step 8).

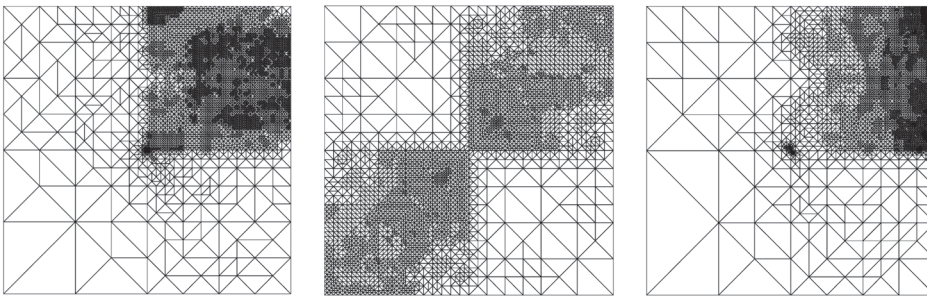


FIG. 6.3. Meshes in the final adaptive step for  $\tilde{\alpha} = 0.001$ . From left to right: Friedrichs (adaptive step 16), Poincaré (adaptive step 14), trace constant (adaptive step 15).

the error decreases in jumps after each mesh refinement. For example, if  $N_{\text{DOF}}$  is about 5000, then the relative error is already close to the threshold  $E_{\text{REL}} = 0.01$ . If it is slightly below the threshold, we stop the algorithm, but if it is slightly above, we have to refine the mesh one more time, which results in higher  $N_{\text{DOF}}$  and also a smaller  $E_{\text{REL}}$ .

**7. Conclusions.** We present a method for computing guaranteed lower and upper bounds of principal eigenvalues of elliptic operators and consequently for computing guaranteed two-sided bounds of the optimal constants in Friedrichs, Poincaré, trace, and similar inequalities. The bounds are guaranteed provided there are no round-off errors and all integrals are evaluated exactly. Further, the bounds are guaranteed only if the domain  $\Omega$  is represented exactly by used finite elements. Furthermore, the upper bounds on eigenvalues computed by the Galerkin method are guaranteed only if the corresponding matrix-eigenvalue problems are solved exactly. On the other hand, the lower bounds on eigenvalues obtained by the complementarity technique are guaranteed even if the matrix-eigenvalue problems and linear algebraic systems are solved approximately only. In any case, the crucial assumption for having guaranteed lower bounds on eigenvalues is (3.6).

These two-sided bounds can be of interest if the corresponding eigenvalue problem cannot be solved analytically and if analytical estimates are not available or they are too inaccurate. In particular, this is the case of complicated geometry of the domain  $\Omega$ , mixed boundary conditions, presence of nonconstant and/or anisotropic diffusion coefficient  $\mathcal{A}$ , presence of reaction coefficient  $c$ , and presence of coefficient  $\alpha$ .

The method is quite general and it can be used for a wide variety of problems. The general Hilbert space setting presented in section 2 enables a variety of applications including linear elasticity. We believe that this approach can be further generalized. Nonlinear eigenvalue problems and nonsymmetric operators can be of particular interest and generalizations in these directions can be a subject for further research.

## REFERENCES

- [1] R. A. ADAMS AND J. J. F. FOURNIER, *Sobolev Spaces*, Elsevier, Amsterdam, 2003.
- [2] M. AINSWORTH AND T. VEJCHODSKÝ, *Fully computable robust a posteriori error bounds for singularly perturbed reaction-diffusion problems*, Numer. Math., 119 (2011), pp. 219–243.
- [3] A. ALONSO, A. D. RUSSO, C. PADRA, AND R. RODRÍGUEZ, *A posteriori error estimates and a local refinement strategy for a finite element method to solve structural-acoustic vibration problems*, Adv. Comput. Math., 15 (2001), pp. 25–59.
- [4] A. ANDREEV AND M. RACHEVA, *Two-sided bounds of eigenvalues*, Appl. Math., to appear.
- [5] I. BABUŠKA AND J. E. OSBORN, *Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems*, Math. Comp., 52 (1989), pp. 275–297.
- [6] I. BABUŠKA AND J. E. OSBORN, *Eigenvalue Problems*, Handb. Numer. Anal. 2, North-Holland, Amsterdam, 1991, pp. 641–787.
- [7] Z. BAI, J. DEMMEL, J. DONGARRA, A. RUHE, AND H. VAN DER VORST EDS., *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, Software Environ. Tools 11, SIAM, Philadelphia, 2000.
- [8] N. W. BAZLEY AND D. W. FOX, *Lower bounds for eigenvalues of Schrödinger’s equation*, Phys. Rev., 124 (1961), pp. 483–492.
- [9] C. BEATTIE AND F. GOERISCH, *Methods for computing lower bounds to eigenvalues of self-adjoint operators*, Numer. Math., 72 (1995), pp. 143–172.
- [10] M. BIEGERT, *On traces of Sobolev functions on the boundary of extension domains*, Proc. Amer. Math. Soc., 137 (2009), pp. 4169–4176.
- [11] D. BOFFI, *Finite element approximation of eigenvalue problems*, Acta Numer., 19 (2010), pp. 1–120.
- [12] D. BRAESS AND J. SCHÖBERL, *Equilibrated residual error estimator for edge elements*, Math. Comp., 77 (2008), pp. 651–672.
- [13] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer-Verlag, New York, 1991.
- [14] I. CHEDDADI, R. FUČÍK, M. I. PRIETO, AND M. VOHRALÍK, *Guaranteed and robust a posteriori error estimates for singularly perturbed reaction-diffusion problems*, M2AN Math. Model. Numer. Anal., 43 (2009), pp. 867–888.
- [15] E. K.-W. CHU, T.-M. HUANG, W.-W. LIN, AND C.-T. WU, *Palindromic eigenvalue problems: A brief survey*, Taiwanese J. Math., 14 (2010), pp. 743–779.
- [16] D. W. FOX AND W. C. RHEINBOLDT, *Computational methods for determining lower bounds for eigenvalues of operators in Hilbert spaces*, SIAM Rev., 8 (1966), pp. 427–462.
- [17] S. A. GAAL, *Linear Analysis and Representation Theory*, Springer-Verlag, New York, 1973.
- [18] J. HASLINGER AND I. HLAVÁČEK, *Convergence of a finite element method based on the dual variational formulation*, Apl. Mat., 21 (1976), pp. 43–65.
- [19] J. HU, Y. HUANG, AND Q. LIN, *The Lower Bounds for Eigenvalues of Elliptic Operators—by Nonconforming Finite Element Methods*, arXiv:1112.1145v2, 2013.
- [20] S. KOROTOV, *Two-sided a posteriori error estimates for linear elliptic problems with mixed boundary conditions*, Appl. Math., 52 (2007), pp. 235–249.
- [21] A. KUFNER, O. JOHN, AND S. FUČÍK, *Function Spaces*, Noordhoff, Leyden, Netherlands, 1977.
- [22] J. R. KUTTNER AND V. G. SIGILLITO, *Bounding eigenvalues of elliptic operators*, SIAM J. Math. Anal., 9 (1978), pp. 768–778.
- [23] J. R. KUTTNER AND V. G. SIGILLITO, *Eigenvalues of the Laplacian in two dimensions*, SIAM Rev., 26 (1984), pp. 163–193.
- [24] R. G. LARSON, *Instabilities in viscoelastic flows*, Rheologica Acta, 31 (1992), pp. 213–263.
- [25] R. B. LEHOUCQ, D. C. SORENSEN, AND C. YANG, *ARPACK Users’ Guide*, Software Environ. Tools, SIAM, Philadelphia, 1998.
- [26] A. W. LEISSA, *Vibration of Plates*, NASA, Washington, DC, 1969.
- [27] Q. LIN, H. XIE, F. LUO, Y. LI, AND Y. YANG, *Stokes eigenvalue approximations from below with nonconforming mixed finite element methods*, Math. Pract. Theory, 40 (2010), pp. 157–168.

- [28] Q. LIN, H. XIE, AND J. XU, *Lower bounds of the discretization error for piecewise polynomials*, Math. Comp., 83 (2014), pp. 1–13.
- [29] F. LUO, Q. LIN, AND H. XIE, *Computing the lower and upper bounds of Laplace eigenvalue problem: by combining conforming and nonconforming finite element methods*, Sci. China Math., 55 (2012), pp. 1069–1082.
- [30] R. M. MARTIN, *Electronic Structure. Basic Theory and Practical Methods*, Cambridge University Press, Cambridge, UK, 2004.
- [31] H. OGAWA AND M. H. PROTTER, *A lower bound for the first eigenvalue of second order elliptic operators*, Proc. Amer. Math. Soc., 19 (1968), pp. 292–295.
- [32] R. RANNACHER, *Nonconforming finite element methods for eigenvalue problems in linear plate theory*, Numer. Math., 33 (1979), pp. 23–42.
- [33] S. REPIN, *A Posteriori Estimates for Partial Differential Equations*, Walter de Gruyter, Berlin, 2008.
- [34] W. RUDIN, *Functional Analysis*, 2nd ed., McGraw-Hill, New York, 1991.
- [35] V. G. SIGILLITO, *Explicit a priori inequalities with applications to boundary value problems*, Pitman, Boston, 1977.
- [36] G. STRANG AND G. J. FIX, *An Analysis of the Finite Element Methods*, 2nd ed., Wellesley-Cambridge Press, Wellesley, MA, 2008.
- [37] T. VEJCHODSKÝ, *Guaranteed and locally computable a posteriori error estimate*, IMA J. Numer. Anal., 26 (2006), pp. 525–540.
- [38] T. VEJCHODSKÝ, *Complementarity based a posteriori error estimates and their properties*, Math. Comput. Simulation, 82 (2012), pp. 2033–2046.
- [39] T. VEJCHODSKÝ, *Computing upper bounds on Friedrichs' constant*, in Applications of Mathematics 2012, J. Brandts, J. Chleboun, S. Korotov, K. Segeth, J. Šístek, and T. Vejchodský, eds., Institute of Mathematics, ASCR, Prague, 2012, pp. 278–289.
- [40] M. VOHRALÍK, *Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients*, J. Sci. Comput., 46 (2011), pp. 397–438.
- [41] A. WEINSTEIN AND W. STENGER, *Methods of Intermediate Problems for Eigenvalues. Theory and Ramifications*, Academic Press, New York, 1972.
- [42] Y. YANG, Z. ZHANG, AND F. LIN, *Eigenvalue approximation from below using non-conforming finite elements*, Sci. China Math., 53 (2010), pp. 137–150.
- [43] K. YOSIDA, *Functional Analysis*, 6th ed., Springer-Verlag, New York, 1994.