Numerische
Mathematik

An adaptive finite element method with asymptotic saturation for eigenvalue problems

**C. Carstensen · J. Gedicke · V. Mehrmann ·
A. Miȩdlar**

Abstract This paper discusses adaptive finite element methods for the solution of elliptic eigenvalue problems associated with partial differential operators. An adaptive method based on nodal-patch refinement leads to an asymptotic error reduction property for the computed sequence of simple eigenvalues and eigenfunctions. This justifies the use of the proven saturation property for a class of reliable and efficient hierarchical a posteriori error estimators. Numerical experiments confirm that the saturation property is present even for very coarse meshes for many examples; in other cases the smallness assumption on the initial mesh may be severe.

Mathematics Subject Classification (1991) 65N15 · 65N25 · 65N30

C. Carstensen (✉)
Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: cc@math.hu-berlin.de

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

J. Gedicke
Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA
e-mail: jgedicke@math.lsu.edu

V. Mehrmann · A. Miȩdlar
Institut für Mathematik, MA 4-5 TU Berlin, Strasse des 17. Juni 136, 10623 Berlin, Germany
e-mail: mehrmann@math.tu-berlin.de

A. Miȩdlar
e-mail: miedlar@math.tu-berlin.de

🖉 Springer

1 Introduction

We discuss the error reduction (also called saturation) property in adaptive finite element methods (AFEM). This property is a frequent assumption that a mesh refinement procedure will eventually lead to convergence of the approximate finite element solution to the exact solution. For boundary value problems associated with linear second order elliptic partial differential equations (PDEs), this assumption has been reasonably justified in Dörfler and Nochetto [8], and is used in a number of publications, [1,3,13,18].

For the eigenvalue problem associated with partial differential operator, the mathematical justification of the ad hoc saturation assumption, see e.g. Mehrmann and Miedlar [15], Neymeyr [16], is widely open even in the asymptotic range for extremely small mesh-sizes.

This paper lays the mathematical justification of the saturation property for the simplest model problem of an elliptic PDE eigenvalue problem, i.e., for the Laplace operator on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$, which is the problem of determining an eigenvalue/eigenfunction pair $(\lambda, u) \in \mathbb{R} \times \left(H_0^1(\Omega; \mathbb{R}) \cap H_{loc}^2(\Omega; \mathbb{R}) \right)$ such that

$$- \Delta u = \lambda u \quad \text{in } \Omega \qquad \text{and} \qquad u = 0 \quad \text{on } \partial\Omega. \tag{1.1}$$

It is well known, that problem (1.1) has a countable number of eigenvalue/eigenfunction pairs with positive eigenvalues that can be ordered increasingly [2].

We will discuss the case of determining one single and simple eigenvalue λ, i.e., we assume that there is no other eigenvalue in a small neighborhood of λ, and we present an AFEM for a *single* sequence of eigenvalue/eigenfunction pairs $(\lambda_\ell, u_\ell)_{\ell \in \mathbb{N}_0}$ of the discretized problem. For this method we will prove the *asymptotic saturation condition*, that there exists a constant $0 < \rho < 1$ such that any two subsequent mesh refinement levels ℓ and $\ell + 1$ with maximal mesh-sizes H_ℓ, $H_{\ell+1}$ satisfy

$$|\lambda - \lambda_{\ell+1}| + \|u - u_{\ell+1}\|^2 \leq \rho \left(|\lambda - \lambda_\ell| + \|u - u_\ell\|^2 \right) + 2\lambda_{\ell+1}^3 H_\ell^4. \tag{1.2}$$

Here and throughout this paper, $|\cdot|$ denotes the Euclidean norm (or the modulus) while $\|\cdot\| := a(\cdot, \cdot)^{1/2}$ is the energy norm associated to the weak form of the PDE while $\| \cdot \|$ abbreviates the norm in $L^2(\Omega)$.

The adaptive algorithm utilizes a patch-oriented refinement process based on the newest-vertex bisection without interior node property and there is no need to compute any higher-order or fine-grid solutions.

Note that the remainder term of *oscillations* $\lambda_{\ell+1}^3 H_\ell^4$ in (1.2) is explicit even with the multiplicative constant 2 in front of it. This justifies the assumption of Mehrmann and Miedlar [15], Neymeyr [16] that this remainder may be neglected for sufficiently small mesh-sizes.

Our results complement those of Carstensen and Gedicke [6] where it is shown that oscillations can be neglected under certain particular assumptions on the meshes; however the same global arguments do not apply in the present situation.

The numerical examples of Sect. 5 verify the (asymptotic) reliability and efficiency of the hierarchical error estimator and therefore confirm the (asymptotic) saturation

property. For the smallest eigenvalue, the mesh-size restrictions on H_0 are empirically not visible, but they are certainly more severe for larger eigenvalues with much more oscillatory eigenfunctions.

The saturation property has to be considered in comparison to the error estimator reduction in the convergence analysis of adaptive finite element eigenvalue solvers [6,7,11]. In explicit residual-based error estimators, the mesh-size enters as a weight and hence reduces under refinement. This implies a reduction property of such error estimators and eventually leads to linear convergence of some total error which is a convex combination of the error estimator and the error; cf. e.g. ([6, Thm 5.2]), ([7, Lemma 5.3]), ([11, Thm 5.2]). In contrast to this, the saturation property describes the reduction (1.2) of the error terms without involving any error estimator contribution, but with immediate important applications in the context of the solution of the algebraic eigenvalue problems that have to be solved at each level of refinement [15,16]. The proofs are rather independent, e.g., the saturation property (1.2) cannot be proved by simply reducing the mesh-size.

The outline of the remaining part of this paper is as follows. Section 2 describes the AFEM based on patch refinement. The discrete efficiency of the edge residual a posteriori error estimator is introduced in Sect. 3. The proof of the saturation property and its equivalence to the reliability and the efficiency of the hierarchical error estimator follow in Sect. 4. Section 5 verifies the theoretical results for some numerical benchmark problems on the unit square, the L-shaped domain, and two isospectral domains.

Throughout this paper, standard notations on Sobolev and Lebesgue spaces apply, see e.g. [4,12] and $f_T f := \int_T f/|T|dx$ denotes the integral mean over a triangle T of area $|T|$. Similar notation applies to integral means over edges or patches.

The notation $x \lesssim y$ abbreviates the inequality $x \leq C_1 y$ and $x \approx y$ the inequalities $C_2 y \leq x \leq C_1 y$ with constants $C_1, C_2 > 0$ which do not depend on the mesh-size.

We will call an error estimator η_ℓ *efficient* and *reliable* if there exist mesh-size independent constants C_{eff} (efficiency constant) and C_{rel} (reliability constant) such that

$$C_{\text{eff}} \eta_\ell \leq \|u - u_\ell\| \leq C_{\text{rel}} \eta_\ell.$$

2 Adaptive finite element method

Consider the elliptic eigenvalue problem (1.1) and let $\mathcal{T}_\ell$ denote a shape-regular triangulation of Ω into triangles, see e.g. [4,18]. The linear conforming finite element space for the triangulation $\mathcal{T}_\ell$ is defined by

$$V_\ell := \left\{ v \in H_0^1(\Omega) : \text{ for all } T \in \mathcal{T}_\ell, \ v|_T \text{ is affine} \right\}.$$

The AFEM computes a sequence of discrete subspaces

$$V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_\ell \subset V$$

via successive local refinement of the underlying mesh $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2, \ldots$ of the domain Ω through a loop of the form

$$\text{Solve} \rightarrow \text{Estimate} \rightarrow \text{Mark} \rightarrow \text{Refine.}$$

In the following we briefly summarize these components, for details see Brenner and Scott [4], Carstensen and Gedicke [6]. The input consists of a shape-regular triangulation $\mathcal{T}_0$ (with some initialization of the reference edges) and some bulk parameter $0 < \theta \leq 1$.

Solve. The weak formulation of problem (1.1) consists of determining an eigenvalue/eigenfunction pair $(\lambda, u) \in \mathbb{R} \times V := \mathbb{R} \times H_0^1(\Omega; \mathbb{R})$ with $b(u, u) = 1$ and

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

where the bilinear forms $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ are defined by

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

They induce the norms $\|\cdot\| := |\cdot|_{H^1(\Omega)}$ on V and $\|\cdot\| := \|\cdot\|_{L^2(\Omega)}$ on $L^2(\Omega)$.

The corresponding discrete eigenvalue problem consists of determining an eigenvalue/eigenfunction pair $(\lambda_\ell, u_\ell) \in \mathbb{R} \times V_\ell$ with $b(u_\ell, u_\ell) = 1$ and

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

Using the coordinate representation, the discrete eigenvalue problem leads to the finite-dimensional generalized algebraic eigenvalue problem

$$A_\ell x_\ell = \lambda_\ell B_\ell x_\ell$$

for the stiffness matrix $A_\ell = [a(\varphi_i, \varphi_j)]_{i,j=1,\ldots,N_\ell}$ and the mass matrix $B_\ell = [b(\varphi_i, \varphi_j)]_{i,j=1,\ldots,N_\ell}$ associated with the nodal basis functions $\varphi_1, \ldots, \varphi_{N_\ell}$ of $V_\ell = \{\varphi_1, \ldots, \varphi_{N_\ell}\}$, with the discrete eigenvector $x_\ell =: [x_{\ell,1}, \ldots, x_{\ell,N_\ell}]^T$. The approximated eigenfunction is then expressed as

$$u_\ell = \sum_{k=1}^{N_\ell} x_{\ell,k} \varphi_k \in V_\ell.$$

Estimate. The error in the eigenvalue/eigenfunction pair can be estimated a posteriori via

$$|\lambda - \lambda_\ell| + \|u - u_\ell\|^2 \lesssim \mu_\ell^2 := \|u_{\ell-1} - u_\ell\|^2.$$

Recall that $\lesssim$ denotes an inequality that holds up to a multiplicative constant.

Such an a posteriori error estimator for the discussed Laplace eigenvalue problem has been presented in Mehrmann and Miedlar [15]. Hierarchical a posteriori error estimators based on edge bubble-functions were considered in Grubišić and Ovall [14], Neymeyr [16] for the eigenvalues and eigenfunctions.

Mark. For the triangulation $\mathcal{T}_\ell$ let $\mathcal{N}_\ell$ (resp. $\mathcal{N}_\ell(\Omega)$) denote the set of nodes (resp. interior nodes) and let $\mathcal{E}_\ell$ (resp. $\mathcal{E}_\ell(\Omega)$) denote the set of edges (resp. interior edges). For a node $z \in \mathcal{N}_\ell$, we denote by $\mathcal{E}_\ell(z) \subseteq \mathcal{E}_\ell$ the subset of edges that share the node z and by ω_z the union of triangles in $\mathcal{T}_\ell$ that share the node z. The maximal mesh-size is denoted by $H_\ell := \max_{T \in \mathcal{T}_\ell} \mathrm{diam}(T)$. For $E \in \mathcal{E}_\ell(\Omega)$ let $T_+, T_- \in \mathcal{T}_\ell$ be the two neighboring triangles such that $E = T_+ \cap T_-$. The jump of the discrete gradient ∇u_ℓ along an inner edge $E \in \mathcal{E}_\ell(\Omega)$ in normal direction ν_E, pointing from T_+ to T_-, is defined by $[\nabla u_\ell] \cdot \nu_E := (\nabla u_\ell|_{T_+} - \nabla u_\ell|_{T_-}) \cdot \nu_E$.

The patch-oriented marking strategy employs the edge residual a posteriori error estimator for the eigenvalue problem, see [6,10],

$$\eta_\ell^2 := \sum_{E \in \mathcal{E}_\ell(\Omega)} \eta_\ell^2(E) \quad \text{with} \quad \eta_\ell^2(E) := |E| \|[\nabla u_\ell] \cdot \nu_E\|_{L^2(E)}^2, \tag{2.1}$$

which is reliable and efficient for sufficiently small mesh-size H_0 [6], in the sense that

$$\|u - u_\ell\| \approx \eta_\ell. \tag{2.2}$$

Based on the local refinement indicators $\eta_\ell(E)$, nodes are marked for refinement. Let $\mathcal{M}_\ell \subseteq \mathcal{N}_\ell(\Omega)$ be the minimal set of refinement nodes such that for $0 < \theta \leq 1$ the bulk criterion [9] is fulfilled, i.e.,

$$\theta \sum_{z \in \mathcal{N}_\ell(\Omega)} \eta_\ell^2(\mathcal{E}_\ell(z)) \leq \sum_{z \in \mathcal{M}_\ell} \eta_\ell^2(\mathcal{E}_\ell(z)).$$

Refine. For each refinement node $z \in \mathcal{M}_\ell \subseteq \mathcal{N}_\ell(\Omega)$, mark all edges $\mathcal{E}_\ell(z)$ for refinement. In order to preserve the quality of the mesh, i.e., the maximal angle condition, additional edges have to be marked by the closure algorithm before refinement. For each triangle T let one edge be the uniquely defined *reference* edge $E(T)$. The closure algorithm marks additional edges such that once an edge of a triangle T is marked for refinement, its *reference* edge $E(T)$ is marked for refinement as well. The refinement $\mathcal{T}_{\ell+1}$ is then computed by the application of one of the rules from Figure 1 which shows the sub-triangles together with the new *reference* edges. Note that all triangles $T \subseteq \overline{\omega}_z$, $z \in \mathcal{M}_\ell$, are refined either *red* or *blue*.

In the following sections we analyze the properties of this adaptive FEM technique.

3 Discrete efficiency

This section introduces the discrete efficiency of η_ℓ as defined in (2.1). A key ingredient for the proof is the following fine-grid function. Let $\varphi_z \in V_\ell$ denote the shape function associated with the node $z \in \mathcal{N}_\ell(\Omega)$. Under the assumption that all edges $\mathcal{E}_\ell(z)$ are
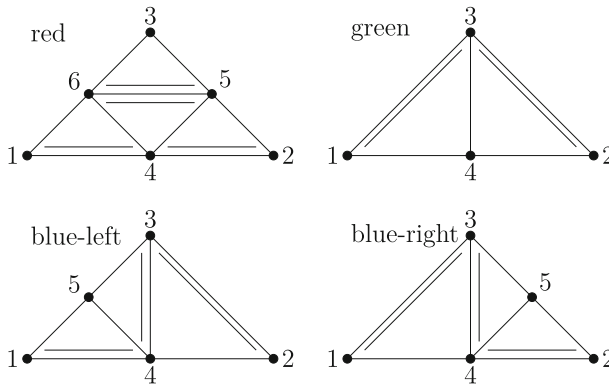
Fig. 1 Refinement rules: sub-triangles with corresponding reference edges depicted with a second edge

refined, let ψ_E be the linear shape function of the refined triangulation $\mathcal{T}_{\ell+1}$ associated with the midpoint of the edge $E \in \mathcal{E}_\ell$, and introduce

$$\Phi_z := \varphi_z - \sum_{E \in \mathcal{E}_\ell(z)} \psi_E \in H_0^1(\omega_z) \subseteq V.$$

Theorem 3.1 *Consider the adaptive FEM of Sect. 2. For some node $z \in \mathcal{N}_\ell(\Omega)$ let all edges $\mathcal{E}_\ell(z)$ be bisected in $\mathcal{T}_{\ell+1}$. If z is not opposite to the reference edge $E(T)$ or T is refined by red-refinement for at least one triangle T of ω_z, (see Fig. 1), then*

$$\fint_{\omega_z} \Phi_z \, dx \approx 1 \quad and \quad \fint_E \Phi_z \, ds = 0 \quad \text{for all } E \in \mathcal{E}_\ell(\omega_z).$$

Proof The second assertion, that $\fint_E \Phi_z \, ds = 0$ for all $E \in \mathcal{E}_\ell(\omega_z)$ follows directly from the definition of Φ_z. For the first assertion all possible sub-triangulations together with the values of $\fint_T \Phi_z dx$ are depicted in Fig. 2. Note that the sub-triangulations for a triangle T of ω_z that result in values $\fint_T \Phi_z dx = 0$ are excluded by assumption and that all other possible sub-triangulations share the same sign for $\fint_T \Phi_z dx$.

Theorem 3.2 (Discrete efficiency) *Consider the adaptive FEM of Sect. 2. Then for any refinement level $\ell \in \mathbb{N}_0$ the following estimate holds*

$$\eta_\ell \lesssim \|u_\ell - u_{\ell+1}\| + \lambda_{\ell+1}^{3/2} H_\ell^2.$$

Proof In the *first step* observe that the bulk criterion implies that

$$\eta_\ell^2 \le \sum_{z \in \mathcal{N}_\ell(\Omega)} \eta_\ell^2(\mathcal{E}_\ell(z)) \le \theta^{-1} \sum_{z \in \mathcal{M}_\ell} \eta_\ell^2(\mathcal{E}_\ell(z)). \tag{3.1}$$
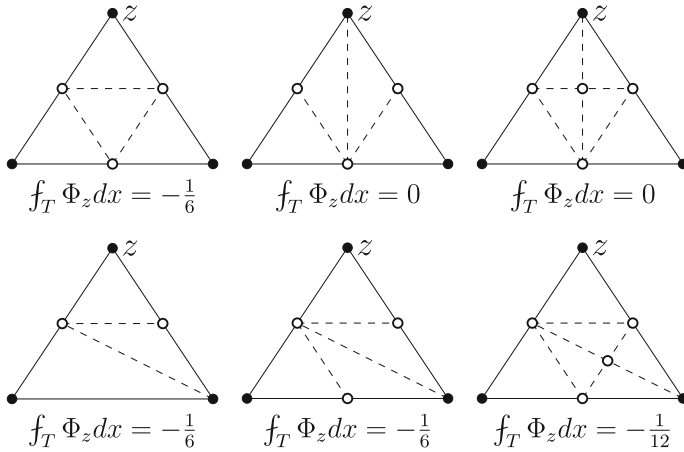
Fig. 2 All possible sub-triangulations for a *triangle* $T \subset \overline{\omega}_z$ in the proof of Theorem 3.1 with values of $f_T \Phi_z dx$

The *second step* is to show that any refinement node $z \in \mathcal{M}_\ell$ and any edge E of $\mathcal{E}_\ell(z)$ satisfy

$$\eta_\ell(E) \lesssim \|\nabla(u_\ell - u_{\ell+1})\|_{L^2(\omega_z)} + \lambda_{\ell+1} \text{diam}(\omega_z)^2 \|\nabla u_{\ell+1}\|_{L^2(\omega_z)}. \tag{3.2}$$

Since $[\nabla u_\ell] \cdot \nu_E$ is constant along the edge E of length $|E|$ with some sign $\pm$ as indicated below, it follows that

$$\pm \eta_\ell(E) = |E|([\nabla u_\ell] \cdot \nu_E).$$

The edge basis function ψ_E from the beginning of this section satisfies $|E| = 2 \int_E \psi_E \, ds$. Hence,

$$\pm \eta_\ell(E)/2 = \int_E \psi_E [\nabla u_\ell] \cdot \nu_E \, ds.$$

Let $T_\pm \in T_\ell$ denote the two triangles that share the edge E. Theorem 3.1 shows that $\int_E \Phi_z ds = 0$. With α being the value from Fig. 3, this implies that

$$\pm \eta_\ell(E)/2 = \int_E (\psi_E + \alpha \Phi_z) [\nabla u_\ell] \cdot \nu_E \, ds.$$

Note that $\Delta u_\ell|_{T_\pm} \equiv 0$ and that the function $v_{\ell+1} := \psi_E + \alpha \Phi_z \in V_{\ell+1}$ satisfies $\int_F v_{\ell+1} ds = 0$ on all other edges $F \in \mathcal{E}_\ell(\omega_z) \backslash E$. Therefore, the piecewise Gauss divergence theorem [12] leads to
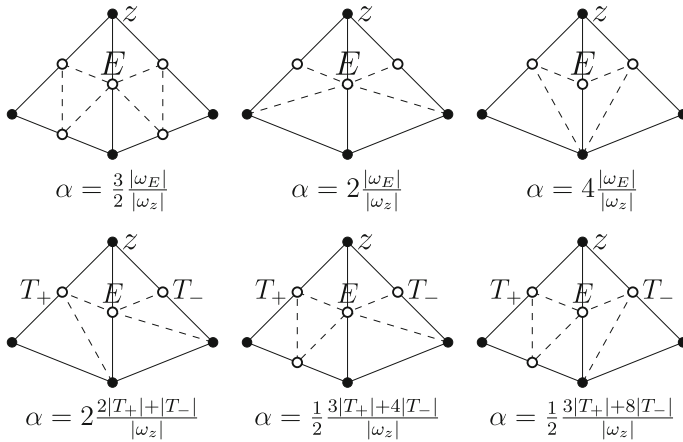
Fig. 3 All possible sub-triangulations of ω_E in the proof of Theorem 3.2 with values of α

$$\pm\eta_\ell(E)/2 = \int\limits_{\omega_z} \nabla v_{\ell+1} \cdot \nabla u_\ell dx.$$

In fact all the volume contributions and all other edge contributions vanish. Hence,

$$\pm\eta_\ell(E)/2 = \int\limits_{\omega_z} \nabla v_{\ell+1} \cdot \nabla(u_\ell - u_{\ell+1})dx + \int\limits_{\omega_z} \nabla v_{\ell+1} \cdot \nabla u_{\ell+1}dx. \qquad (3.3)$$

The first term in (3.3) is estimated via the Cauchy-Schwarz inequality [4] and the discrete estimate $\|\nabla v_{\ell+1}\|_{L^2(\omega_z)} \lesssim 1$, as

$$\int\limits_{\omega_z} \nabla v_{\ell+1} \cdot \nabla(u_\ell - u_{\ell+1})dx \le \|\nabla(u_\ell - u_{\ell+1})\|_{L^2(\omega_z)}.$$

Since $v_{\ell+1}$ is supported on ω_z, the second term in (3.3) can be written as

$$\int\limits_{\omega_z} \nabla v_{\ell+1} \cdot \nabla u_{\ell+1}dx = a(u_{\ell+1}, v_{\ell+1}).$$

Since $v_{\ell+1} \in V_{\ell+1}$, we then have

$$a(u_{\ell+1}, v_{\ell+1}) = \lambda_{\ell+1}b(u_{\ell+1}, v_{\ell+1}).$$

The choice of α as in Fig. 3 shows that $\fint_{\omega_z} v_{\ell+1}dx = 0$, and hence $c_z := \fint_{\omega_z} u_{\ell+1}dx$ satisfies

$$a(u_{\ell+1}, v_{\ell+1}) = \lambda_{\ell+1} b(u_{\ell+1} - c_z, v_{\ell+1}).$$

Applying Cauchy-Schwarz, Poincaré-Friedrich's inequality [4] plus the aforementioned discrete estimate

$$\|v_{\ell+1}\|_{L^2(\omega_z)} \lesssim \mathrm{diam}(\omega_z) \|\nabla v_{\ell+1}\|_{L^2(\omega_z)} \lesssim \mathrm{diam}(\omega_z)$$

shows that

$$\lambda_{\ell+1} b(u_{\ell+1} - c_z, v_{\ell+1}) \lesssim \lambda_{\ell+1} \mathrm{diam}(\omega_z)^2 \|\nabla u_{\ell+1}\|_{L^2(\omega_z)}.$$

The combination of the previous four estimates shows that

$$\int_{\omega_z} \nabla v_{\ell+1} \cdot \nabla u_{\ell+1} dx \lesssim \lambda_{\ell+1} \mathrm{diam}(\omega_z)^2 \|\nabla u_{\ell+1}\|_{L^2(\omega_z)}.$$

Altogether, this second step proves (3.2). *Step three* combines (3.1)–(3.2) with the finite overlap of all the patches to conclude that

$$\eta_\ell \lesssim \||u_\ell - u_{\ell+1}\|| + \lambda_{\ell+1} H_\ell^2 \||u_{\ell+1}\||.$$

This and the identity $\||u_{\ell+1}\|| = \lambda_{\ell+1}^{1/2}$ finish the proof. $\qquad\square$

4 Saturation property

This section is devoted to the main result of this paper, the proof of the saturation property. It is also remarked that the saturation property is equivalent to the reliability of the hierarchical a posteriori error estimator.

Throughout this section, suppose that $(\lambda_\ell, u_\ell) \in \mathbb{R} \times V_\ell$ as well as $(\lambda_{\ell+1}, u_{\ell+1}) \in \mathbb{R} \times V_{\ell+1}$ is some discrete eigenvalue/eigenfunction pair associated with the continuous pair $(\lambda, u) \in \mathbb{R} \times V$ on the level ℓ and $\ell + 1$, respectively, and set $e_\ell := u - u_\ell$ and $e_{\ell+1} := u - u_{\ell+1}$. The eigenvalue error and the errors with respect to the norms $\||\cdot\||$ and $\|\cdot\|$ satisfy [17]

$$\||u - u_\ell\||^2 = \lambda \|u - u_\ell\|^2 + \lambda_\ell - \lambda. \tag{4.1}$$

Furthermore, the following regularity result [6,10] holds for some $0 < C_{\mathrm{reg}} < \infty$,

$$\|u - u_\ell\| \leq C_{\mathrm{reg}} H_\ell^s \||u - u_\ell\||, \tag{4.2}$$

where the regularity exponent $0 < s \leq 1$ depends on the interior angles of the polygonal domain Ω and $s > 1/2$ holds for the pure Dirichlet boundary conditions of (1.1).

The proof of the saturation assumption requires the following quasi-orthogonality.

Theorem 4.1 (Quasi-orthogonality) *Let $\mathcal{T}_{\ell+1}$ be a refinement of the triangulation $\mathcal{T}_\ell$ on some level ℓ in the adaptive FEM of Sect. 2. Then there exists $\varepsilon \lesssim H_0^{2s}$ such that*

$$\|u_{\ell+1} - u_\ell\|^2 \leq (1 + \varepsilon)\|e_\ell\|^2 - \|e_{\ell+1}\|^2. \tag{4.3}$$

Proof The quasi-orthogonality result [7, Lemma 3.1] implies that

$$\|u_{\ell+1} - u_\ell\|^2 \leq \|e_\ell\|^2 - \|e_{\ell+1}\|^2 + \lambda\|e_{\ell+1}\|^2 + \lambda_{\ell+1}\|u_{\ell+1} - u_\ell\|^2.$$

Let $G_\ell : V \to V_\ell$ denote the Galerkin projection, $a(v - G_\ell v, \cdot)|_{V_\ell} = 0$ for all $v \in V$. The proof of [6, Theorem 3.1] shows that

$$\|u - u_\ell\| \lesssim H_\ell^s \|u - G_\ell u\|.$$

Since $V_\ell \subset V_{\ell+1}$, the best approximation property of the Galerkin projection [4] leads to

$$\|u - u_\ell\| + \|u - u_{\ell+1}\| \leq 2C_{\text{reg}} H_\ell^s \|u - u_\ell\|.$$

This and the min–max principle [17] imply (4.3) with $\varepsilon := (\lambda + 4\lambda_0)C_{\text{reg}}^2 H_0^{2s}$. □

Theorem 4.2 (Saturation property) *Consider the adaptive FEM of Sect. 2 for some $\mathcal{T}_0$ with sufficiently small maximal mesh-size H_0. Then there exists $0 \leq \varrho < 1$ such that for all $\ell \in \mathbb{N}_0$ the following inequalities hold*

$$\|u - u_{\ell+1}\|^2 \leq \varrho\|u - u_\ell\|^2 + \lambda_{\ell+1}^3 H_\ell^4; \tag{4.4}$$
$$|\lambda - \lambda_{\ell+1}| \leq \varrho|\lambda - \lambda_\ell| + \lambda_{\ell+1}^3 H_\ell^4. \tag{4.5}$$

Proof For sufficiently small H_0 Theorem 3.2 and (2.2) imply that

$$\|u - u_\ell\|^2 \lesssim \|u_\ell - u_{\ell+1}\|^2 + \lambda_{\ell+1}^3 H_\ell^4.$$

This and the quasi-orthogonality of Theorem 4.1 imply the existence of some generic constant $0 < c \leq 1$ such that

$$c\|u - u_\ell\|^2 \leq (1 + \varepsilon)\|u - u_\ell\|^2 - \|u - u_{\ell+1}\|^2 + \lambda_{\ell+1}^3 H_\ell^4.$$

This is equivalent to

$$\|u - u_{\ell+1}\|^2 \leq (1 + \varepsilon - c)\|u - u_\ell\|^2 + \lambda_{\ell+1}^3 H_\ell^4.$$

The assertion follows from this with $0 \leq \varrho := (1 + \varepsilon - c) < 1$ for sufficiently small H_0. To prove the second saturation property (4.5), recall the inequalities (4.1) and (4.2) which imply

$$|\lambda - \lambda_{\ell+1}| \leq \|u - u_{\ell+1}\|^2;$$
$$\|u - u_\ell\|^2 \leq |\lambda - \lambda_\ell| + \lambda C_{\text{reg}}^2 H_\ell^{2s} \|u - u_\ell\|^2.$$

For any $H_0 < \lambda^{-1/(2s)} C_{\text{reg}}^{-1/s}$, this shows

$$\|u - u_\ell\|^2 \leq \frac{|\lambda - \lambda_\ell|}{1 - \lambda C_{\text{reg}}^2 H_0^{2s}}. \tag{4.6}$$

This and (4.4) lead to (4.5) with the constant

$$0 \leq \varrho := (1 + \varepsilon - c)(1 - \lambda C_{\text{reg}}^2 H_0^{2s})^{-1} < 1.$$

$\square$

Note that the saturation property implies (1.2). A surprising consequence of this saturation property is that the higher-order terms $\lambda_{\ell+1}^3 H_\ell^4$ do not depend on the (possibly reduced) convergence rates of the errors $\|u - u_\ell\|^2 + |\lambda - \lambda_\ell|$ on (non-convex) polygonal domains.

In the subsequent theorem we show that the saturation property is actually equivalent to the reliability of the adaptive FEM.

Theorem 4.3 (Saturation $\Leftrightarrow$ reliability) *Consider the adaptive FEM of Sect. 2 with sufficiently small H_0. Then for some $0 \leq \varrho < 1$, $0 < c \leq 1$, and all $\ell \in \mathbb{N}_0$, the following inequalities (4.7) and (4.8) are equivalent*

$$\|u - u_{\ell+1}\| \leq \varrho \|u - u_\ell\| + \lambda_{\ell+1}^{3/2} H_\ell^2; \tag{4.7}$$
$$c\|u - u_\ell\| \leq \|u_\ell - u_{\ell+1}\| + \lambda_{\ell+1}^{3/2} H_\ell^2. \tag{4.8}$$

Proof We first show that (4.7) with $0 \leq \varrho < 1$ implies (4.8) with $c := 1 - \varrho$. For any $0 \leq \varrho < 1$ the triangle inequality plus the saturation property (4.7) yield that

$$\|u - u_\ell\| \leq \|u - u_{\ell+1}\| + \|u_\ell - u_{\ell+1}\|$$
$$\leq \varrho \|u - u_\ell\| + \|u_\ell - u_{\ell+1}\| + \lambda_{\ell+1}^{3/2} H_\ell^2.$$

This proves (4.8) with $c := 1 - \varrho$.

For the converse we show that (4.8) with $0 < c \leq 1$ implies (4.7) with $\varrho := (1 + \varepsilon - c^2/2)$. The quasi-orthogonality of Theorem 4.1 leads to

$$\|u - u_{\ell+1}\|^2 \leq (1 + \varepsilon)\|u - u_\ell\|^2 - \|u_\ell - u_{\ell+1}\|^2.$$

Inequality (4.8) and Young's inequality [12], result in

$$-\|u_\ell - u_{\ell+1}\|^2 \leq -c^2 \|u - u_\ell\|^2/2 + \lambda_{\ell+1}^3 H_\ell^4.$$

The combination of these inequalities leads to

$$\|u - u_{\ell+1}\|^2 \le (1 + \varepsilon - c^2/2)\|u - u_\ell\|^2 + \lambda_{\ell+1}^3 H_\ell^4.$$

This proves (4.7) with $0 \le \varrho := 1 + \varepsilon - c^2/2 < 1$ for sufficiently small H_0. $\qquad\square$

Remark 4.4 It is shown in Carstensen and Gedicke [6] that $\|u - u_\ell\| \approx \eta_\ell$ holds for sufficiently small H_0 without any higher-order terms. But this is an estimate for the error $\|u - u_\ell\|$ in contrast to the estimate (4.8) that involves the discrete error $\|u_\ell - u_{\ell+1}\|$. The arguments in the proofs of Carstensen and Gedicke [6] are global, whereas the present analysis employs the discrete efficiency of Theorem 3.2 that does not allow for a global L^2-projection-type argument. Hence (4.8) includes the higher-order oscillation terms.

Theorem 4.5 (Efficiency) *Consider the adaptive FEM of Sect. 2. Let H_0 be sufficiently small such that the saturation property of Theorem 4.2 holds for $0 \le \varrho < 1$. Then*

$$\|u_\ell - u_{\ell+1}\| \le 2\|u - u_\ell\| + \lambda_{\ell+1}^{3/2} H_\ell^2 \quad and \quad \|u_\ell - u_{\ell+1}\| \lesssim \|u - u_\ell\|.$$

Proof The triangle inequality reads

$$\|u_\ell - u_{\ell+1}\| \le \|u - u_{\ell+1}\| + \|u - u_\ell\|.$$

The first assertion then follows from the saturation property of Theorem 4.2. For a proof of the second inequality, the min-max principle [17] and (4.1) imply that

$$|\lambda - \lambda_{\ell+1}| \le |\lambda - \lambda_\ell| \le \|u - u_\ell\|^2.$$

Together with (4.6) this shows, that for sufficiently small H_0

$$\|u - u_{\ell+1}\|^2 \le \frac{\|u - u_\ell\|^2}{1 - \lambda C_{\text{reg}}^2 H_0^{2s}}. \tag{4.9}$$

$$\square$$

These estimates hold for all *simple* eigenvalues in the spectrum. However, since the constant present in the upper bound depends on the exact eigenvalue λ, we require the initial triangulation to be reasonably fine in order to obtain *reliable* and *efficient* approximations of larger eigenvalues. On the linear algebra level, due to the use of the Krylov subspace method, we obviously require a larger number of iteration steps for larger eigenvalues, otherwise we would have to employ a shift and invert strategy.

5 Numerical examples

This section is devoted to numerical examples for the solution of the model problem (1.1) on three different domains Ω: the unit square, the L-shaped domain and the isospectral domains.

5.1 Preliminary remarks

The numerical experiments show the performance of the proposed AFEM algorithm in comparison to uniform mesh-refinement and compare the two a posteriori error estimators

$$\eta_\ell^2 := \sum_{E \in \mathcal{E}_\ell(\Omega)} |E| \|[\nabla u_\ell] \cdot \nu_E\|_{L^2(E)}^2 \quad \text{and} \quad \mu_\ell := \|u_\ell - u_{\ell-1}\|.$$

Note that Theorem 4.3 shows that for any level ℓ and sufficiently small H_0

$$\|u - u_\ell\| \lesssim \mu_{\ell+1} + \lambda_{\ell+1}^{3/2} H_\ell^2.$$

The use of this estimate, however, requires the knowledge of $u_{\ell+1}$. On the other hand, for sufficiently small initial mesh-size H_0 (4.9) shows that $\|u - u_\ell\| \lesssim \|u - u_{\ell-1}\|$. The combination with the aforementioned estimate (employed at level $\ell - 1$) gives

$$\|u - u_\ell\| \lesssim \mu_\ell + \lambda_\ell^{3/2} H_\ell^2.$$

In other words, μ_ℓ is a reliable a posteriori error estimator if H_ℓ is small. Throughout all of our numerical experiments, μ_ℓ is used as a posteriori error estimator on level ℓ.

5.2 Unit square

Consider the model problem (1.1) on the unit square $\Omega = (0, 1)^2$. The first eigenvalue/eigenfunction pair is

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

Since the solution is smooth, either uniform and adaptive mesh-refinement leads to optimal convergence rates of $\mathcal{O}(N_\ell^{-1})$ for $|\lambda - \lambda_\ell|$ in Fig. 4 and of order $\mathcal{O}(N_\ell^{-1/2})$ for $\|u - u_\ell\|$ in Fig. 5. Note that for uniform refinement $\mathcal{O}(N_\ell^{-1/2}) = \mathcal{O}(H_\ell)$.

We observe in the experiments that the hierarchical a posteriori error estimators μ_ℓ^2 and μ_ℓ are closer to the eigenvalue and eigenfunction errors in the energy norm than the edge residual a posteriori error estimators η_ℓ^2 and η_ℓ. For adaptive refinement μ_ℓ^2 and μ_ℓ are almost exact. This is an empirical observation which is not mathematically justified by our theoretical analysis, because our estimates contain generic constants.

Note that for the uniform refinement μ_ℓ is an upper bound while for the adaptive refinement μ_ℓ provides a lower bound of the eigenfunction error in the energy norm. In contrast to this, η_ℓ is always an upper bound of the eigenfunction error in the energy norm. The same observations are made for μ_ℓ^2, η_ℓ^2 and the eigenvalue error.

5.3 L-shaped domain

Consider the model problem (1.1) on the L-shaped domain $\Omega = (-1, 1)^2 \setminus ([0, 1] \times [-1, 0])$ with the first approximated eigenvalue $\lambda = 9.6397238440219$, see [5]. Since
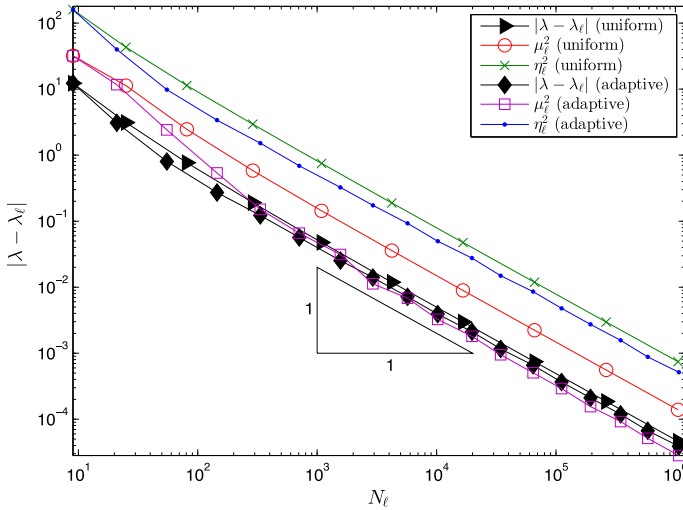
Fig. 4 Convergence history for $|\lambda - \lambda_\ell|$, η_ℓ^2 and μ_ℓ^2 for uniform and adaptive mesh-refinements on the *unit square*
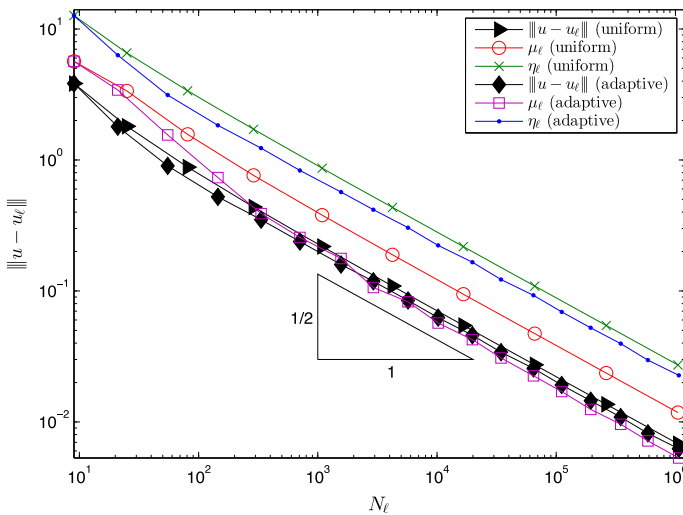

Fig. 5 Convergence history for $\|u - u_\ell\|$, η_ℓ and μ_ℓ for uniform and adaptive mesh-refinements on the *unit square*

the eigenfunction has a singularity, uniform refinement leads to suboptimal convergence rates of order $\mathcal{O}(N_\ell^{-2/3})$, while adaptive refinement leads to empirical optimal convergence rates of order $\mathcal{O}(N_\ell^{-1})$ as displayed in Fig. 6.

As in the previous example μ_ℓ^2 is an upper bound for in the case of uniform meshes and a lower bound in the case of adaptive meshes, while η_ℓ^2 is always an upper bound. In both cases we observe that μ_ℓ^2 is much closer to the eigenvalue error than η_ℓ^2 and that for adaptive refinement μ_ℓ^2 is almost exact.
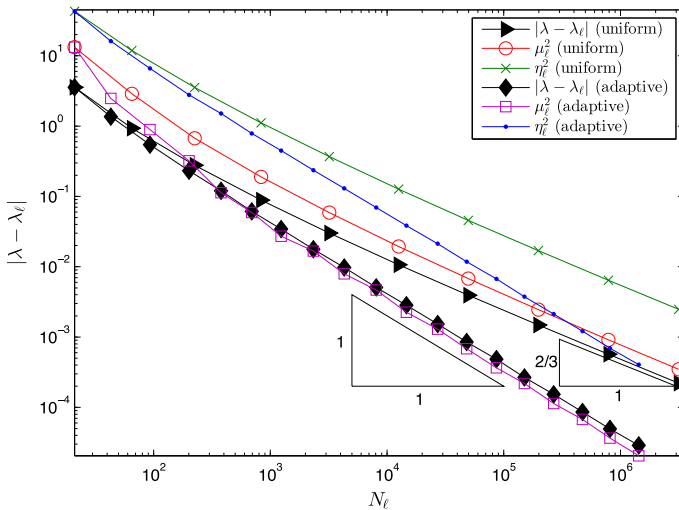
Fig. 6 Convergence history for $|\lambda - \lambda_\ell|$, η_ℓ^2 and μ_ℓ^2 for uniform and adaptive mesh-refinements on the L-shaped domain

Figure 7 displays a sequence of mesh-refinements that are adaptively refined towards the corner singularity.

5.4 Isospectral domains

Consider the model problem (1.1) on the two isospectral domains of Fig. 8 with the approximation of the 50th eigenvalue $\lambda_{50} = 54.187936$, see [5]. Figure 9 shows the convergence history for the eigenvalue error.

We observe that adaptive refinement leads to slightly smaller errors for larger values of N_ℓ than uniform refinement. For uniform refinement both domains are discretized with the same number of degrees of freedom which results in the same approximated values (up to round-off errors) for the eigenvalue error and the a posteriori error estimators. For adaptive refinement the values for the two domains lie asymptotically on the same convergence line.

Note that the pre-asymptotic range for μ_ℓ^2 on adaptive meshes is rather long due to the large eigenvalue and that the pre-asymptotic values for μ_ℓ^2 differ for both domains. Again we observe that asymptotically μ_ℓ^2 is almost exact for adaptive refinement.

This experiment provides numerical evidence that the mesh-size restriction due to the contributions $2\lambda_{\ell+1}^3 H_\ell^3$ in (1.2) can be severe. Despite of this, we observe convergence of the AFEM in all experiments. This follows from the analysis of Carstensen and Gedicke [7].

5.5 Three hierarchical adaptive algorithms

The residual-based AFEM is compared to three different versions of hierarchical AFEMs based on the a posteriori error estimators $\mu_{\ell,k}$, $k = 2, 3, 4$. The hierarchical a
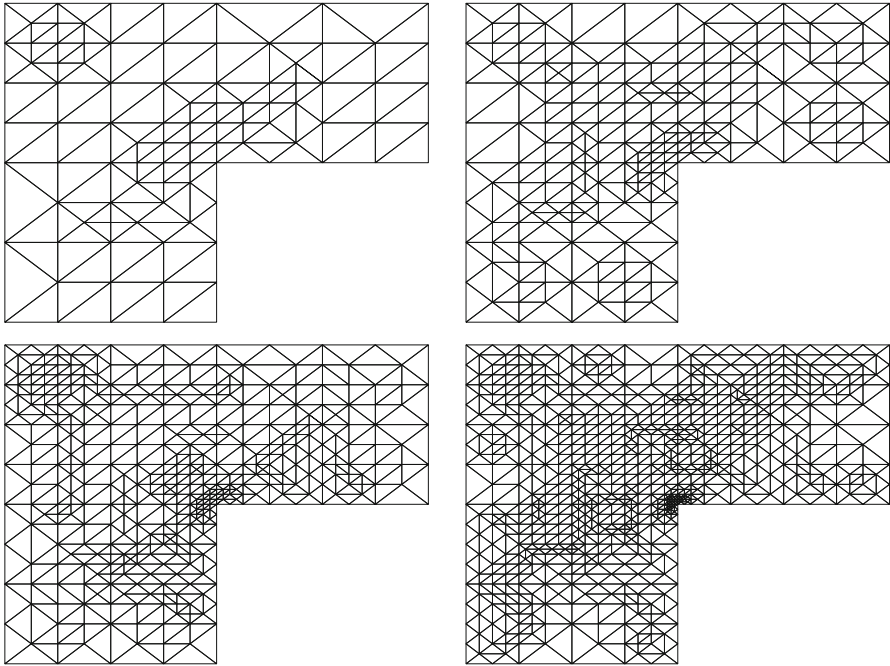
Fig. 7 Sequence of adaptively refined meshes for the L-shaped domain for μ_ℓ with $N_\ell = 93, 201, 378, 694$
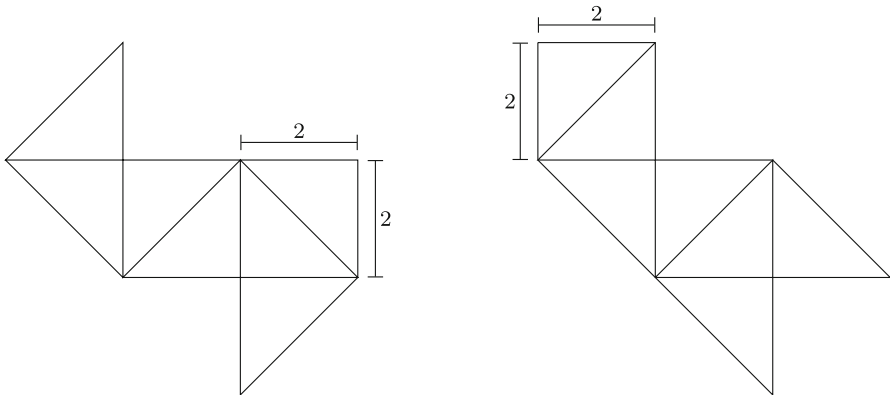

Fig. 8 Two isospectral domains A (*left*) and B (*right*)

posteriori error estimators utilize the fine-grid eigenvalue/eigenfunction pairs $(\hat{\lambda}_\ell, \hat{u}_\ell)$ of the uniform *red*-refinement $\hat{T}_\ell$ of T_ℓ. The first version of the hierarchical a posteriori error estimator reads

$$\mu_{\ell,2}^2 := \sum_{T \in T_\ell} \|\nabla(u_\ell - \hat{u}_\ell)\|_{L^2(T)}.$$

The discrete efficiency of Theorem 3.2 leads (for $\theta = 1$) to the a posteriori error estimator
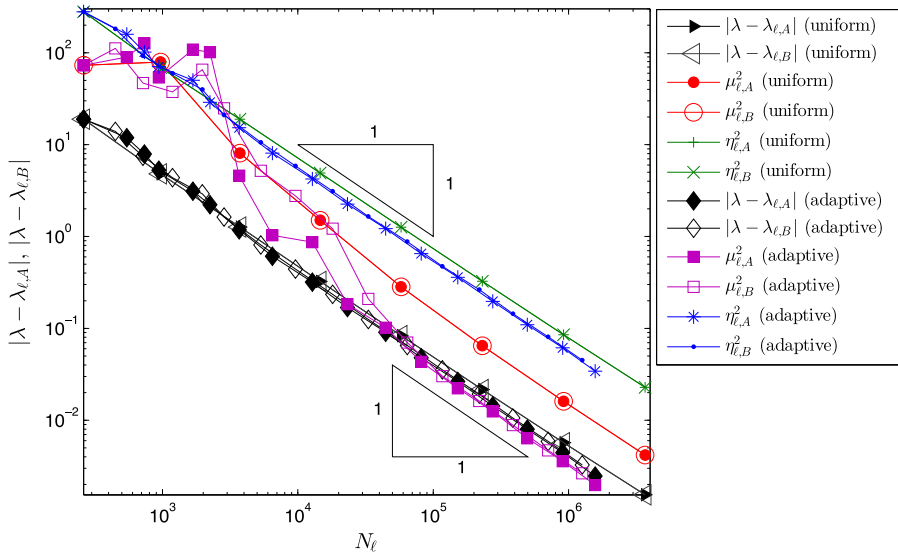
Fig. 9 Convergence history for $|\lambda - \lambda_\ell|$, η_ℓ^2 and μ_ℓ^2 for uniform and adaptive mesh-refinements on both isospectral domains A and B
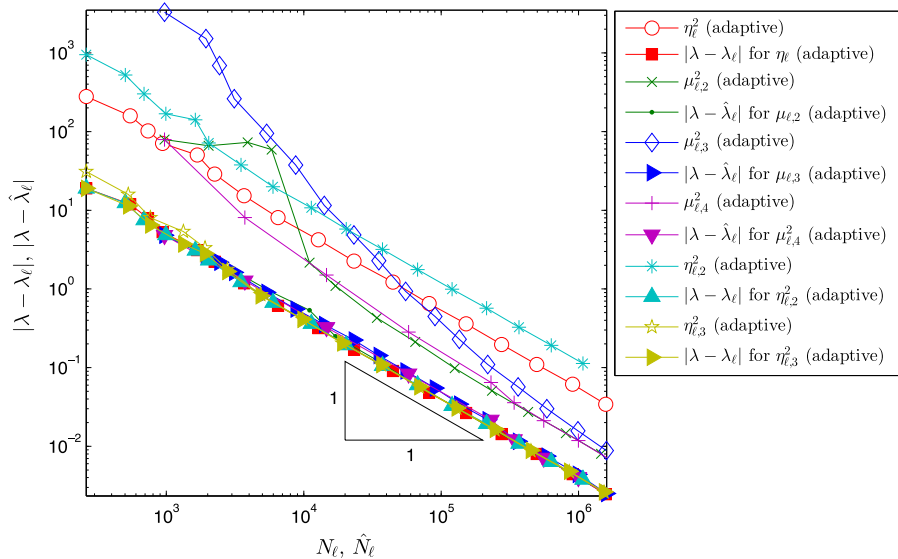

Fig. 10 Convergence history for different adaptive mesh-refinements and the isospectral domain A

$$\mu_{\ell,3}^2 := \sum_{T \in \mathcal{T}_\ell} \left(\|\nabla(u_\ell - \hat{u}_\ell)\|_{L^2(T)} + \hat{\lambda}_\ell \text{diam}(T)^2 \|\nabla \hat{u}_\ell\|_{L^2(T)} \right).$$

The third version utilizes a separate marking strategy based on

$$\mu_{\ell,4} := \|u_\ell - \hat{u}_\ell\| + \hat{\lambda}_\ell^{3/2} H^2.$$

If $\|u_\ell - \hat{u}_\ell\| < \hat{\lambda}_\ell^{3/2} H^2$ then do uniform *red*-refinement, otherwise mark elements accordingly to $\mu_{\ell,2}$. The residual a posteriori error estimator [7,10] reads

$$\eta_{\ell,2}^2 := \sum_{T \in \mathcal{T}_\ell} |T| \|\lambda_\ell u_\ell\|_{L^2(T)}^2 + \sum_{E \in \mathcal{E}_\ell(\Omega)} |E| \|[\nabla u_\ell] \cdot \nu_E\|_{L^2(E)}^2.$$

For the averaging operator $A_\ell : P_0(\mathcal{T}_\ell)^2 \to \{V_\ell^2 \cap C(\Omega)^2\}$ for the nodal basis functions $\varphi_z, z \in \mathcal{N}_\ell$,

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

the averaging a posteriori error estimator [6] reads

$$\eta_{\ell,3}^2 := \sum_{T \in \mathcal{T}_\ell} \|\nabla u_\ell - A_\ell(\nabla u_\ell)\|_{L^2(T)}^2.$$

Figure 10 shows a comparison of the AFEM driven by the a posteriori error estimators $\eta_\ell, \eta_{\ell,2}, \eta_{\ell,3}, \mu_{\ell,2}, \mu_{\ell,3}$, and $\mu_{\ell,4}$ for λ_{50} on the isospectral domain A. Note that $\mu_{\ell,k}$, $k = 2, 3, 4$, are plotted versus the degrees of freedom $\hat{N}_\ell$ of the fine-grid solution and compared to the fine-grid eigenvalue errors $|\lambda - \hat{\lambda}_\ell|$. We observe that all AFEM lead to comparable eigenvalue errors but the behavior of the estimators differs. The averaging estimator $\eta_{\ell,3}^2$ appears to be asymptotically exact. The two residual estimators η_ℓ^2 and $\eta_{\ell,2}^2$ are reliable and efficient from the very beginning and $\eta_{\ell,2}^2$ is larger than η_ℓ^2. The three versions of hierarchical a posteriori error estimators are asymptotically equal but exploit a different pre-asymptotic behavior. In the pre-asymptotic range, the convergence of the error estimator $\mu_{\ell,2}^2$ is too slow while that of $\mu_{\ell,3}^2$ is too fast. In contrast, the separate marking strategy of the estimator $\mu_{\ell,4}^2$ leads to the best overall convergence.

5.6 Conclusions

We have proved the saturation assumption, as well as reliability and efficiency for an AFEM applied to the model problem of computing eigenvalues of the Laplace operator on Lipschitz domains.

The numerical examples confirm the (asymptotic) saturation property of Theorem 4.2.

The presented results apply to any simple eigenvalue, but it is clear that for the approximation of a highly oscillating eigenfunction of a larger eigenvalue the initial mesh needs to be sufficiently fine such that the oscillations are resolved.

The (asymptotic) reliability and efficiency of the hierarchical a posteriori error estimator are empirically verified for eigenvalue and eigenfunction errors in the energy norm.

The proposed AFEM leads to empirical optimal convergence rates for the eigenvalue and eigenfunction errors in the energy norm while uniform refinements lead to suboptimal rates in the presence of corner singularities.

The proposed AFEM is globally convergent in the sense that given any number $k \in \mathbb{N}$ such that the initial problem at level zero is larger than or equal to k and that the step SOLVE computes the discrete eigenvalue λ_ℓ with number k (counted increasingly including multiplicity), then the output of the AFEM is a convergent sequence of real numbers with a limit which is an eigenvalue of (1.1). (For a proof, note that the bulk criterion in the AFEM of this paper implies that of Carstensen and Gedicke [6] and then Theorem 5.1 of that paper implies convergence.) The optimality of AFEM is an open question and the recent progress in Carstensen and Gedicke [7] does not lead to an immediate result here. The present analysis does not employ any volume contributions and the reliability proof relies upon some global L^2-projection which, seemingly, does not allow a localized version to prove discrete reliability.

The comparison between the hierarchical and the edge residual a posteriori error estimator shows that the hierarchical estimator is much closer to the error than the edge residual estimator. However, the hierarchical a posteriori error estimator does not provide guaranteed error control because it underestimates the error.

References

1. Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. Pure and Applied Mathematics. Wiley, New York (2000)
2. Babuška, I., Osborn, J.E.: Eigenvalue Problems. Handbook of Numerical Analysis, vol. II. North Holland, Amsterdam (1991)
3. Bank, R.E., Smith, R.K.: A posteriori error estimates based on hierarchical bases. SIAM J. Numer. Anal. **30**(4), 921–935 (1993)
4. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, Third Ed., Texts in Applied Mathematics, vol. 15. Springer, New York (2008)
5. Betcke, T., Trefethen, L.N.: Reviving the method of particular solutions. SIAM Rev. **47**(3), 469–491 (2005)
6. Carstensen, C., Gedicke, J.: An oscillation-free adaptive FEM for symmetric eigenvalue problems. Numer. Math. **118**(3), 401–427 (2011)
7. Carstensen, C., Gedicke, J.: An adaptive finite element eigenvalue solver of asymptotic quasi-optimal computational complexity. SIAM J. Numer. Anal. **50**(3), 1029–1057 (2012)
8. Dörfler, W., Nochetto, R.H.: Small data oscillation implies the saturation assumption. Numer. Math. **91**(1), 1–12 (2002)
9. Dörfler, W.: A convergent adaptive algorithm for Poisson's equation. SIAM J. Numer. Anal. **33**, 1106–1124 (1996)
10. 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)
11. Dai, X., Xu, J., Zhou, A.: Convergence and optimal complexity of adaptive finite element eigenvalue computations. Numer. Math. **110**(3), 313–355 (2008)
12. Evans, L.C.: Partial Differential Equations, Second Ed., Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence (2010)
13. Ferraz-Leite, S., Ortner, C., Praetorius, D.: Convergence of simple adaptive Galerkin schemes based on $h - h/2$ error estimators. Numer. Math. **116**(2), 291–316 (2010)

14. Grubišić, L., Ovall, J.S.: On estimators for eigenvalue/eigenvector approximations. Math. Comp. **78**(266), 739–770 (2009)
15. Mehrmann, V., Miedlar, A.: Adaptive computation of smallest eigenvalues of self-adjoint elliptic partial differential equations. Numer. Linear Algebra Appl. **18**(3), 387–409 (2011)
16. Neymeyr, K.: A posteriori error estimation for elliptic eigenproblems. Numer. Linear Algebra Appl. **9**(4), 263–279 (2002)
17. Strang, G., Fix, G. J.: An Analysis of the Finite Element Method, Prentice-Hall Inc., Englewood Cliffs, N.J. (1973). Prentice-Hall Series in Automatic Computation
18. Verfürth, R.: A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Wiley and Teubner, Stuttgart (1996)