Numer. Math. (2011) 118:401–427
DOI 10.1007/s00211-011-0367-2

Numerische Mathematik

An oscillation-free adaptive FEM for symmetric eigenvalue problems

Carsten Carstensen · Joscha Gedicke

Received: 12 March 2008 / Published online: 15 May 2011
© Springer-Verlag 2011

Abstract A refined a posteriori error analysis for symmetric eigenvalue problems and the convergence of the first-order adaptive finite element method (AFEM) is presented. The H^1 stability of the L^2 projection provides reliability and efficiency of the edge-contribution of standard residual-based error estimators for P_1 finite element methods. In fact, the volume contributions and even oscillations can be omitted for Courant finite element methods. This allows for a refined averaging scheme and so improves (Mao et al. in Adv Comput Math 25(1–3):135–160, 2006). The proposed AFEM monitors the edge-contributions in a bulk criterion and so enables a contraction property up to higher-order terms and global convergence. Numerical experiments exploit the remaining L^2 error contributions and confirm our theoretical findings. The averaging schemes show a high accuracy and the AFEM leads to optimal empirical convergence rates.

Mathematics Subject Classification (2000) 65N12 · 65N25 · 65N30 · 65N50

Supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin and the Hausdorff Institute for Mathematics in Bonn, Germany.

C. Carstensen (✉) · J. Gedicke
Institut für Mathematik, Humboldt-Universität zu Berlin,
Unter den Linden 6, 10099 Berlin, Germany
e-mail: cc@mathematik.hu-berlin.de

J. Gedicke
e-mail: gedicke@mathematik.hu-berlin.de

C. Carstensen
Department of Computational Science and Engineering,
Yonsei University, Seoul 120-749, Korea

1 Introduction

While error estimates for adaptive methods for space and time dependent PDEs have been studied in great detail in recent years, error estimates and adaptive algorithms for eigenvalue problems are still under development. A priori error estimates for elliptic operators [4,5,12,16,19,21–23] assume that the mesh-size is sufficiently small. Knyazev and Osborn [17] overcame this difficulty and presented the first truly a priori error estimate for symmetric eigenvalue problems.

The a posteriori error analysis for symmetric second order elliptic eigenvalue problems started with Verfürth [24] and Larson [18] for L^2 and H^1 error estimates based on duality. An energy-based technique due to Durán et al. [13] controls the error by some edge and volume residual plus a higher-order term. This paper will provide a refinement without the volume contribution for all eigenvalues which generalises and simplifies the proof in [13]. Mao et al. [20] suggested some local averaging technique which we improve by neglecting the volume contributions. The first convergence of an adaptive algorithm with oscillation terms can be found in [14], which we further develop here for a refined adaptive scheme.

Nonsymmetric elliptic eigenvalue problems are analysed by Heuveline and Rannacher in [6,15] and lay beyond the scope of this paper.

Throughout this paper, we study the following general formulation. The weak form of the symmetric eigenvalue problem involves two real Hilbert spaces (V, a) and (H, b) with $V \subset H \subset V^*$. The scalar products a and b induce norms in respective spaces, namely

$$|||.||| := a(.,.)^{1/2} \quad \text{and} \quad \|.\| := b(.,.)^{1/2},$$

and the embedding of V in H is continuous and compact,

$$V \overset{c}{\hookrightarrow} H.$$

The *continuous eigenvalue problem* consists in finding a pair (λ, u) of $\lambda \in \mathbb{R}$ (actually $\lambda > 0$) and $u \in V$ with $\|u\| = 1$ and

$$a(u, v) = \lambda\, b(u, v) \quad \text{for all } v \in V. \tag{1.1}$$

Given any finite-dimensional subspace V_ℓ of V, the *discrete eigenvalue problem* consists in finding $(\lambda_\ell, u_\ell) \in \mathbb{R} \times V_\ell$ with $\|u_\ell\| = 1$ and

$$a(u_\ell, v_\ell) = \lambda_\ell\, b(u_\ell, v_\ell) \quad \text{for all} \quad v_\ell \in V_\ell. \tag{1.2}$$

Throughout this paper, the min-max principle [23] allows some ordering of the discrete eigenvalues with $0 \le \lambda \le \lambda_\ell$.

Typical examples for eigenvalue problems include the Poisson problem

$$-\Delta u = \lambda u \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Omega$$

(for the Laplace operator Δ) and the Lamé problem

$$-\Delta^* u = \lambda \rho u \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Omega$$

from harmonic dynamic of linear elasticity (with the Lamé operator Δ^* and the density ρ).

Given an initial coarse mesh $\mathcal{T}_0$, an adaptive finite element method (AFEM) successively generates a sequence of meshes $\mathcal{T}_1, \mathcal{T}_2, \ldots$ and associated discrete subspaces

$$V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_\ell \subsetneq V_{\ell+1} \subsetneq \cdots \subsetneq V$$

with discrete solutions consisting of discrete eigenpairs (λ_ℓ, u_ℓ). A typical loop from V_ℓ to $V_{\ell+1}$ (at frozen level ℓ) consists of the steps

$$\text{SOLVE} \to \text{ESTIMATE} \to \text{MARK} \to \text{REFINE}. \tag{1.3}$$

This paper contributes to the a posteriori error analysis [13,20,25] of eigenvalue problems and to the design and convergence of AFEM [14]. Here we give a shorter proof of the edge-residual estimator in [13] and improve the results from [20], in the sense that in the estimator no additional volumetric part is needed. Additionally, we show that the higher-order terms can really be neglected and underline that by numerical experiments. In contrast to [14] we proof the convergence of AFEM without the inner node property. Our global convergence proof seems to be the first that does not need the usual assumption that the mesh size is small enough.

The outline of the remainder of this paper is as follows. Section 2 describes an adaptive mesh-refinement algorithm that allows for the H^1 stability of the L^2 projection. In Sect. 3, the algebraic aspects of the a posteriori error analysis are provided. Section 4 presents the edge residual and the refined averaging technique. Section 5 analyses the convergence of the AFEM illustrated in Sect. 6 by numerical experiments.

2 Adaptive mesh refinement algorithm

This section describes the algorithm REFINE of one loop of AFEM from (1.3) in order to state precisely conditions for the H^1 stable L^2 projection required below.

2.1 Input: assumptions on course triangulation $\mathcal{T}_0$

The initial mesh $\mathcal{T}_0$ is a *regular triangulation* of $\Omega \subset \mathbb{R}^n$ into closed triangles in the sense that two distinct closed-element domains are either disjoint or their intersection is one common vertex or one common edge. We suppose that each element with domain in $\mathcal{T}_0$ has at least one vertex in the interior of Ω.

Given any $T \in \mathcal{T}_0$, one chooses one of its edges $E(T)$ as a *reference edge* from the set of Edges $\mathcal{E}(T)$ such that the following holds. An element $T \in \mathcal{T}_0$ is called *isolated* if $E(T)$ either belongs to the boundary $\partial\Omega$ or equals the side of another element $K \in \mathcal{T}_0$ with $E(T) = \partial T \cap \partial K \neq E(K)$. Given a regular triangulation $\mathcal{T}_0$,

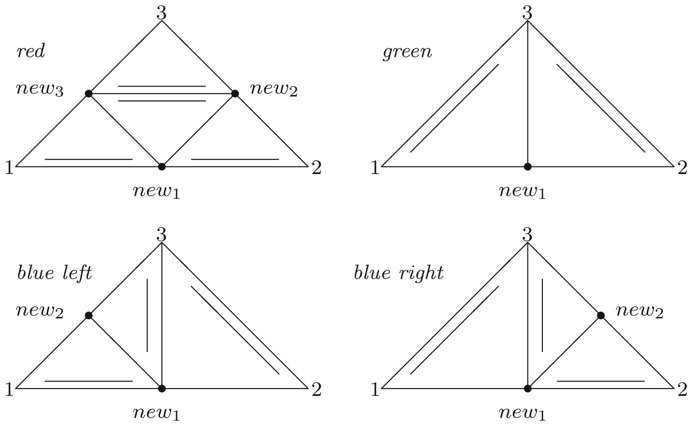

Fig. 1 *Red*, *green* and *blue* refinement. The new reference edge is marked through a second line in parallel opposite the new vertices new_1, new_2 or new_3

Algorithm 2.1 of [8] computes the reference edges $(E(T) : T \in \mathcal{T}_0)$ such that two distinct isolated triangles do not share an edge. This is important for the H^1 stability of the L^2 projection in Sect. 2.4.

2.2 Red–green–blue refinements

Given a triangulation $\mathcal{T}_\ell$ on the level ℓ, let $\mathcal{E}_\ell$ denote its set of interior edges and suppose that $E(T)$ $(E(T) : T \in \mathcal{T}_\ell)$ denotes the given reference edges. There is no need to label the reference edges $E(T)$ by some level ℓ because $E(T)$ will be the same edge of T in all triangulations $\mathcal{T}_m$ which include T. However, once T in $\mathcal{T}_\ell$ is refined, the reference edges will be specified for the sub-triangles as indicated in Fig. 1. The mesh-refinement strategy consists of the following five different refinements. Elements with no marked edge are not refined, elements with one marked edge are refined *green*, elements with two marked edges are refined *blue*, and elements with three marked edges are refined *red*.

2.3 Marking and closure

The set of refined edges $\mathcal{M}_\ell \subset \mathcal{E}_\ell$ is specified in the algorithm MARK. The closure algorithm computes the smallest subset $\widehat{\mathcal{M}}_\ell$ of $\mathcal{E}_\ell$ which includes $\mathcal{M}_\ell$ such that

$$\left\{E(T) : T \in \mathcal{T} \quad \text{with} \quad \mathcal{E}(T) \cap \widehat{\mathcal{M}}_\ell \neq \emptyset\right\} \subseteq \widehat{\mathcal{M}}_\ell.$$

In other words, once an edge E of an element T is marked for refinement (written $E \in \widehat{\mathcal{M}}_\ell$), the reference edge $E(T)$ of T is marked as well. Consequently, each element has either $k = 0, 1, 2,$ or 3 of its edges marked for refinement, if $k \geq 1$, the reference edge belongs to it. Therefore, exactly one of the five refinement rules of Fig. 1 is applied. This specifies sub-triangles and their reference edges in the new triangulation $\mathcal{T}_{\ell+1}$.

2.4 Properties of the triangulations

This subsection lists a few results on the triangulation $\mathcal{T}_\ell$ obtained by REFINE under the assumptions on $\mathcal{T}_0$ of Sect. 2.1. The non-elementary proofs can be found in [8].

(i) $\mathcal{T}_\ell$ is a regular triangulation of Ω into triangles; for each $T \in \mathcal{T}_\ell$ there exists one reference edge $E(T)$ which depends only on T but not on the level ℓ.

(ii) For each $K \in \mathcal{T}_0$, $\mathcal{T}_\ell|_K := \{T \in \mathcal{T}_\ell \,|\, T \subseteq K\}$ is the picture under an affine map $\Phi : K \to T_{ref}$ onto the reference triangle $T_{ref} = \mathrm{conv}\{(0,0), (0,1), (1,0)\}$ by $\Phi(E(K)) = \mathrm{conv}\{(0,0), (1,0)\}$ and $\det D\Phi > 0$. The triangulation $\widehat{T}_K := \{\Phi(T) : T \in \mathcal{T}, T \subseteq K\}$ of K consists of right isosceles triangles. (A right isosceles triangle results from a square halved along a diagonal.)

(iii) The L^2 projection Π onto $V_\ell := \mathcal{P}_1(\mathcal{T}_\ell) \cap V$ is H^1 stable. The piecewise affine space are defined by

$$\mathcal{P}_1(T; \mathbb{R}^m) := \left\{ v \in C^\infty(T; \mathbb{R}^m) : v \text{ affine on } T \right\},$$
$$\mathcal{P}_1(\mathcal{T}_\ell; \mathbb{R}^m) := \left\{ v \in L^\infty(\Omega; \mathbb{R}^m) : \forall T \in \mathcal{T}_\ell, v|_T \in \mathcal{P}_1(T; \mathbb{R}^m) \right\}.$$

For any $v \in V := H_0^1(\Omega)$ the L^2 projection Πv on V_ℓ satisfies

$$\|\nabla \Pi v\|_{L^2(\Omega)} \leq C_{stab} \|\nabla v\|_{L^2(\Omega)}.$$

(iv) The approximation property of the L^2 projection states

$$\sum_{T \in \mathcal{T}_\ell} \|h_T^{-1}(v - \Pi v)\|^2_{L^2(T)} + \sum_{E \in \mathcal{E}_\ell} \|h_E^{-1/2}(v - \Pi v)\|^2_{L^2(E)} \leq C_{app} \|\nabla v\|^2_{L^2(\Omega)}$$

for all $v \in V$. The constants C_{stab} and C_{app} depend exclusively on $\mathcal{T}_0$.

3 Algebraic aspects of an a posteriori error analysis

Throughout this section, (λ, u) solves (1.1) and (λ_ℓ, u_ℓ) solves (1.2). Suppose that the orientation of the unit vectors u and u_ℓ is normalised to $b(u, u_\ell) \geq 0$. Set $e_\ell := u - u_\ell$ and

$$Res_\ell := \lambda_\ell b(u_\ell, \cdot) - a(u_\ell, \cdot) \in V^*$$

such that $V_\ell \subset \ker(Res_\ell)$.

Lemma 3.1 *Let* (λ, u) *and* (λ_ℓ, u_ℓ) *be eigenpairs of* (1.1) *and* (1.2). *Then it holds*

$$|||e_\ell|||^2 = \lambda \|e_\ell\|^2 + \lambda_\ell - \lambda = (\lambda + \lambda_\ell)\|e_\ell\|^2/2 + Res_\ell(e_\ell).$$

Proof The first identity follows from

$$\begin{aligned} a(e_\ell, e_\ell) &= \lambda_\ell + \lambda - 2a(u, u_\ell) \\ &= \lambda_\ell - \lambda + 2\lambda(1 - b(u, u_\ell)) \\ &= \lambda_\ell - \lambda + \lambda b(e_\ell, e_\ell) \end{aligned}$$

and the second follows from

$$\begin{aligned} a(e_\ell, e_\ell) &= a(u, e_\ell) + a(u_\ell, u_\ell) - a(u_\ell, u) \\ &= \lambda b(u, e_\ell) + \lambda_\ell b(u_\ell, u_\ell) - a(u_\ell, u) \\ &= b(\lambda u - \lambda_\ell u_\ell, e_\ell) + \lambda_\ell b(u_\ell, u) - a(u_\ell, u) \\ &= b(\lambda u - \lambda_\ell u_\ell, e_\ell) + Res_\ell(u) \\ &= (\lambda + \lambda_\ell)\,(1 - b(u, u_\ell)) + Res_\ell(e_\ell) \\ &= \frac{\lambda + \lambda_\ell}{2} b(e_\ell, e_\ell) + Res_\ell(e_\ell). \end{aligned}$$

□

For the discussion of $\|e_\ell\| \ll |||e_\ell|||$, suppose that the eigenvalues and the $N_\ell = \dim(V_\ell)$ discrete eigenvalues are enumerated

$$0 < \lambda_1 \le \lambda_2 \le \cdots \qquad \text{and} \qquad 0 < \lambda_{\ell,1} \le \cdots \le \lambda_{\ell,N_\ell}.$$

Let $(u_1, u_2, u_3, \ldots)$ and $(u_{\ell,1}, \ldots, u_{\ell,N_\ell})$ denote some b-orthonormal basis of V and V_ℓ of corresponding eigenfunctions. Suppose that there exist a cluster of eigenvalues $\lambda_{n+1} \le \cdots \le \lambda_{n+m}$ of multiplicity $m \in \mathbb{N}$ with eigenspace $W := \operatorname{span}\{u_{n+1}, \ldots, u_{n+m}\}$. Define their index set $I := \{n+1, \ldots, n+m\}$ and the complement $N_\ell(I) := \{1, \ldots, N_\ell\} \backslash I$. The minmax principle and known a priori error estimates [23] show for some sufficiently small global mesh-size h_0 that there exists some separation bound

$$0 < M_1(I) := \sup_{\ell \in \mathbb{N}_0} \max_{j \in N_\ell(I)} \max_{k \in I} \frac{\lambda_k}{|\lambda_{\ell,j} - \lambda_k|} < \infty.$$

Let $W_\ell := \operatorname{span}\{u_{\ell,n+1}, \ldots, u_{\ell,n+m}\}$ and set $\operatorname{dist}_{\|.\|}(v, W_\ell) := \min\{\|v - w_\ell\| : w_\ell \in W_\ell\}$. In the following, the map $P : V \to W$ denotes the b-orthogonal projection onto W, $b(Pv - v, \cdot)|_W = 0$ for all $v \in V$, $P_\ell : V \to W_\ell$ the b-orthogonal projection onto W_ℓ, $b(P_\ell v - v, \cdot)|_{W_\ell} = 0$ for all $v \in V$, and $G_\ell : V \to V_\ell$ the Galerkin projection, $a(G_\ell v - v, \cdot)|_{V_\ell} = 0$ for all $v \in V$.

Proposition 3.1 *Let $u_k \in W$ be some b-normalised eigenfunction to the k-th eigenvalue λ_k with $k \in I$. Then it holds*

$$dist_{\|.\|}(G_\ell u_k, W_\ell) \le M_1(I)\|u_k - G_\ell u_k\|.$$

Proof Set $v := G_\ell u_k - P_\ell(G_\ell u_k)$ for the b-orthogonal projection P_ℓ onto W_ℓ. Then $\mathrm{dist}_{\|.\|}(G_\ell u_k, W_\ell) = \|v\|$ with $v := \sum_{j\in N_\ell(I)} \alpha_j u_{\ell,j}$ and $W_\ell \perp \mathrm{span}\{u_j : j \in N_\ell(I)\}$ implies

$$b\left(P_\ell(G_\ell u_k), \sum_{j\in N_\ell(I)} \alpha_j u_{\ell,j}\right) = 0.$$

The pairwise b-orthogonality of the basis functions $u_{\ell,1}, \ldots, u_{\ell,N_\ell}$ yields

$$\left\| \sum_{j\in N_\ell(I)} \alpha_j \frac{\lambda_k}{\lambda_{\ell,j} - \lambda_k} u_{\ell,j} \right\|^2 = \sum_{j\in N_\ell(I)} \left(\frac{\lambda_k}{\lambda_{\ell,j} - \lambda_k}\right)^2 \alpha_j^2 \le M_1^2(I)\|v\|^2.$$

The Galerkin orthogonality $a(G_\ell u_k, u_{\ell,j}) = a(u_k, u_{\ell,j})$ for all $j = 1, \ldots, N_\ell$ shows

$$(\lambda_{\ell,j} - \lambda_k) b(G_\ell u_k, u_{\ell,j}) = \lambda_k b(u_k - G_\ell u_k, u_{\ell,j}),$$

because $\lambda_k b(G_\ell u_k, u_{\ell,j})$ occurs on both sides [23, Lemma 6.4]. This, some algebra, and elementary estimates show

$$\|v\|^2 = b\left(G_\ell u_k, \sum_{j\in N_\ell(I)} \alpha_j u_{\ell,j}\right) = b\left(u_k - G_\ell u_k, \sum_{j\in N_\ell(I)} \alpha_j \frac{\lambda_k}{\lambda_{\ell,j} - \lambda_k} u_{\ell,j}\right).$$

Therefore,

$$\|v\|^2 \le \|u_k - G_\ell u_k\| \left\| \sum_{j\in N_\ell(I)} \alpha_j \frac{\lambda_k}{\lambda_{\ell,j} - \lambda_k} u_{\ell,j} \right\| \le M_1(I)\|u_k - G_\ell u_k\|\|v\|.$$

□

Proposition 3.2 *Let* $(\lambda_{\ell,k}, u_{\ell,k})$ *denote some discrete eigenpair number* $k \in I$ *and let* $Pu_{\ell,k} = \|Pu_{\ell,k}\|u_k^*$ *for some* $u_k^* \in W$ *with* $\|u_k^*\| = 1$. *Then it holds*

$$\frac{\|u_k^* - u_{\ell,k}\|^2}{2} = \frac{dist_{\|.\|}(u_{\ell,k}, W)^2}{1 + \|Pu_{\ell,k}\|} \le M_2^2(I) \max_{j\in I} \|u_j - G_\ell u_j\|^2$$

with $M_2(I) := m(2m + 1)(1 + M_1(I))$.

Proof Notice that for $e_\ell^* := u_k^* - u_{\ell,k}$, $\|e_\ell^*\|^2 = \|e_\ell^* - Pe_\ell^*\|^2 + \|Pe_\ell^*\|^2$ and $\|e_\ell^* - Pe_\ell^*\|^2 = \|u_{\ell,k} - Pu_{\ell,k}\|^2 = \mathrm{dist}_{\|.\|}(u_{\ell,k}, W)^2$ as well as

$$\|Pe_\ell^*\|^2 = \|u_k^* - Pu_{\ell,k}\|^2 = (1 - \|Pu_{\ell,k}\|)^2.$$

Moreover, $b(Pu_{\ell,k}, u_{\ell,k}) = \|Pu_{\ell,k}\| b(u_k^*, u_{\ell,k})$ and

$$b(u_k^*, u_{\ell,k}) = b(u_k^*, Pu_{\ell,k}) = b(Pu_{\ell,k}, Pu_{\ell,k})/\|Pu_{\ell,k}\| = \|Pu_{\ell,k}\| \geq 0.$$

Therefore, $\operatorname{dist}_{\|.\|}(u_{\ell,k}, W)^2 = 1 - \|Pu_{\ell,k}\|^2$ and it follows

$$\begin{aligned} \|e_\ell^*\|^2 &= 1 - \|Pu_{\ell,k}\|^2 + (1 - \|Pu_{\ell,k}\|)^2 \ = 2(1 - \|Pu_{\ell,k}\|) \\ &= 2\frac{1 - \|Pu_{\ell,k}\|^2}{1 + \|Pu_{\ell,k}\|} = 2\frac{\operatorname{dist}_{\|.\|}(u_{\ell,k}, W)^2}{1 + \|Pu_{\ell,k}\|}. \end{aligned}$$

This proves the first equality. For a proof of the second inequality notice that $(u_{n+1}, \ldots, u_{n+m})$ is some b-orthonormal basis of W and therefore $Pu_{\ell,k} = \sum_{j \in I} b(Pu_{\ell,k}, u_j) u_j$. Suppose that the global mesh-size is small enough in the sense that $\varepsilon := \max\{\|u_j - P_\ell u_j\| : j \in I\} \ll 1$. With Kronecker's delta, $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$, it follows for all $i, j \in I$

$$|b(P_\ell u_i, P_\ell u_j) - \delta_{ij}| = |b(u_i, P_\ell u_j) - \delta_{ij}| = |b(u_i, P_\ell u_j - u_j)| \leq \varepsilon.$$

Thus, $(P_\ell u_{n+1}, \ldots, P_\ell u_{n+m})$ is a basis of W_ℓ and $u_{\ell,k} = \sum_{j \in I} \alpha_j P_\ell u_j$ for some α_j. Let $i \in I$, from

$$b(u_i, u_{\ell,k}) = \sum_{j \in I} \alpha_j b(u_i, P_\ell u_j) = \alpha_i + \sum_{j \in I} \alpha_j (b(u_i, P_\ell u_j) - \delta_{ij})$$

it follows

$$|\alpha_i| \leq |b(u_i, u_{\ell,k})| + \sum_{j \in I} |\alpha_j| |b(u_i, P_\ell u_j) - \delta_{ij}| \leq 1 + \varepsilon \sum_{j \in I} |\alpha_j|.$$

Suppose that $0 < \varepsilon \leq 1/(2m)$, then summation over i yields

$$\sum_{i \in I} |\alpha_i| \leq m + \varepsilon m \sum_{j \in I} |\alpha_j|$$

and hence

$$\sum_{i \in I} |\alpha_i| \leq m/(1 - \varepsilon m) \leq 2m.$$

Thus, $|\alpha_j - b(u_{\ell,k}, u_j)| \leq 2m\varepsilon$ and it holds

$$\begin{aligned}
\operatorname{dist}_{\|.\|}(u_{\ell,k}, W) &= \left\| \sum_{j\in I} (\alpha_j P_\ell u_j - b(Pu_{\ell,k}, u_j)u_j) \right\| \\
&= \left\| \sum_{j\in I} (\alpha_j P_\ell u_j - b(u_{\ell,k}, u_j)u_j) \right\| \\
&= \left\| \sum_{j\in I} \big((\alpha_j - b(u_{\ell,k}, u_j))P_\ell u_j + b(u_{\ell,k}, u_j)(P_\ell u_j - u_j)\big) \right\| \\
&\leq m(2m+1) \max_{j\in I} \|u_j - P_\ell u_j\|.
\end{aligned}$$

The triangle inequality leads to

$$\operatorname{dist}_{\|.\|}(u_j, W_\ell) \leq \|u_j - G_\ell u_j\| + \operatorname{dist}_{\|.\|}(G_\ell u_j, W_\ell).$$

The previous two inequalities plus Proposition 3.1 conclude the proof. □

Theorem 3.1 *For sufficiently small mesh-size*

$$h_\ell := \max\{h_T : T \in \mathcal{T}_\ell\} \quad \textit{with} \quad h_T := diam(T)$$

there exists $0 < \delta_\ell < 1$ *with*

$$\sum_{j\in I} \left\| \frac{Pu_{\ell,j}}{\|Pu_{\ell,j}\|} - u_{\ell,j} \right\| \leq \frac{\sum_{j\in I} \left\| Res_{\ell,j} \right\|_*}{1-\delta_\ell} \quad \textit{and} \quad \lim_{h_\ell\to 0} \delta_\ell = 0.$$

Proof The eigenvalue problem (1.1) corresponds to the boundary value problem to find $z \in V$ such that

$$a(z, v) = \int_\Omega f v \, dx \quad \text{for all } v \in V.$$

Suppose this problem is H^{1+s}-regular for all $f \in L^2(\Omega)$, i.e., $z \in H^{1+s}(\Omega) \cap V$ with $\|z\|_{H^{1+s}(\Omega)} \leq C_{reg}\|f\|_{L^2(\Omega)}$. Then the following convergence estimate holds for the Galerkin projection $G_\ell : V \to V_\ell$

$$\|z - G_\ell z\|_{H^1(\Omega)} \leq C_{conv} h_\ell^s \|z\|_{H^{1+s}(\Omega)}$$

for the maximal interior angle ω and $0 < s < \pi/\omega$ [7, Theorem 14.3.3]. Under the above assumption, that the problem is H^{1+s}-regular, the Aubin–Nitzsche duality technique leads to

$$\|u_j - G_\ell u_j\| \leq C_{reg} C_{conv} h_\ell^s \left\| u_j - G_\ell u_j \right\| .$$

Suppose that h_ℓ is sufficiently small such that $\varepsilon := \max\{\|Pu_{\ell,j} - u_{\ell,j}\| : j \in I\} \leq 1/(2m)$. Then, with the same argumentation as in Proposition 3.2, $(Pu_{\ell,n+1}, \ldots, Pu_{\ell,n+m})$ is a basis of W and $u_k = \sum_{j\in I} \alpha_j Pu_{\ell,j}$ for some α_j with $|\alpha_j| \leq 1 + 2\varepsilon m \leq 2$ and $k \in I$. Since G_ℓ is Galerkin projection with best approximating property, it holds for $v_\ell := \sum_{j\in I} \alpha_j \|Pu_{\ell,j}\| u_{\ell,j} \in W_\ell$

$$|\!|\!| u_k - G_\ell u_k |\!|\!| \leq |\!|\!| u_k - v_\ell |\!|\!| \leq 2 \sum_{j\in I} \left|\!\left|\!\left| \frac{Pu_{\ell,j}}{\|Pu_{\ell,j}\|} - u_{\ell,j} \right|\!\right|\!\right|. \tag{3.1}$$

With the Friedrichs inequality $\|v\| \leq C_F |\!|\!| v |\!|\!|$ for all $v \in V$, Lemma 3.1 yields

$$\sum_{j\in I} \left|\!\left|\!\left| \frac{Pu_{\ell,j}}{\|Pu_{\ell,j}\|} - u_{\ell,j} \right|\!\right|\!\right| \leq \sum_{j\in I} |\!|\!| Res_{\ell,j} |\!|\!|_* + C_F \frac{\lambda + \lambda_{\ell,n+m}}{2} \sum_{j\in I} \left\| \frac{Pu_{\ell,j}}{\|Pu_{\ell,j}\|} - u_{\ell,j} \right\|.$$

Suppose that h_ℓ is sufficiently small such that

$$\delta_\ell := \sqrt{2} h_\ell^s m M_2(I) C_{reg} C_{conv} C_F (\lambda + \lambda_{\ell,n+m}) \ll 1.$$

Then Proposition 3.2 together with (3.1) lead to

$$\sum_{j\in I} \left|\!\left|\!\left| \frac{Pu_{\ell,j}}{\|Pu_{\ell,j}\|} - u_{\ell,j} \right|\!\right|\!\right| \leq \frac{\sum_{j\in I} |\!|\!| Res_{\ell,j} |\!|\!|_*}{1 - \delta_\ell}.$$

□

Notice that $1/(1-\delta_\ell) \to 1$ as the maximal mesh-size $h_\ell \to 0$.

4 Two a posteriori error estimators

The a posteriori error estimates of this section employ the abstract framework of [10] by estimating the dual norm of the residual $|\!|\!| Res_\ell |\!|\!|_*$. The first a posteriori error estimator is explicit residual-based and the second improves the averaging error estimator of [20].

4.1 Residual-based error estimator

The book of Verfürth [24] summarises a few equivalences of a posteriori error estimates. This and the following estimate allow for reliable and efficient error estimators via other estimators as well. Given any interior edge E, written $E \in \mathcal{E}_\ell$, of length h_E and with normal unit vector ν_E let $[\nabla u_\ell] := \nabla u_\ell|_{T_+} - \nabla u_\ell|_{T_-}$ denote the jump of the piecewise constant gradient across $E = \partial T_+ \cap \partial T_-$ from the neighbouring element domains $T_\pm \in \mathcal{T}_\ell$. The notation $x \lesssim y$ abbreviates the inequality $x \leq Cy$ with a constant $C > 0$ which does not depend on the mesh-size.

Theorem 4.1 *Let* (λ, u) *and* (λ_ℓ, u_ℓ) *be eigenpairs of* (1.1) *and* (1.2). *Then it holds*

$$\|\!|Res_\ell\|\!|_*^2 \lesssim \eta_\ell^2 := \sum_{E\in\mathcal{E}_\ell} h_E \|[\nabla u_\ell]\cdot \nu_E\|_{L^2(E)}^2 \lesssim \|\!|e_\ell\|\!|^2 .$$

Proof (reliability) Let v_ℓ be the L^2 projection of v in V_ℓ. The approximation property (iv) of Sect. 2.4 for the edges reads

$$\sum_{E\in\mathcal{E}_\ell} \|h_E^{-1/2}(v-v_\ell)\|_{L^2(E)}^2 \lesssim \|\nabla v\|_{L^2(\Omega)}^2.$$

The definition of the residual and some elementary algebra yields

$$\begin{aligned}
Res_\ell(v) &= Res_\ell(v-v_\ell) = \lambda_\ell b(u_\ell, v-v_\ell) - a(u_\ell, v-v_\ell)\\
&= -a(u_\ell, v-v_\ell) = -\sum_{E\in\mathcal{E}_\ell}\int_E ([\nabla u_\ell]\cdot\nu_E)(v-v_\ell)\,ds\\
&\leq \sum_{E\in\mathcal{E}_\ell} h_E^{1/2}\|[\nabla u_\ell]\cdot\nu_E\|_{L^2(E)}\|h_E^{-1/2}(v-v_\ell)\|_{L^2(E)}\\
&\leq \left(\sum_{E\in\mathcal{E}_\ell} h_E\|[\nabla u_\ell]\cdot\nu_E\|_{L^2(E)}^2\right)^{1/2}\left(\sum_{E\in\mathcal{E}_\ell}\|h_E^{-1/2}(v-v_\ell)\|_{L^2(E)}^2\right)^{1/2}\\
&\lesssim \eta_\ell\|\nabla v\|_{L^2(\Omega)}.
\end{aligned}$$

□

Proof ((global) efficiency) Utilizing the bubble function technique of Verfürth [24, Lemma 1.3], Durán, Padra, and Rodríguez proved local efficiency for the edge-residuals [13, Lemma 3.4], namely

$$h_E^{1/2}\|[\nabla u_\ell]\cdot\nu_E\|_{L^2(E)} \lesssim \|\nabla e_\ell\|_{L^2(\omega_E)} + h_{\omega_E}\|\lambda u - \lambda_\ell u_\ell\|_{L^2(\omega_E)},$$

for the edge patch $\omega_E := T_+ \cup T_-$ of E. With $h_\ell := \max\{h_T : T\in\mathcal{T}_\ell\}$, the global version reads

$$\eta_\ell^2 \lesssim \|\!|e_\ell\|\!|^2 + h_\ell^2\|\lambda u - \lambda_\ell u_\ell\|^2.$$

Some elementary algebra in the spirit of Lemma 3.1 shows

$$\|\lambda u - \lambda_\ell u_\ell\|^2 = (\lambda_\ell - \lambda)^2 + \lambda\lambda_\ell\|e_\ell\|^2.$$

Lemma 3.1 yields $(\lambda_\ell - \lambda)^2 \leq \|\!|e_\ell\|\!|^4$ and $\lambda\lambda_\ell\|e_\ell\|^2 \leq \lambda_\ell\,\|\!|e_\ell\|\!|^2$. Since λ_ℓ is bounded by λ_0 it holds

$$\eta_\ell^2 \lesssim \|\!|e_\ell\|\!|^2$$

even for larger mesh-sizes $h_\ell \lesssim 1$. □

4.2 Averaging technique for a posteriori error control

Let $A_\ell : V_\ell^d \to \mathcal{S}^1(\mathcal{T}_\ell)^d := V_\ell^d \cap \mathcal{C}(\Omega)^d$ be some local averaging operator. For example,

$$A_\ell(\nabla u_\ell) := \sum_{z \in \mathcal{N}_\ell} \frac{1}{|\omega_z|} \left(\int_{\omega_z} \nabla u_\ell \, dx \right) \varphi_z,$$

with nodal hat functions φ_z. Alternative estimators from [9] could be employed as well.

Theorem 4.2 *Let (λ, u) and (λ_ℓ, u_ℓ) be eigenpairs of (1.1) and (1.2). Then it holds*

$$|||Res_\ell|||_*^2 \lesssim \mu_\ell^2 := \sum_{T \in \mathcal{T}} \|A_\ell(\nabla u_\ell) - \nabla u_\ell\|_{L^2(T)}^2 \lesssim |||e_\ell|||^2 .$$

Proof Let v_ℓ be the L^2 projection of v in V_ℓ. Since $A_\ell(\nabla u_\ell)$ is globally continuous, the divergence theorem is globally applicable. Notice that for the finite dimensional subspace V_ℓ there holds the local inverse inequality

$$\|h_T \operatorname{div}(v_\ell)\|_{L^2(T)} \le C_{inv} \|v_\ell\|_{L^2(T)}.$$

Together with the the approximation property (iv) of Sect. 2.4,

$$\sum_{T \in \mathcal{T}_\ell} \|h_T^{-1}(v - v_\ell)\|_{L^2(T)}^2 \lesssim \|\nabla v\|_{L^2(\Omega)}^2,$$

it follows

$$\begin{aligned} -\int_\Omega A_\ell(\nabla u_\ell)\nabla(v - v_\ell)\, dx &= \int_\Omega (v - v_\ell) \operatorname{div}(A_\ell(\nabla u_\ell))\, dx \\ &= \sum_T \int_T h_T \operatorname{div}(A_\ell(\nabla u_\ell)) h_T^{-1}(v - v_\ell)\, dx \\ &\le \sum_T \|h_T \operatorname{div}(A_\ell(\nabla u_\ell) - \nabla u_\ell)\|_{L^2(T)} \|h_T^{-1}(v - v_\ell)\|_{L^2(T)} \\ &\le C_{inv} \sum_T \|A_\ell(\nabla u_\ell) - \nabla u_\ell\|_{L^2(T)} \|h_T^{-1}(v - v_\ell)\|_{L^2(T)} \\ &\lesssim \|A_\ell(\nabla u_\ell) - \nabla u_\ell\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}. \end{aligned}$$

This inequality and the stability (iii) of Sect. 2.4,

$$\|\nabla v_\ell\|_{L^2(\Omega)} \lesssim \|\nabla v\|_{L^2(\Omega)},$$

lead to

$$\begin{aligned} Res_\ell(v) &= \lambda_\ell b(u_\ell, v - v_\ell) - a(u_\ell, v - v_\ell) = -a(u_\ell, v - v_\ell) \\ &= -\int_\Omega A_\ell(\nabla u_\ell)\nabla(v - v_\ell) + \int_\Omega (A_\ell(\nabla u_\ell) - \nabla u_\ell)\nabla(v - v_\ell) \\ &\lesssim \mu_\ell \|\nabla v\|_{L^2(\Omega)}. \end{aligned}$$

Hence we have proved reliability. The efficiency is proved by the known fact that the averaging estimator is equivalent to the edge-residual estimator [24]. Since the edge-residual estimator is efficient, so is μ_ℓ. A direct proof of efficiency for a class of averaging operators follows as in [9]. □

5 AFEM convergence

The main results are discussed in the first subsection and proven in the subsequent ones.

5.1 Global strong convergence and contraction property

Let k be some fixed positive integer and suppose dim $V_0 \geq k$. Let $(V_\ell)_{\ell=0,1,2,\ldots}$ denote the nested sequence of discrete spaces computed by the adaptive algorithm based on the residual

$$Res_\ell := \lambda_\ell b(u_\ell, \cdot) - a(u_\ell, \cdot)$$

for the k-th algebraic eigenvalue λ_ℓ of the discrete eigenvalue problem on the level ℓ with *some* eigenvector $u_\ell \in V_\ell$. Suppose that $V_\ell \subseteq \ker(Res_\ell)$ and $\|u_\ell\| = 1$ and notice that at least the orientation of u_ℓ is arbitrary even if the discrete eigenspan of λ_ℓ is one-dimensional. The procedure MARK employs the edge-contributions $\eta_E^{(\ell)} := h_E^{1/2}\|[\nabla u_\ell] \cdot \nu_E\|_{L^2(E)}$ and computes $\mathcal{M}_\ell \subseteq \mathcal{E}_\ell$ (with minimal cardinality) such that

$$\eta_\ell^2 := \sum_{E\in\mathcal{E}_\ell} \eta_E^{(\ell)^2} \leq \theta^{-1} \sum_{E\in\mathcal{M}_\ell} \eta_E^{(\ell)^2}$$

with some global parameter $0 < \theta < 1$. The global convergence result will be proved throughout the remaining part of this section.

Theorem 5.1 (global convergence) *The sequence of discrete eigenvalues* (λ_ℓ) *converges towards some eigenvalue* λ *of the continuous problem. Each subsequence* (u_{ℓ_j})

of discrete eigenvectors has a further subsequence which converges strongly towards some u in V and u is an eigenvector of λ.

Theorem 5.1 shows that spurious eigenvalues do not occur: Every accumulation point of discrete eigenvalues is an exact eigenvalue. Moreover, for a simple eigenvalue λ (i.e., the eigenspan is one-dimensional) it shows that, up to a proper choice of the sign of $\pm u_\ell$, the complete sequence converges strongly to the eigenvector $\pm u$ of λ.

Notice that there is monotone convergence of the discrete eigenvalues to an exact eigenvalue λ. The Rayleigh-Ritz principle guarantees that λ is amongst the exact eigenvalues number k or higher but it remains open to conclude that λ equals the k-th one. Spurious eigenvalues cannot appear as any limit is some exact eigenvalue, but, without further assumptions we cannot guarantee that some exact eigenvalues are left out. To avoid that, one requires some global assumption such as that the mesh-size is globally fine enough.

In the restricted case of a simple eigenvalue λ the following contraction property holds.

Theorem 5.2 (contraction property) *If the triangulation $\mathcal{T}_0$ is sufficiently fine, i.e., h_0 is sufficiently small, and λ is simple, then there exists $\gamma > 0$ and $0 < \rho < 1$ such that, for all $\ell = 0, 1, 2, \ldots$,*

$$\gamma \eta_{\ell+1}^2 + \|\!|u - u_{\ell+1}|\!\|^2 \le \rho \left(\gamma \eta_\ell^2 + \|\!|u - u_\ell|\!\|^2\right).$$

An alternative name for the contraction property is Q-linear convergence and this holds for the combination of error and estimator. An immediate consequence is R-linear convergence of the errors in the sense that, for all $\ell = 0, 1, 2, \ldots$, it holds

$$\|\!|u - u_\ell|\!\|^2 \lesssim \rho^\ell.$$

The proofs of the two theorems will be the content of the subsequent subsections.

5.2 Strong convergence of subsequences

The Raleigh–Ritz principle shows for the nested discrete spaces $V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots$ that (λ_ℓ) is monotone decreasing and hence convergent to some reel number $\lambda_\infty > 0$ which is even bigger than or equal to the k-th exact eigenvalue. In particular, (λ_ℓ) is a Cauchy sequence. Notice that $\lambda_\ell = \|\!|u_\ell|\!\|^2$ and hence (u_ℓ) is bounded in the Hilbert space V. Since each bounded sequence in V has some subsequence which is weakly convergent in V and strongly convergent in H towards some element in V, there exist some subsequence (u_{ℓ_j}) and some weak limit $u_\infty \in V$ such that

$$\lim_{j\to\infty} \|u_\infty - u_{\ell_j}\| = 0 \quad \text{while} \quad (u_{\ell_j}) \rightharpoonup u_\infty \text{ in } V.$$

The arguments from the first part of Lemma 3.1 show for $\ell \le m$ that

$$\|\!|u_m - u_\ell|\!\|^2 = \lambda_\ell - \lambda_m + \lambda_m \|u_m - u_\ell\|^2$$

and, for subsequences, the right-hand side tends to zero as $\ell \to \infty$ and hence (u_{ℓ_j}) is a Cauchy sequence in V. Consequently,

$$\lim_{j\to\infty} |||u_\infty - u_{\ell_j}||| = 0.$$

In particular, $\|u_\infty\| = 1$ and the residual Res_∞ reads

$$Res_\infty := \lambda_\infty b(u_\infty, \cdot) - a(u_\infty, \cdot) \in V^*.$$

It remains to prove $Res_\infty = 0$. The aforementioned convergence properties show the weak convergence

$$(Res_{\ell_j}) \rightharpoonup Res_\infty \quad \text{in } V^*.$$

So it remains to conclude

$$\lim_{j\to\infty} |||Res_{\ell_j}|||_* = 0$$

which will follow from the reliability of Theorem 4.1 and the convergence of the estimators in Lemma 5.2 below. The proof of that follows from an estimator perturbation result similar to [11].

Lemma 5.1 *There exist some $C > 0$ and $0 < \rho < 1$ such that, for all non-negative integers ℓ and m, it holds*

$$\eta_{\ell+m}^2 \le \rho\eta_\ell^2 + C\,|||u_{\ell+m} - u_\ell|||^2 .$$

Proof For all $E \in \mathcal{E}_\ell$ we have either $E \in \mathcal{E}_{\ell+m}$ or otherwise there exist $E_1, \ldots, E_J \in \mathcal{E}_{\ell+m}$ with $E = E_1 \cup \cdots \cup E_J$ and $J \ge 2$. In the second case $E \in \mathcal{E}_\ell \setminus \mathcal{E}_{\ell+m}$, for any $0 < \delta < \theta/(2-\theta)$,

$$\begin{aligned}
\sum_{j=1}^{J} \eta_{E_j}^{(\ell+m)2} &= \sum_{j=1}^{J} h_{E_j}^2 |[\nabla u_{\ell+m}]\cdot \nu_{E_j}|^2 \\
&\le \sum_{j=1}^{J} h_{E_j}^2 \Big((1+\delta)|[\nabla u_\ell]\cdot \nu_{E_j}|^2 + (1+1/\delta)|[\nabla u_{\ell+m} - \nabla u_\ell]\cdot \nu_{E_j}|^2\Big) \\
&\le (1+\delta)/2\,\eta_E^{(\ell)2} + (1+1/\delta)\sum_{j=1}^{J} h_{E_j}^2 |[\nabla u_{\ell+m} - \nabla u_\ell]\cdot \nu_{E_j}|^2.
\end{aligned}$$

Notice that the factor $1/2$ results from $J > 1$ refinements (at least one bisection) of $E \in \mathcal{E}_\ell \backslash \mathcal{E}_{\ell+m}$. Therefore,

$$\sum_{\substack{E\in\mathcal{E}_{\ell+m}\\ E\subseteq\cup\mathcal{E}_\ell}} \eta_E^{(\ell+m)2} \leq (1+\delta)/2 \sum_{E\in\mathcal{E}_\ell\backslash\mathcal{E}_{\ell+m}} \eta_E^{(\ell)2} + (1+\delta) \sum_{E\in\mathcal{E}_\ell\cap\mathcal{E}_{\ell+m}} \eta_E^{(\ell)2}$$
$$+(1+1/\delta) \sum_{\substack{E\in\mathcal{E}_{\ell+m}\\ E\subseteq\cup\mathcal{E}_\ell}} h_E^2 |[\nabla u_{\ell+m} - \nabla u_\ell]\cdot\nu_E|^2.$$

For any $E \in \mathcal{E}_{\ell+m}$ with $E \nsubseteq \cup\mathcal{E}_\ell$, $[\nabla u_\ell]\cdot\nu_E = 0$ on E. Hence

$$\eta_E^{(\ell+m)2} = h_E^2 |[\nabla u_{\ell+m} - \nabla u_\ell]\cdot\nu_E|^2.$$

Therefore,

$$\eta_{\ell+m}^2 \leq (1+\delta)/2 \sum_{E\in\mathcal{E}_\ell\backslash\mathcal{E}_{\ell+m}} \eta_E^{(\ell)2} + (1+\delta) \sum_{E\in\mathcal{E}_\ell\cap\mathcal{E}_{\ell+m}} \eta_E^{(\ell)2}$$
$$+(1+1/\delta) \sum_{E\in\mathcal{E}_{\ell+m}} h_E^2 |[\nabla u_{\ell+m} - \nabla u_\ell]\cdot\nu_E|^2.$$

Since $\nabla u_{\ell+m} - \nabla u_\ell$ is piecewise constant with respect to the shape regular triangulation $\mathcal{T}_{\ell+m}$,

$$h_E^2 |[\nabla u_{\ell+m} - \nabla u_\ell]\cdot\nu_E|^2 \lesssim \|\nabla u_{\ell+m} - \nabla u_\ell\|_{L^2(\omega_E)}$$

for the edge patch ω_E of E in $\mathcal{T}_{\ell+m}$. Since there is only a finite overlap of all edge patches,

$$\eta_{\ell+m}^2 \leq (1+\delta)/2 \sum_{E\in\mathcal{M}_\ell} \eta_E^{(\ell)2} + (1+\delta) \sum_{E\in\mathcal{E}_\ell\backslash\mathcal{M}_\ell} \eta_E^{(\ell)2} + C\,|\!|\!| u_{\ell+m} - u_\ell |\!|\!|^2 .$$

The bulk criterion leads to

$$1/2 \sum_{E\in\mathcal{M}_\ell} \eta_E^{(\ell)2} + \sum_{E\in\mathcal{E}_\ell\backslash\mathcal{M}_\ell} \eta_E^{(\ell)2} = \eta_\ell^2 - 1/2 \sum_{E\in\mathcal{M}_\ell} \eta_E^{(\ell)2} \leq (1-\theta/2)\eta_\ell^2.$$

Since $\delta < \theta/(2-\theta)$, the resulting estimate proves the assertion:

$$\eta_{\ell+m}^2 \leq (1+\delta)(1-\theta/2)\eta_\ell^2 + C\,|\!|\!| u_{\ell+m} - u_\ell |\!|\!|^2 .$$

□

Lemma 5.2 *For the subsequence (u_{ℓ_j}) it holds*

$$\lim_{\ell_j\to\infty} \eta_{\ell_j}^2 = 0.$$

Proof Since (u_{ℓ_j}) is a Cauchy sequence and Lemma 5.1 yields

$$\eta_{\ell_{j+1}}^2 \le \rho \eta_{\ell_j}^2 + C \left\| u_{\ell_{j+1}} - u_{\ell_j} \right\|^2 \quad \text{for all } j = 1, 2, \ldots$$

one concludes the assertion with some elementary analysis and the geometric series. □

This concludes the proof of Theorem 5.1 on the global convergence.

5.3 Contraction property

Throughout this subsection, let (λ, u) denote some eigenpair of the continuous eigenvalue problem, (λ_ℓ, u_ℓ) denotes some discrete eigenpair with error estimator η_ℓ, and $e_\ell := u - u_\ell$. Suppose that λ is a simple eigenvalue and that the global mesh-size is sufficiently small such that λ_ℓ is well separated from the remaining part of the spectrum.

Theorem 5.3 *There exist constants* $0 < \varrho < 1$ *and* $\gamma > 0$ *such that, for all* $\ell = 0, 1, 2, \ldots,$

$$\gamma \eta_{\ell+1}^2 + \|e_{\ell+1}\|^2 \le \varrho \left(\gamma \eta_\ell^2 + \|e_\ell\|^2 \right) + 3\lambda_{\ell+1} \|e_{\ell+1}\|^2 + 3\lambda_\ell \|e_\ell\|^2.$$

Proof Let ρ denote the constant in Lemma 5.1 which, for $m = 1$, becomes

$$\eta_{\ell+1}^2 \le \rho \eta_\ell^2 + C \|u_{\ell+1} - u_\ell\|^2 .$$

This and some algebra (since (λ, u) and (λ_ℓ, u_ℓ) are eigenpairs) lead to

$$\|u_{\ell+1} - u_\ell\|^2 = \|e_\ell\|^2 - \|e_{\ell+1}\|^2 - 2b(\lambda u - \lambda_{\ell+1} u_{\ell+1}, u_{\ell+1} - u_\ell).$$

Thus,

$$\gamma \eta_{\ell+1}^2 + \|e_{\ell+1}\|^2 \le \rho \gamma \eta_\ell^2 + \|e_\ell\|^2 - 2b(\lambda u - \lambda_{\ell+1} u_{\ell+1}, u_{\ell+1} - u_\ell).$$

Set

$$\rho < \varrho := \frac{\rho\gamma + C_{rel}}{\gamma + C_{rel}} < 1.$$

Then

$$\gamma \eta_{\ell+1}^2 + \|e_{\ell+1}\|^2 \le \varrho \left(\gamma \eta_\ell^2 + \|e_\ell\|^2 \right) + (\rho - \varrho)\gamma \eta_\ell^2 + (1 - \varrho) \|e_\ell\|^2 \\ - 2b(\lambda u - \lambda_{\ell+1} u_{\ell+1}, u_{\ell+1} - u_\ell). \tag{5.1}$$

Lemma 3.1 plus Young's inequality yield

$$2\,\|\!|e_\ell|\!\|^2 \leq (\lambda+\lambda_\ell)\|e_\ell\|^2 + 2\,\|\!|Res_\ell|\!\|_*\,\|\!|e_\ell|\!\| \leq (\lambda+\lambda_\ell)\|e_\ell\|^2 + \|\!|Res_\ell|\!\|_*^2 + \|\!|e_\ell|\!\|^2\,.$$

This and the reliability estimate of Theorem 4.1

$$\|\!|Res_\ell|\!\|_*^2 \leq C_{rel}\eta_\ell^2$$

result in

$$\|\!|e_\ell|\!\|^2 \leq (\lambda+\lambda_\ell)\|e_\ell\|^2 + C_{rel}\eta_\ell^2.$$

The last term in (5.1) reads

$$\begin{aligned}
&-2b(\lambda u-\lambda_{\ell+1}u_{\ell+1}, u_{\ell+1}-u_\ell)\\
&\quad = -2\lambda b(u, u_{\ell+1}-u_\ell) + 2\lambda_{\ell+1}b(u_{\ell+1}, u_{\ell+1}-u_\ell)\\
&\quad = \lambda\|e_{\ell+1}\|^2 - \lambda\|e_\ell\|^2 + \lambda_{\ell+1}\|u_{\ell+1}-u_\ell\|^2.
\end{aligned}$$

Young's inequality for $\|u_{\ell+1}-u_\ell\|^2$ yields

$$-2b(\lambda u-\lambda_{\ell+1}u_{\ell+1}, u_{\ell+1}-u_\ell) \leq (2\lambda_{\ell+1}+\lambda)\|e_{\ell+1}\|^2 + (2\lambda_{\ell+1}-\lambda)\|e_\ell\|^2.$$

Since $\lambda \leq \lambda_{\ell+1} \leq \lambda_\ell$, this and (5.1) lead to

$$\begin{aligned}
\gamma\eta_{\ell+1}^2 + \|\!|e_{\ell+1}|\!\|^2 \leq \varrho\left(\gamma\eta_\ell^2 + \|\!|e_\ell|\!\|^2\right) + ((\rho-\varrho)\gamma + C_{rel}(1-\varrho))\,\eta_\ell^2\\
+3\lambda_{\ell+1}\|e_{\ell+1}\|^2 + 3\lambda_\ell\|e_\ell\|^2.
\end{aligned}$$

By definition of ϱ, $(\rho-\varrho)\gamma + C_{rel}(1-\varrho) \leq 0$. This completes the proof of Theorem 5.3. □

Proof of Theorem 5.2 For sufficiently small mesh-sizes $h_\ell \leq h_0 \ll 1$, Proposition 3.2 and Theorem 3.1 show

$$\|e_\ell\|^2 \leq 2M_2^2(\lambda)C_{reg}^2C_{conv}^2h_0^{2s}\,\|\!|e_\ell|\!\|^2 \quad\text{and}\quad \|e_{\ell+1}\|^2 \leq 2M_2^2(\lambda)C_{reg}^2C_{conv}^2h_0^{2s}\,\|\!|e_{\ell+1}|\!\|^2.$$

Hence Theorem 5.3 yields the contraction property with a constant

$$0 < \frac{\varrho + 6\lambda_0M_2^2(\lambda)C_{reg}^2C_{conv}^2h_0^{2s}}{1-6\lambda_0M_2^2(\lambda)C_{reg}^2C_{conv}^2h_0^{2s}} < 1.$$

This concludes the proof of Theorem 5.2. □

```
function [x,lambda] = EWP(coordinates,elements,dirichlet,k)

A = sparse(size(coordinates,1),size(coordinates,1));
B = sparse(size(coordinates,1),size(coordinates,1));
x = zeros(size(coordinates,1),1);

for j = 1:size(elements,1)
    A(elements(j,:),elements(j,:)) = A(elements(j,:),elements(j,:))+...
        stima(coordinates(elements(j,:),:));
    B(elements(j,:),elements(j,:)) = B(elements(j,:),elements(j,:))+...
        det([1,1,1;coordinates(elements(j,:),:)'])*(ones(3)+eye(3))/24;
end

freeNodes = setdiff(1:size(coordinates,1),unique(dirichlet));
[V,D] = eigs(A(freeNodes,freeNodes),B(freeNodes,freeNodes),k,'sm');
x(freeNodes) = V(:,1);lambda = D(1,1);

function stima=stima(vertices)
P = [ones(1,size(vertices,2)+1);vertices'];
Q = P\[zeros(1,size(vertices,2));eye(size(vertices,2))];
stima = det(P)*Q*Q'/prod(1:size(vertices,2));
```

Fig. 2 17 lines of MATLAB

6 Numerical experiments

6.1 Numerical realisation

This section is devoted to four numerical experiments on the square, the L-shape, and the slit domain for the Laplace operator as well as tuning fork vibrations. The edge-based residual estimator and the averaging estimator read

$$\eta_\ell = \left(\sum_{E\in\mathcal{E}} h_E \| [\nabla u_\ell] \cdot \nu_E \|^2_{L^2(E)} \right)^{1/2} \quad \text{and} \tag{6.1}$$

$$\mu_\ell = \left(\sum_{T\in\mathcal{T}} \| A_\ell(\nabla u_\ell) - \nabla u_\ell \|^2_{L^2(T)} \right)^{1/2}. \tag{6.2}$$

The numerical examples show that the a posteriori error estimators are reliable and efficient and that the remaining term is indeed of higher order when compared to the estimators.

The MATLAB implementation follows the spirit of [2,3] and Fig. 2 displays the kernel MATLAB function `EWP.m` of the computer program utilised in this section.

6.2 Unit square

The first example consists of the eigenvalue problem of the Poisson problem on the unit square with Dirichlet boundary condition, that means: seek for the first eigenpair

$$(\lambda, u) = (2\pi^2, 2\sin(x\pi)\sin(y\pi))$$

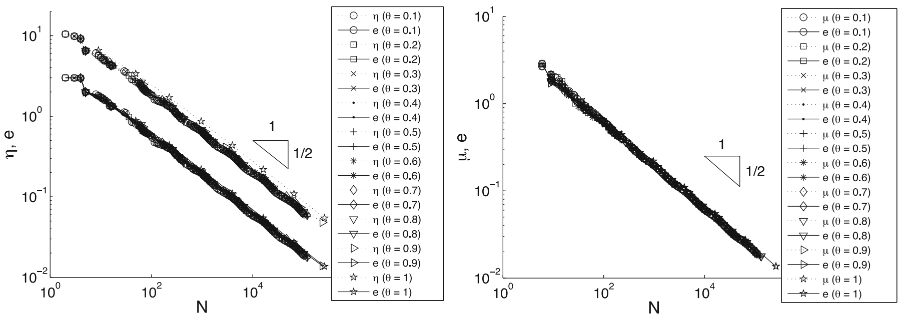

Fig. 3 Convergence history for η_ℓ (*left*) and μ_ℓ (*right*) with different choices of θ for the unite square

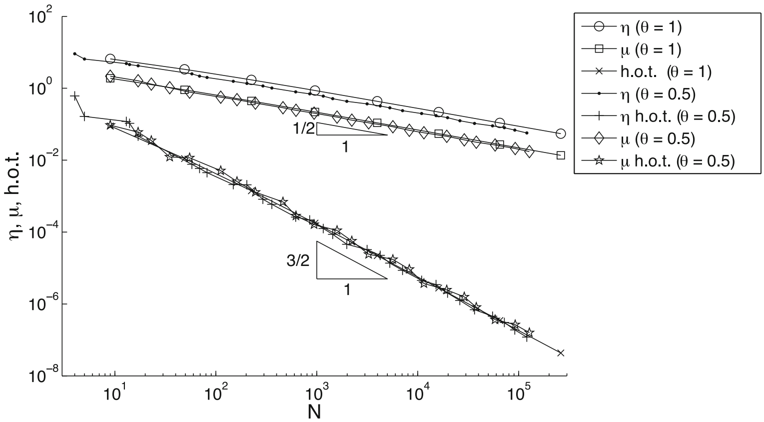

Fig. 4 Comparison of the estimator and the h.o.t. for η_ℓ and μ_ℓ for the unite square

of the Laplace operator in $\Omega = [0, 1] \times [0, 1]$ with

$$-\Delta u = \lambda u \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{along } \partial\Omega.$$

Figure 3 shows the convergence history for $|||e_\ell|||$, η_ℓ (6.1) and μ_ℓ (6.2) for different choices of θ. Notice that $\theta = 1$ results in uniform refinement while $\theta < 1$ leads to adaptively refined meshes. One observes that μ_ℓ is asymptotically exact. In Fig. 4 it is numerically shown that

$$\text{h.o.t.} = \lambda_\ell \|e_\ell\|^2 / \, |||e_\ell|||$$

is really of higher order compared to the estimator η_ℓ or μ_ℓ. Figure 5 shows that the constant C with $\|u - u_\ell\| \leq C \, |||u - G_\ell u|||$ which is bounded in Proposition 3.2 is numerically less than 1 and the L^2-error is of higher order compared to the energy error of the Galerkin projection as shown in Theorem 3.1.

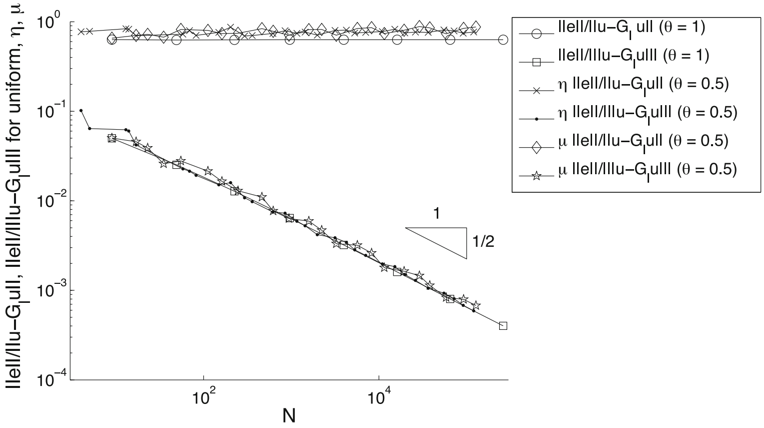

Fig. 5 Size of the constant C with $\|u - u_\ell\| \leq C \, |||u - G_\ell u|||$ and higher order convergence of the L^2-norm compared to the energy norm for the unite square

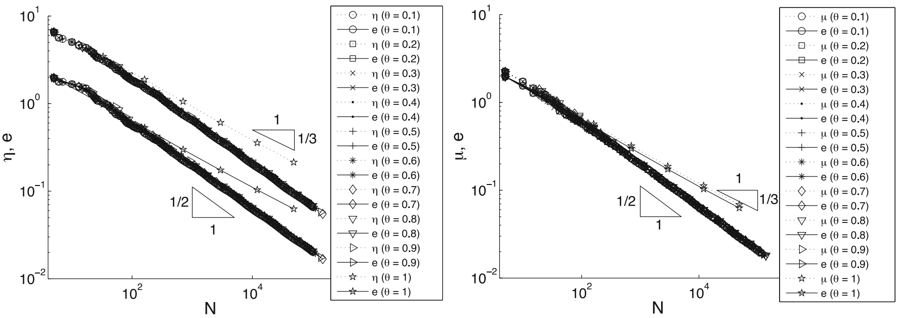

Fig. 6 Convergence history for η_ℓ (*left*) and μ_ℓ (*right*) with different choices of θ for the L-shaped domain

6.3 L-shaped domain

Seek for the first eigenpair (λ, u) of the Laplace operator in $\Omega = [-1, 1] \times [0, 1] \cup [-1, 0] \times [-1, 0]$.

$$-\Delta u = \lambda u \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{along } \partial\Omega.$$

Because the first eigenfunction of the L-shaped domain is singular, the energy error $|||e_\ell|||$ is estimated by

$$|||e_\ell|||^2 = \lambda^* + \lambda_\ell - \lambda b(u^*, u_\ell),$$

for some known approximation $\lambda^* = 9.639723844$ to λ with high accuracy and an approximation u^* to u with second order P_2 FEM. Figure 6 shows the convergence history of η_ℓ and μ_ℓ. Notice that adaptive refinement (for $\theta < 1$) is much better than uniform refinement (for $\theta = 1$). Adaptive refinement results in optimal convergence

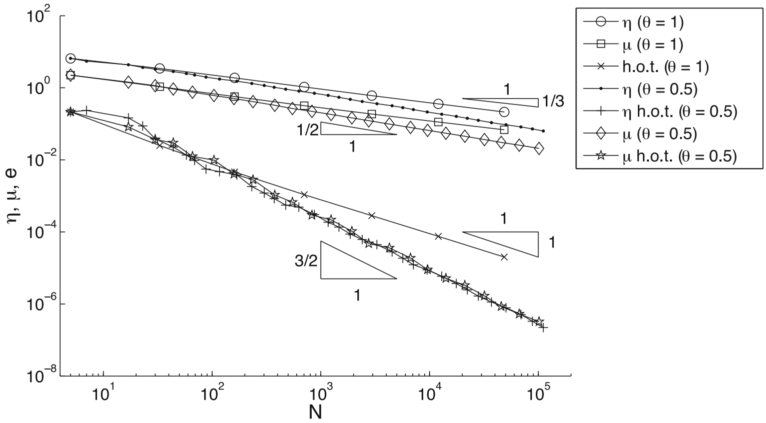

Fig. 7 Comparison of the estimator and the h.o.t. for η_ℓ and μ_ℓ for the L-shaped domain

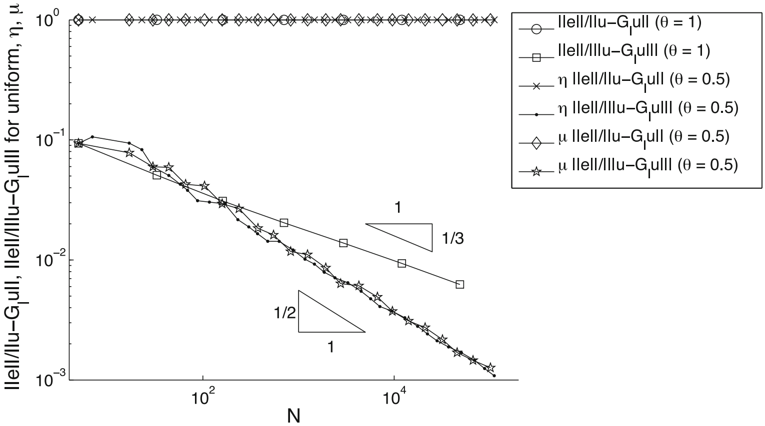

Fig. 8 Size of the constant C with $\|u - u_\ell\| \le C\,|||u - G_\ell u|||$ and higher order convergence of the L^2-norm compared to the energy norm for the L-shaped domain

$O(N_\ell^{-1/2})$ where uniform refinement results in only $O(N_\ell^{-1/3})$ convergence, with $N_\ell = \dim(V_\ell)$ and $N_\ell^{-1/2} \approx h_\ell$ for uniform refined meshes. Notice that μ_ℓ is not asymptotically exact for uniform refinement because of the singularity at the re-entrant corner, but only for the elements at the corner and therefore there is only a small difference. Again in Fig. 7 it is shown that the h.o.t. is of higher order. Figure 8 shows that the constant C in $\|u - u_\ell\| \le C\,|||u - G_\ell u|||$ is about 1 and that the L^2-error is again of higher order, although the solution is singular. Towards the corner singularity at the origin adaptive refined meshes are shown in Fig. 9.

6.4 Slit domain

Although the slit domain $\Omega := (-1, 1)^2 \setminus [0, 1] \times \{0\}$ is not a Lipschitz (the domain is not on one side of the slit) the benchmark serves as an extreme example, where one

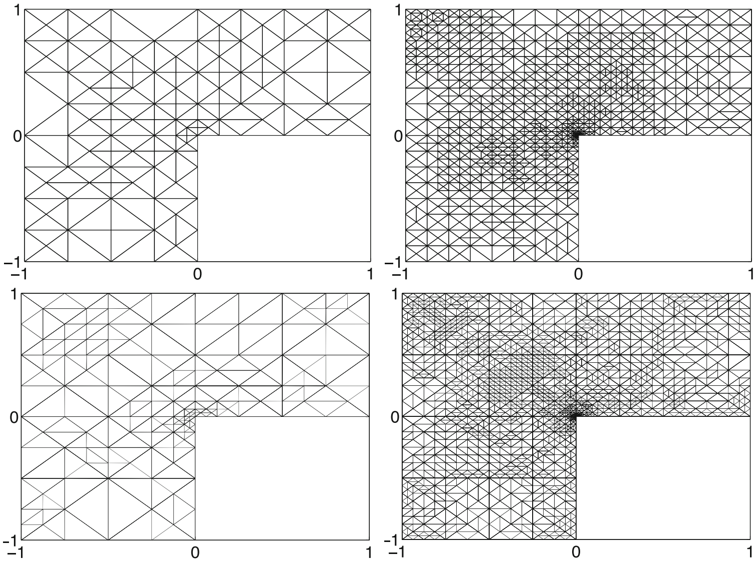

Fig. 9 Adaptive meshes generated with $\theta = 0.5$ for the a posteriori error estimator η_ℓ (*top*) and μ_ℓ (*bottom*) for about 100 and 1000 nodes for the L-shaped domain

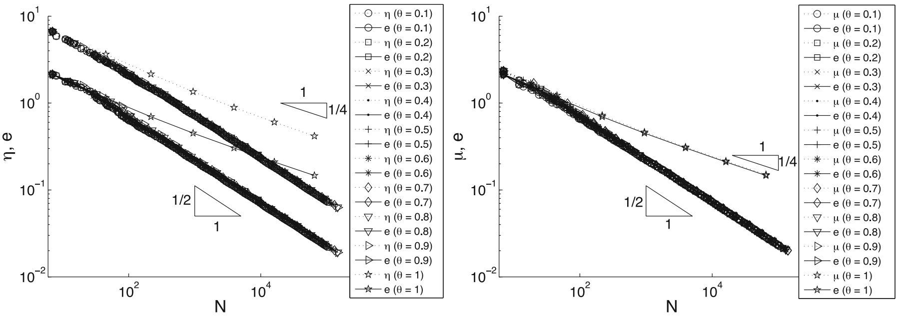

Fig. 10 Convergence history for η_ℓ (*left*) and μ_ℓ (*right*) with different choices of θ for the slit domain

seeks the first eigenpair (λ, u) of the Laplace. Similar to the L-shaped domain, the first eigenfunction is singular and the energy error $|||e_\ell|||$ is estimated by

$$|||e_\ell|||^2 = \lambda^* + \lambda_\ell - \lambda b(u^*, u_\ell),$$

with $\lambda^* = 8.371329711$ of sufficient accuracy and u^* is an approximation to u with second order P_2 FEM. As in the previous example Fig. 10 shows the convergence history of η_ℓ and μ_ℓ. Adaptive refinement results in optimal convergence $O(N_\ell^{-1/2})$ while uniform refinement results in only $O(N_\ell^{-1/4})$ convergence. Figure 11 shows that the h.o.t. $= \lambda_\ell \|e_\ell\|^2 / |||e_\ell|||$ is of higher order. Figure 12 shows that the constant C in $\|u - u_\ell\| \leq C\, |||u - G_\ell u|||$ is about 1 and that even in this extreme example with

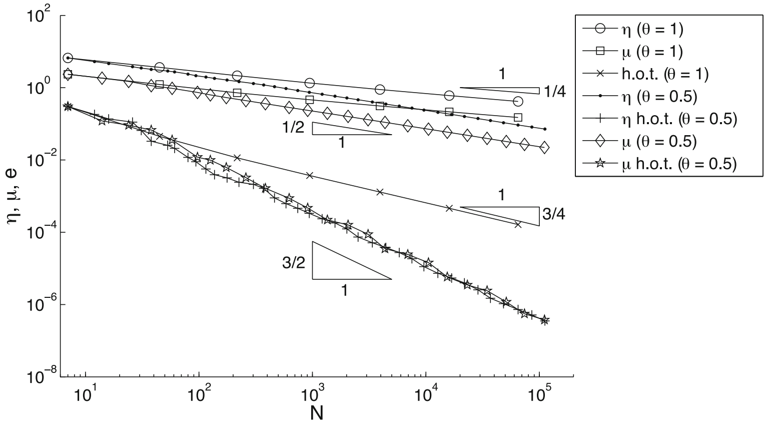

Fig. 11 Comparison of the estimator and the h.o.t. for η_ℓ and μ_ℓ for the slit domain

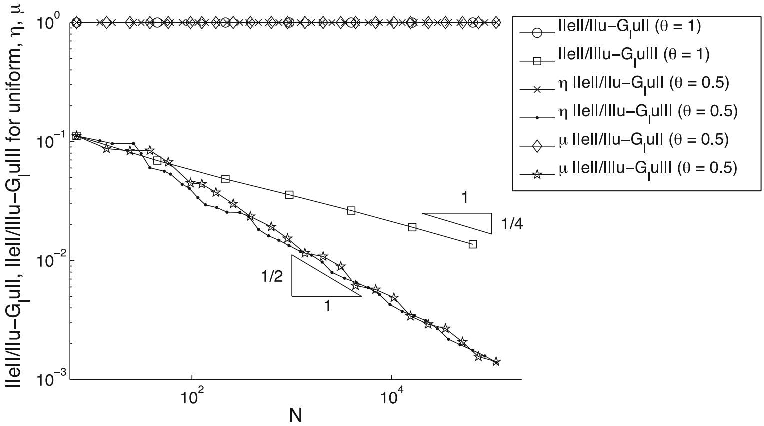

Fig. 12 Size of the constant C with $\|u - u_\ell\| \le C \, |\!|\!| u - G_\ell u |\!|\!|$ and higher order convergence of the L^2-norm compared to the energy norm for the slit domain

poor regularity the L^2-error is of higher order. Different adaptive meshes are shown in Fig. 13.

6.5 Elastic vibrations of a tuning fork

The harmonic dynamic of linear elasticity (involves the Lamé operator $\Delta^* := \operatorname{div}\mathbb{C}\varepsilon$ for the linear Green strain $\varepsilon := \operatorname{sym}\nabla$ of the displacement $u \in V := H_0^1(\Omega; \mathbb{R}^2)$ and the density ρ) leads to the eigenvalue problem of the Lamé operator

$$-\Delta^* u = \lambda \rho u \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Gamma_D.$$

The domain Ω is displayed with the initial triangulation $\mathcal{T}_0$ in Fig. 14 where $\Gamma_D = \partial\Omega \cap ([-1, 1] \times \{0\})$ and the traction vanishes along $\partial\Omega \backslash \Gamma_D$. The weak formulation

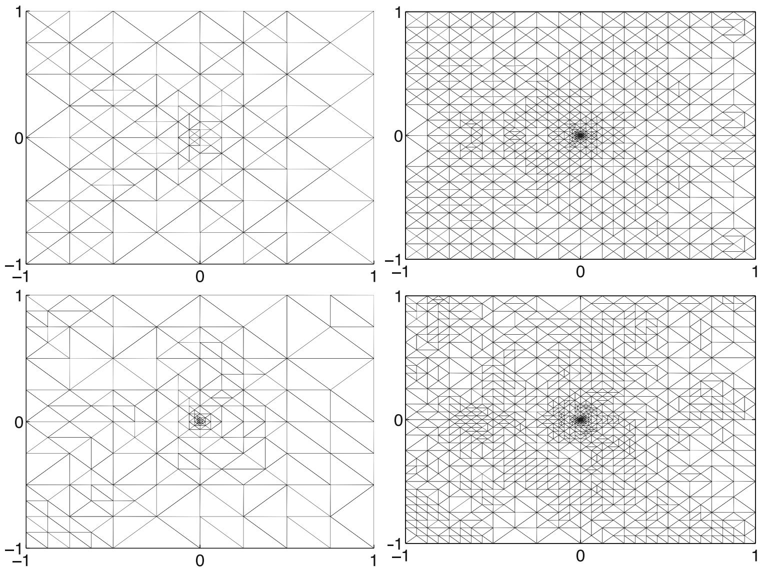

Fig. 13 Adaptive meshes generated with $\theta = 0.5$ for the a posteriori error estimator η_ℓ (*top*) and μ_ℓ (*bottom*) for about 100 and 1000 nodes for the slit domain

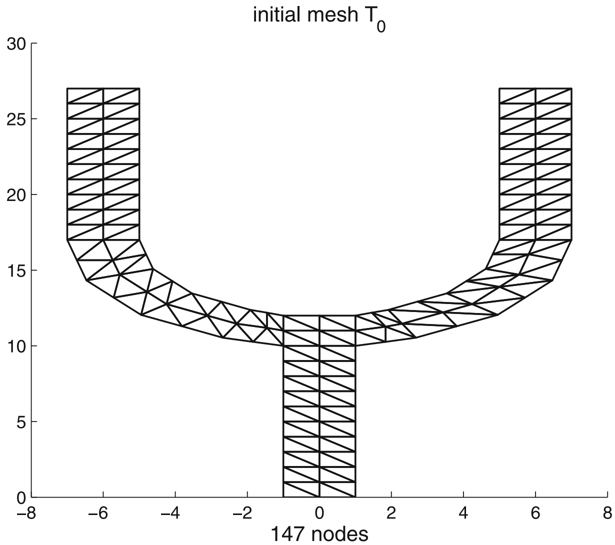

Fig. 14 Initial triangulation $\mathcal{T}_0$ for the tuning fork

involves the bilinear forms

$$a(u, v) = \int_\Omega \varepsilon(u) : \mathbb{C}\varepsilon(v)\, dx \quad \text{and} \quad b(u, v) = \int_\Omega \rho u \cdot v\, dx \quad \text{for } u, v \in V.$$

We refer to [3] for details on the model and the elasticity tensor $\mathbb{C}$ with Poisson's ratio 0.3, Young's modulus $E = 214\text{GPa}$, density $\rho = 1$, as well as to the MAT-

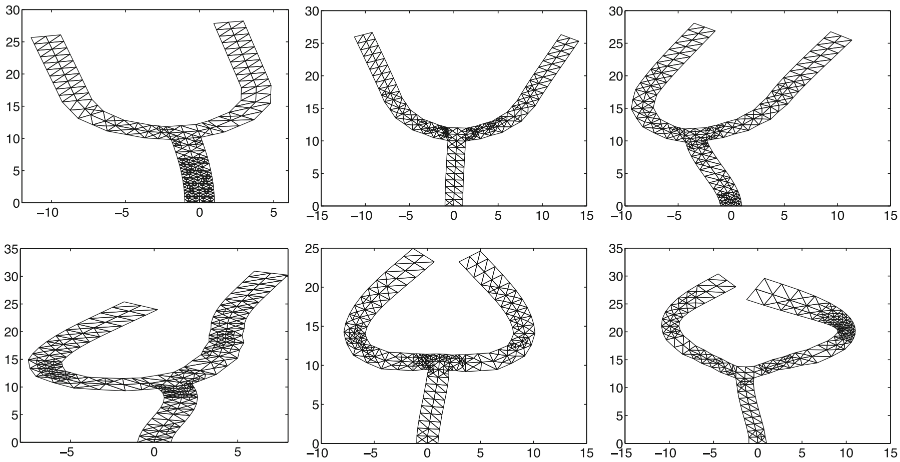

Fig. 15 The first six eigenforms of the tuning fork (from *left* to *right*, *top* to *bottom*) computed on adaptively refined meshes for the corresponding discrete eigenvalue on level $\ell = 7$ with about 500 nodes, stretched by a factor 20

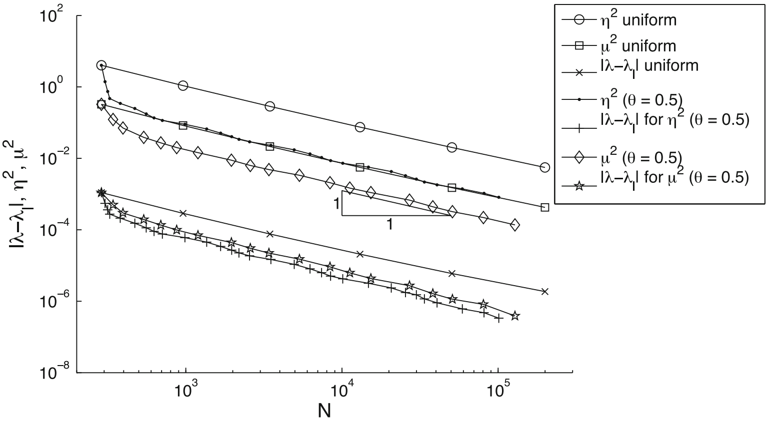

Fig. 16 Convergence history for the first eigenvalue of the tuning fork

LAB simulation tools for the numerical experiments. The first six eigenforms for the discrete eigenvalues on level $\ell = 7$

$$\lambda_{\ell,1}, \ldots, \lambda_{\ell,6} \approx 0.0013049,\ 0.014685,\ 0.068861,\ 0.1748,\ 0.28598,\ 1.2361$$

of the tuning fork are shown in Fig. 15. The convergence history for the error in the first eigenvalue $\lambda_1 \approx 0.00119135$ is displayed in Fig. 16. The expected eigenforms give rise to completely different adapted meshes and seem to correspond reasonably to the eigenmodes.

Acknowledgments The work of the two authors was supported by the German Research Foundation (DFG) under C22 in the Research Center MATHEON. The results on global convergence were accomplished during the first author's happy research stay at the Hausdorff Institute for Mathematics in Bonn, Germany, in Spring 2008.

References

1. Ainsworth, M., Oden, J.T.: A posteriori error estimation in finite element analysis. In: Pure and Applied Mathematics. Wiley-Interscience, New York (2000)
2. Alberty, J., Carstensen, C., Funken, S.A.: Remarks around 50 lines of Matlab: short finite element implementation. Numer. Algorithms **20**(2–3), 117–137 (1999)
3. Alberty, J., Carstensen, C., Funken, S.A., Klose, R.: Matlab implementation of the finite element method in elasticity. Computing **69**(3), 239–263 (2002)
4. Babuška, I., Osborn, J.: Eigenvalue problems. In: Handbook of Numerical Analysis, vol II, pp. 641–787. North-Holland, Amsterdam (1991)
5. Babuška, I., Osborn, J.E.: Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems. Math. Comput. **52**(186), 275–297 (1989)
6. Bangerth, W., Rannacher, R.: Adaptive finite element methods for differential equations. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2003)
7. Brenner, S.C., Scott, L.R.: The mathematical theory of finite element methods. In: Texts in Applied Mathematics, vol. 15, 3rd edn. Springer, New York (2008)
8. Carstensen, C.: An adaptive mesh-refining algorithm allowing for an H^1 stable L^2 projection onto Courant finite element spaces. Constr. Approx. **20**(4), 549–564 (2004)
9. Carstensen, C.: All first-order averaging techniques for a posteriori finite element error control on unstructured grids are efficient and reliable. Math. Comput. **73**(247), 1153–1165 (2004)
10. Carstensen, C.: A unifying theory of a posteriori finite element error control. Numer. Math. **100**(4), 617–637 (2005)
11. Cascon, J.M., Kreuzer, C., Nochetto, R.H., Siebert, K.G.: Quasi-optimal convergence rate for an adaptive finite element method. SIAM J. Numer. Anal. **46**(5), 2524–2550 (2008)
12. Chatelin, F.: Spectral approximation of linear operators. Computer Science and Applied Mathematics. Academic Press, New York (1983)
13. Durán, R.G., Padra, C., Rodríguez, R.: A posteriori error estimates for the finite element approximation of eigenvalue problems. Math. Models Methods Appl. Sci. **13**(8), 1219–1229 (2003)
14. Giani, S., Graham, I.G.: A convergent adaptive method for elliptic eigenvalue problems. SIAM J. Numer. Anal. **47**(2), 1067–1091 (2009)
15. Heuveline, V., Rannacher, R.: A posteriori error control for finite approximations of elliptic eigenvalue problems. Adv. Comput. Math. **15**(1–4), 107–138 (2001)
16. Knyazev, A.V.: New estimates for Ritz vectors. Math. Comput. **66**(219), 985–995 (1997)
17. Knyazev, A.V., Osborn, J.E.: New a priori FEM error estimates for eigenvalues. SIAM J. Numer. Anal. **43**(6), 2647–2667 (2006)
18. Larson, M.G.: A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems. SIAM J. Numer. Anal. **38**(2), 608–625 (2000)
19. Larsson, S., Thomée, V.: Partial differential equations with numerical methods. In: Texts in Applied Mathematics, vol. 45. Springer, Berlin (2003)
20. Mao, D., Shen, L., Zhou, A.: Adaptive finite element algorithms for eigenvalue problems based on local averaging type a posteriori error estimates. Adv. Comput. Math. **25**(1–3), 135–160 (2006)
21. Raviart, P.A., Thomas, J.M.: Introduction à l'analyse numérique des équations aux dérivées partielles. Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris (1983)
22. Sauter, S.: hp-finite elements for elliptic eigenvalue problems: error estimates which are explicit with respect to λ, h, and p. SIAM J. Numer. Anal. **48**(1), 95–108 (2010)
23. Strang, G., Fix, G.J.: An Analysis of the Finite Element Method. Prentice-Hall, Englewood Cliffs (1973)
24. Verfürth, R.: A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Wiley and Teubner, New York (1996)
25. Walsh, T.F., Reese, G.M., Hetmaniuk, U.L.: Explicit a posteriori error estimates for eigenvalue analysis of heterogeneous elastic structures. Comput. Methods Appl. Mech. Eng. **196**(37–40), 3614–3623 (2007)