Optimal convergence of adaptive FEM for eigenvalue clusters in mixed form

D. Boffi*[1], D. Gallistl†[2], F. Gardini‡[1], and L. Gastaldi§[3]

[1]Dipartimento di Matematica "F. Casorati", University of Pavia, Italy
[2]Institut für Numerische Simulation, University of Bonn, Germany
[3]DICATAM, University of Brescia, Italy

September 24, 2018

Abstract

It is shown that the h-adaptive mixed finite element method for the discretization of eigenvalue clusters of the Laplace operator produces optimal convergence rates in terms of nonlinear approximation classes. The results are valid for the typical mixed spaces of Raviart–Thomas or Brezzi–Douglas–Marini type with arbitrary fixed polynomial degree in two and three space dimensions.

1 Introduction

The study of optimal convergence rates for adaptive finite element schemes has been carried on by several researchers during the last decades in the case of source problems (see, e.g., [22, 41, 17, 4, 16, 34]) and more recently has been applied to eigenvalue problems as well (see, e.g., [28, 32, 13] for convergence and [20, 14, 19, 12] for optimal rates). Some survey papers are available; we refer, in particular, for further reading and references, to [37, 38, 11]. In the case of eigenvalue approximation, it has been recently observed that adaptive schemes driven by the error indicator associated to an individual eigenvalue may produce unsatisfactory results, and that eigenvalues belonging to clusters have to be considered simultaneously (see, in particular, [25, 26, 27]).

In this paper, we study the adaptive approximation of the Laplace eigenvalue problem by mixed finite elements. The analysis of the underlying formulation, which fits the framework of $(0, g)$-type mixed problems, is not a mere generalization of the case of standard conforming Galerkin approach (see [6], where the convergence and the a priori estimates are recalled). This causes additional technical difficulties which were in previous works [24] circumvented by showing

*daniele.boffi@unipv.it
†gallistl@ins.uni-bonn.de
‡francesca.gardini@unipv.it
§lucia.gastaldi@unibs.it

equivalence with some nonconforming but elliptic finite element formulation. Typically, residual-based a posteriori error estimates are derived by exploiting the fact that the error of the eigenvalues as well as the error of the eigenfunctions in some weaker norm (usually the L^2 norm) is of higher-order compared with the error in the energy-like norm. The higher-order L^2 convergence, however, is not valid in its original format in mixed FEMs, and one technical tool we make use of is a fairly abstract superconvergence result for eigenvalue problems where a certain error quantity is shown to be of higher order in the L^2 norm. For the low-order case a similar result was shown in [29] by using the representation in terms of nonconforming finite elements from [24].

We follow the argument of [17] in order to show the optimality of an adaptive finite element scheme which is constructed taking into account clusters of eigenvalues in the spirit of [25]. In order to obtain the result, we need to derive estimates which are essentially different from the case of standard FEMs: this is one of the main contributions of our paper.

Previous a posteriori estimates for mixed formulation (source or eigenvalues problem) mostly showed efficiency and reliability with respect to the vector variable only (see [1] and [18, 34]; other results in this context can be found in [9, 44, 30, 36, 35]). Estimates involving the scalar variable were present in [24] (where, as already mentioned, the equivalence with nonconforming schemes is exploited) and in [10] (where the source problem is considered). Another main contribution of our analysis is that we show optimality also with respect to the scalar variable (see Definitions 6 and 7). This is performed by a suitable definition of the error indicator (see Definitions 5 and 9); this allows to prove the optimal convergence rate not only for the eigenfunction but for the eigenvalues as well (see Section 5).

The outline of the paper is as follows: Section 2 introduces the problem we are dealing with, Section 3 describes the error indicators and our adaptive algorithm, Section 4 states the main theorem of our paper, concerning the convergence of the adaptive scheme in terms of a *theoretical* error indicator which is equivalent to the error indicator used for the design of the AFEM algorithm. Section 5 shows that the convergence of the error indicator, which is related to the convergence of the eigenfunctions, actually implies the convergence of the eigenvalues as well. Finally, Section 6 contains all technical results which are used in the proof of our main theorem and Section 7 discusses the extension to three space dimensions.

Throughout this paper, we use standard notation for Lebesgue and Sobolev spaces and their norms. The L^2 norm of a function v over some domain ω is denoted by $\|v\|_\omega$ and, if there is no risk of confusion, we write $\|v\| = \|v\|_\Omega$ for the physical domain Ω. The scalar product of $L^2(\Omega)$ is denoted by $(\cdot,\cdot)$. If $\mathcal{A}$ is a disjoint union of subdomains of Ω, typically a (subset of a) triangulation, then $\|v\|^2_{\mathcal{A}} = \sum_{\omega\in\mathcal{A}} \|v\|^2_\omega$. We denote the scalar curl of some two-dimensional vector field ψ by $\operatorname{curl}\psi = \partial_2\psi_1 - \partial_1\psi_2$ and the vector curl of a scalar-valued function v by $\mathbf{curl}\, v = (-\partial_2 v, \partial_1 v)^T$. In three dimensions we define as usual $\mathbf{curl}\,\psi = \nabla\times\psi$.

The notation $A \lesssim B$ refers to an inequality $A \le CB$ up to a constant C that is independent of the mesh size. We do not trace the explicit dependence of the constants on the eigenvalues, cf. Remark 1.

The mesh-size is typically denoted by h; when a triangulation $\mathcal{T}_h$ is obtained as a refinement of a given mesh, we denote by $\mathcal{T}_H$ the coarser mesh. When dealing with the adaptive scheme, we denote by ℓ the level of refinement, so that $\mathcal{T}_{\ell+1}$ is the next triangulation in the algorithm obtained from $\mathcal{T}_\ell$.

2 Setting of the problem

Our main result is valid both in two and three dimensions. From now on, we discuss the two dimensional setting. Section 7 extends the result in three dimensions.

Given a polygonal domain Ω, in this paper we are interested in the following eigenvalue problem associated with the Laplace operator in mixed form: find $\lambda \in \mathbb{R}$ and $u \in L^2(\Omega)$ with $\|u\| = 1$ such that for some $\boldsymbol{\sigma} \in H(\mathrm{div};\Omega)$ it holds

$$\begin{cases} \displaystyle\int_\Omega \boldsymbol{\sigma}\cdot\boldsymbol{\tau}\,d\mathbf{x} + \int_\Omega u\,\mathrm{div}\,\boldsymbol{\tau}\,d\mathbf{x} = 0 & \forall\boldsymbol{\tau}\in H(\mathrm{div};\Omega)\\ \displaystyle\int_\Omega v\,\mathrm{div}\,\boldsymbol{\sigma}\,d\mathbf{x} = -\lambda\int_\Omega uv\,d\mathbf{x} & \forall v\in L^2(\Omega). \end{cases}$$

2.1 Abstract mixed eigenvalue problem

We cast this problem within the standard setting of abstract eigenvalue problems in mixed form of the second type (see [8, 6]).

Let Σ, M, $\mathcal{H}$ be Hilbert spaces such that $M \subseteq \mathcal{H} \subseteq M^\star$ and consider two bilinear and continuous forms $a : \Sigma\times\Sigma \to \mathbb{R}$ symmetric, and $b : \Sigma \times M \to \mathbb{R}$ which satisfy the usual hypotheses for mixed problems [7]: a is elliptic in the kernel of b and b fulfills the inf-sup condition. Moreover, the form a is supposed to be positive definite so that the associated norm $|\cdot|_a$ is well defined. In the pivot space $\mathcal{H}$ we consider the scalar product $(\cdot,\cdot)_\mathcal{H}$ and corresponding norm $\|\cdot\|_\mathcal{H}$.

In this framework, the continuous eigenvalue problem reads: find $\lambda \in \mathbb{R}$ and $u \in M$ with $\|u\|_\mathcal{H} = 1$ such that for some $\sigma \in \Sigma$ it holds

$$\begin{cases} a(\sigma,\tau) + b(\tau,u) = 0 & \forall\tau\in\Sigma\\ b(\sigma,v) = -\lambda(u,v)_\mathcal{H} & \forall v\in M \end{cases} \tag{2.1}$$

and, given finite dimensional subspaces $\Sigma_h \subset \Sigma$ and $M_h \subset M$ (typically associated to a finite element mesh $\mathcal{T}_h$), its discrete counterpart is: find $\lambda_h \in \mathbb{R}$ and $u_h \in M_h$ with $\|u_h\|_\mathcal{H} = 1$ such that for some $\sigma_h \in \Sigma_h$ it holds

$$\begin{cases} a(\sigma_h,\tau) + b(\tau,u_h) = 0 & \forall\tau\in\Sigma_h\\ b(\sigma_h,v) = -\lambda_h(u_h,v)_\mathcal{H} & \forall v\in M_h. \end{cases} \tag{2.2}$$

The following three assumptions ensure the good approximation of the eigenmodes (see [8, 6]), where $\rho(h)$ tends to zero as h goes to zero and Σ_0 and M_0 are the subspaces of Σ and M, respectively, containing all solutions to the source problem associated with (2.1) when the datum is in $\mathcal{H}$; the discrete kernel associated to the bilinear form b is as usual defined as

$$\mathbb{K}_h = \{\tau \in \Sigma_h : b(\tau, v) = 0\ \forall v \in M_h\}.$$

Fortid condition. There exists a Fortin operator $\Pi_h^F : \Sigma_0 \to \Sigma_h$ such that

$$b(\sigma - \Pi_h^F \sigma, v) = 0 \quad \forall v \in M_h$$

and

$$|\sigma - \Pi_{F,h}\sigma|_a \le \rho(h)\|\sigma\|_{\Sigma_0} \quad \forall \sigma \in \Sigma_0.$$

Weak approximability of M_0.

$$b(\tau_h, v) \le \rho(h)|\tau_h|_a\|v\|_{M_0} \quad \forall v \in M_0 \ \forall \tau_h \in \mathbb{K}_h.$$

Strong approximability of M_0.

$$\inf_{v_h \in M_h} \|v - v_h\|_{\mathcal{H}} \le \rho(h)\|v\|_{M_0} \quad \forall v \in M_0.$$

We consider a problem associated with a compact operator, so that the eigenvalues are enumerated as

$$0 < \lambda_1 \le \lambda_2 \le \lambda_3 \le \dots$$

(we repeat the eigenvalues according to their multiplicities); the corresponding eigenfunctions are denoted by $\{(\sigma_1, u_1), (\sigma_2, u_2), \dots\}$ and the $\{u_i\}$'s form an orthonormal system in $\mathcal{H}$. In particular, we have $|\sigma_i|_a^2 = \lambda_i$ and $\|u_i\|_{\mathcal{H}} = 1$ for $i = 1, 2, \dots$. We denote by $E(\lambda)$ the span of the $\{u_i\}$'s corresponding to λ.

Analogously, the discrete eigenvalues can be enumerated as follows

$$0 < \lambda_{h,1} \le \lambda_{h,2} \le \dots \le \lambda_{h,N(h)}$$

with corresponding eigenfunctions $\{(\sigma_{h,1}, u_{h,1}), \dots, (\sigma_{h,N(h)}, u_{h,N(h)})\}$, where $N(h) = \dim(M_h)$ and the $\{u_{h,i}\}$'s form an orthonormal system in $\mathcal{H}$. Here we have $|\sigma_{h,i}|_a^2 = \lambda_{h,i}$ and $\|u_{h,i}\|_{\mathcal{H}} = 1$ for $i = 1, 2, \dots, N(h)$.

For a cluster of eigenvalues $\lambda_{n+1}, \dots, \lambda_{n+\mathsf{N}}$ of length $\mathsf{N} \in \mathbb{N}$, we define the index set $J = \{n+1, \dots, n+\mathsf{N}\}$ and the spaces

$$W = \operatorname{span}\{u_j \mid j \in J\} \quad \text{and} \quad W_{\mathcal{T}_h} = W_h = \operatorname{span}\{u_{h,j} \mid j \in J\}.$$

2.2 Some useful operators

Definition 1. *For any $w \in M$ we define $\mathbf{G}(w) \in \Sigma$ as the solution to*

$$a(\mathbf{G}(w), \tau) + b(\tau, w) = 0 \quad \text{for all } \tau \in \Sigma. \tag{2.3}$$

For any $w_h \in M_h$ we define its discrete counterpart $\mathbf{G}_h(w_h) \in \Sigma_h$ via

$$a(\mathbf{G}_h(w_h), \tau_h) + b(\tau_h, w_h) = 0 \quad \text{for all } \tau_h \in \Sigma_h. \tag{2.4}$$

We explicitly notice that when two meshes $\mathcal{T}_h$ and $\mathcal{T}_H$ are present, it is important to distinguish between $\mathbf{G}_h$ and $\mathbf{G}_H$.

In many applications and corresponding instances of a and b, the above definition is related to an integration by parts formula where $\mathbf{G}(w)$ is some derivative of w. For instance, in the case of mixed Laplacian, $\mathbf{G}(w)$ is the gradient of w.

Definition 2. *The solution operators $T : \mathcal{H} \to M$ and $A : \mathcal{H} \to \Sigma$ are defined by*

$$\begin{cases} a(Ag, \tau) + b(\tau, Tg) = 0 & \forall \tau \in \Sigma \\ b(Ag, v) = -(g, v)_{\mathcal{H}} & \forall v \in M \end{cases} \tag{2.5}$$

and $T_h : \mathcal{H} \to M_h$ and $A_h : \mathcal{H} \to \Sigma_h$ are their discrete counterparts

$$\begin{cases} a(A_h g, \tau_h) + b(\tau_h, T_h g) = 0 & \forall \tau_h \in \Sigma_h \\ b(A_h g, v_h) = -(g, v_h)_{\mathcal{H}} & \forall v_h \in M_h. \end{cases} \tag{2.6}$$

Definition 3. *The operator $T_h^\lambda : \mathcal{H} \to M_h$ ($\lambda \in \mathbb{R}$) is defined by*

$$\begin{cases} a(\mathbf{G}_h(T_h^\lambda g), \tau_h) + b(\tau_h, T_h^\lambda g) = 0 & \forall \tau_h \in \Sigma_h \\ b(\mathbf{G}_h(T_h^\lambda g), v_h) = -(\lambda g, v_h)_{\mathcal{H}} & \forall v_h \in M_h, \end{cases} \tag{2.7}$$

that is, $T_h^\lambda = \lambda T_h$.

Let P_h^W denote the $\mathcal{H}$-orthogonal projection onto W_h. The following definition is crucial for the definition of our theoretical error indicator.

Definition 4. *The operator $\Lambda_h : E(\lambda) \to W_h$ is defined as follows:*

$$\Lambda_h = P_h^W \circ T_h^\lambda.$$

For the sake of simplicity, we do not include the dependence from λ in the notation for Λ_h: it will be clear from the context that when Λ_h is applied to an element of $E(\lambda)$, the corresponding value of λ should be used for its definition.

Lemma 2.1. *The operators P_h^W and T_h^λ commute, that is $\Lambda_h = P_h^W \circ T_h^\lambda = T_h^\lambda \circ P_h^W$. In other words, if (σ, u) is an eigenfunction associated with λ, then $\Lambda_h u$ solves*

$$\begin{cases} a(\mathbf{G}_h(\Lambda_h u), \tau_h) + b(\tau_h, \Lambda_h u) = 0 & \forall \tau_h \in \Sigma_h \\ b(\mathbf{G}_h(\Lambda_h u), v_h) = -(\lambda P_h^W u, v_h)_{\mathcal{H}} & \forall v_h \in M_h. \end{cases}$$

Proof. We adapt the result of [27, Lemma 2.2]. The expansion of $\Lambda_h u$ reads as $\Lambda_h u = \sum_{j \in J} (T_h^\lambda u, u_{h,j})_{\mathcal{H}} u_{h,j}$, thus $\Lambda_h u$ solves the discrete linear system (2.6) with right-hand side $g = \sum_{j \in J} (T_h^\lambda u, u_{h,j})_{\mathcal{H}} \lambda_{h,j} u_{h,j}$. For any $j \in J$ we have

$$\begin{aligned} \lambda_{h,j} (T_h^\lambda u, u_{h,j})_{\mathcal{H}} &= -b(\sigma_{h,j}, T_h^\lambda u) = a(\mathbf{G}_h(T_h^\lambda u), \sigma_{h,j}) = -b(\mathbf{G}_h(T_h^\lambda u), u_{h,j}) \\ &= \lambda (u, u_{h,j})_{\mathcal{H}}, \end{aligned}$$

which gives the final result that $\Lambda_h u$ solves the discrete linear system (2.6) with right-hand side $g = \sum_{j \in J} \lambda (u, u_{h,j})_{\mathcal{H}} u_{h,j} = \lambda P_h^W u$.

□

3 AFEM algorithm and error quantities

As already mentioned, we are interested in the Laplace eigenvalue problem in mixed form with Dirichlet boundary conditions. Namely, with the notation introduced in Section 2, we are making the following choices:

$$\Sigma = H(\mathrm{div};\Omega)$$
$$M = \mathcal{H} = L^2(\Omega)$$
$$a(\sigma,\tau) = (\sigma,\tau)$$
$$b(\tau,v) = (\operatorname{div}\tau, v)$$

for an open, bounded, simply-connected polygonal Lipschitz domain Ω.

It follows, in particular that the seminorm $|\cdot|_a$ is the norm in $(L^2(\Omega))^2$. Our analysis applies to more general operators (for instance, Neumann boundary conditions or non-constant coefficients), but we stick to this simpler example for the sake of readability.

We discretize the problem with standard mixed finite elements (including Raviart–Thomas, Brezzi–Douglas–Marini, etc.), see [7] for more detail. It is well-known that this choice satisfies the assumptions discussed in Section 2 (see, for instance, [8]).

Moreover, we observe that the following relation (part of the commuting diagram) holds true:

$$\operatorname{div}(\Sigma_h) = M_h \tag{3.1}$$

Let us first introduce our error indicator.

Definition 5. *Let $\mathcal{T}_h$ be a triangulation of Ω and let $(\sigma_{h,j}, u_{h,j}) \in \Sigma_h \times M_h$ be a discrete eigensolution computed on the mesh $\mathcal{T}_h$. Then, for all $T \in \mathcal{T}_h$ we define*

$$\eta_{h,j}(T)^2 = \|h_T(\sigma_{h,j} - \nabla u_{h,j})\|_T^2 + \|h_T \operatorname{curl}\sigma_{h,j}\|_T^2 + \sum_{E\in\mathcal{E}(T)} h_E \|[\sigma_{h,j}]_E \cdot t_E\|_E^2,$$

where h_T is the diameter of T, $\mathcal{E}(T)$ denotes the set of edges of T, h_E is the length of the edge E, and t_E is its unit tangent vector. As usual, $[\sigma_h]_E \cdot t_E$ denotes the jump of the trace of $\sigma_h \cdot t_E$ for internal edges and the trace for boundary edges.

Given a set $\mathcal{M}$ of elements of $\mathcal{T}_h$, we define

$$\eta_{h,j}(\mathcal{M})^2 = \sum_{T\in\mathcal{M}} \eta_{h,j}(T)^2.$$

3.1 Adaptive algorithm

The adaptive algorithm consists of the standard four steps: solve, estimate, mark, and refine. In the description of the fours steps, we describe how the algorithms runs from level ℓ to $\ell+1$.

Solve. Given a mesh $\mathcal{T}_\ell$ the algorithm computes the eigensolutions of (2.2) belonging to the cluster $(\lambda_{\ell,j}, \sigma_{\ell,j}, u_{\ell,j})$ for $j \in J$. We assume that the discrete solution is computed exactly.

Estimate. The algorithm computes the local contributions of the error estimator for the eigenfunctions in the cluster $\{\eta_{\ell,j}(T)\}_{T\in\mathcal{T}_\ell}$ $(j \in J)$.

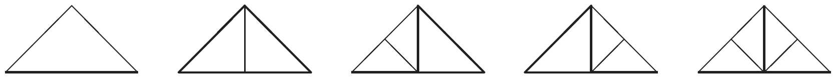

Figure 3.1: Possible refinements of a triangle T in one level in 2D. The thick lines indicate the refinement edges of the sub-triangles as in [5, 42].

Mark. The algorithm uses the well known Dörfler marking strategy [22]. Given a bulk parameter $\theta \in (0, 1]$, a minimal subset $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ is identified such that

$$\theta \sum_{j \in J} \eta_{\ell,j}(\mathcal{T}_\ell)^2 \leq \sum_{j \in J} \eta_{\ell,j}(\mathcal{M}_\ell)^2.$$

The elements belonging to $\mathcal{M}_\ell$ are marked for refinement.

Refine. A new triangulation $\mathcal{T}_{\ell+1}$ is generated, as the smallest admissible refinement of $\mathcal{T}_\ell$ satisfying $\mathcal{M}_\ell \cap \mathcal{T}_{\ell+1} = \emptyset$ by using the refinement rules of [5, 42]. Figure 3.1 shows possible refinements of a triangle.

To summarize, the adaptive algorithm accepts as **input** the bulk parameter θ and the initial mesh $\mathcal{T}_0$ (with proper initialization of refinement edges as in [5, 42]), and returns as **output** a sequence of meshes $\{\mathcal{T}_\ell\}$ and of discrete eigenpairs $\{(\lambda_{\ell,j}, \sigma_{\ell,j}, u_{\ell,j})\}_{j \in J}$.

Finally, we shall make use of the following notation: given an initial mesh $\mathcal{T}_0$, regular in the sense of Ciarlet, we denote by $\mathbb{T}$ the set of *admissible meshes* in the sense that a mesh in $\mathbb{T}$ is a refinement of $\mathcal{T}_0$ obtained using the rules of [5, 42].

3.2 Error quantities and theoretical error indicator

The following definition introduces a metric in M.

Definition 6. *$d : M \times M \to \mathbb{R}$ is defined as*

$$d(v, w) = \sqrt{\|v - w\|^2 + |\mathbf{G}(v) - \mathbf{G}(w)|_a^2}$$

When v (resp. w) belongs to M_h, then $\mathbf{G}_h(v)$ (resp. $\mathbf{G}_h(w)$) should be used.

Remark 1. *We note that it may be useful to balance the terms in the square root of Definition 6 in terms of λ. In particular, if v and w are related to eigenfunctions with frequency λ, the right scaling would involve multiplying by λ the term $\|v - w\|$. This is of particular interest it one aims to quantify the conditions on the initial mesh-size. In this paper, we do not aim at such a quantification and refer the interested reader to [27] for such a λ explicit analysis in the context of conforming standard finite elements.*

This distance allows us to evaluate the gap between discrete and continuous eigenfunctions in the cluster.

Definition 7. *The following quantity measures how combinations of eigenfunctions in the cluster W are approximated by their discrete counterparts in W_h.*

$$\delta(W, W_h) = \sup_{\substack{u \in W \\ \|u\|=1}} \inf_{v_h \in W_h} d(u, v_h).$$

Given a refinement $\mathcal{T}_\ell \in \mathbb{T}$ of the initial mesh $\mathcal{T}_0$, our theory is based on the introduction of the following *non-computable* error indicator μ_ℓ which will be proved equivalent to the *computable* indicator η_ℓ.

Definition 8. *Let $\mathcal{T}_h \in \mathbb{T}$ be a triangulation and for all $T \in \mathcal{T}_h$ and $g_h \in M_h$ let us consider the following seminorm*

$$|g_h|_{\eta,T}^2 = \|h_T(\mathbf{G}_h(g_h) - \nabla g_h)\|_T^2 + \|h_T \operatorname{curl} \mathbf{G}_h(g_h)\|_T^2 + \sum_{E \in \mathcal{E}(T)} h_E \|[\mathbf{G}_h(g_h)]_E \cdot t_E\|_E^2,$$

so that

$$\eta_{h,j}(T) = |u_{h,j}|_{\eta,T}.$$

Then, given an eigenfunction (σ, u) associated to the eigenvalue λ (in particular, this is used in the definition of Λ_h), we define

$$\mu_h(u;T) = |\Lambda_h u|_{\eta,T}.$$

Given a set $\mathcal{M}$ of elements of $\mathcal{T}_h$, we define

$$\mu_h(u;\mathcal{M})^2 = \sum_{T \in \mathcal{M}} \mu_h(u;T)^2.$$

The next lemma is of technical nature and gives a criterion for linear independence. It generalizes [13, Prop. 3.2].

Lemma 3.1. *Recall the notation $\mathsf{N} = \operatorname{card}(J)$ and suppose that*

$$\varepsilon = \max_{j \in J} \|u_j - \Lambda_h u_j\| \le \sqrt{1 + 1/(2\mathsf{N})} - 1. \tag{3.2}$$

Then, $\{\Lambda_h u_j\}_{j\in J}$ forms a basis of W_h. For any $w_h \in W_h$ with $\|w_h\| = 1$, the coefficients of the representation $w_h = \sum_{j\in J} \gamma_j \Lambda_h u_j$ are controlled as

$$\sum_{j\in J} |\gamma_j|^2 \le 2 + 4\mathsf{N}. \tag{3.3}$$

Proof. The proof employs Gershgorin's theorem. Since the proof follows verbatim the lines of [27, Lemma 5.1], it is omitted here. □

The following lemma states the equivalence between the two introduced estimators. It is clear that the adaptive algorithm will make use of the computable indicator η, while the indicator μ will be used for the analysis.

Lemma 3.2 (Local comparison of the error estimators)**.** *Provided the initial mesh-size is small enough such that (3.2) is satisfied, it holds for any $T \in \mathcal{T}_h$ that*

$$\mathsf{N}^{-1} \sum_{j\in J} \mu_h(u_j;T)^2 \le \left(\frac{\mathsf{B}}{\mathsf{A}}\right)^2 \sum_{j\in J} \eta_{h,j}(T)^2 \le \left(\frac{\mathsf{B}}{\mathsf{A}}\right)^2 (2\mathsf{N} + 4\mathsf{N}^2) \sum_{j\in J} \mu_h(u_j;T)^2$$

where $[\mathsf{A},\mathsf{B}]$ denotes a real interval containing the (continuous and discrete) eigenvalue cluster and N is the number of eigenvalues in the cluster.

Proof. The proof follows from a perturbation analysis as in [27, Prop. 5.1]. We include the proof for self-contained reading. Let $k \in J$ and consider the expansion of $\Lambda_h u_k = \sum_{j\in J} \gamma_j u_{h,j}$ with coefficients $\gamma_j = (\Lambda_h u_k, u_{h,j})$. The definition of Λ_h and the symmetry yield

$$\begin{aligned}\gamma_j &= (\Lambda_h u_k, u_{h,j}) = (T_h^\lambda u_k, u_{h,j}) = -\lambda_{h,j}^{-1} b(\sigma_{h,j}, T_h^\lambda u_k)\\ &= \lambda_{h,j}^{-1} a(\sigma_{h,j}, \mathsf{G}_h(T_h^\lambda u_k)) = -\lambda_{h,j}^{-1} b(\mathsf{G}_h(T_h^\lambda u_k), u_{h,j}) = \lambda_{h,j}^{-1}\lambda_k (u_k, u_{h,j}).\end{aligned}$$

Since $\{u_{h,j}\}_{j\in J}$ is an orthonormal system, we arrive at $\sum_{j\in J}\gamma_j^2 \le (\mathsf{B}/\mathsf{A})^2$, which implies

$$|\Lambda_h u_k|_{\eta,T}^2 \le \left(\sum_{j\in J}\gamma_j^2\right)\sum_{j\in J}|u_{h,j}|_{\eta,T}^2 \le \left(\frac{\mathsf{B}}{\mathsf{A}}\right)^2\sum_{j\in J}|u_{h,j}|_{\eta,T}^2.$$

This proves the first stated inequality.

Lemma 3.1 shows that there exist real coefficients $\{\delta_j \mid j \in J\}$ such that

$$u_{h,k} = \sum_{j\in J}\delta_j \Lambda_h u_j \quad \text{and} \quad \sum_{j\in J}\delta_j^2 \le 2 + 4\mathsf{N}.$$

The triangle and Cauchy inequalities lead to

$$|u_{h,k}|_{\eta,T}^2 \le \left(\sum_{j\in J}\delta_j^2\right)\sum_{j\in J}|\Lambda_h u_j|_{\eta,T}^2 \le (2+4\mathsf{N})\sum_{j\in J}|\Lambda_h u_j|_{\eta,T}^2.$$

This shows the second stated inequality and concludes the proof. ☐

4 Optimal convergence of the adaptive scheme

In this section we state the main theorem showing the optimal convergence of our adaptive scheme and sketch the principal lines of its proof. The structure of the proof is closely related to [17] and relies on several intermediate results which, for the sake of readability, will be postponed to Section 6.

As usual in this context, the convergence is measured by introducing a suitable nonlinear approximation class in the spirit of [5]. For any $m \in \mathbb{N}$, we denote by

$$\mathbb{T}(m) = \{\mathcal{T} \in \mathbb{T} \mid \mathrm{card}(\mathcal{T}) - \mathrm{card}(\mathcal{T}_0) \le m\}$$

the set of admissible triangulations in $\mathbb{T}$ whose cardinality differs from that of $\mathcal{T}_0$ by m or less.

The best algebraic convergence rate $s \in (0, +\infty)$ obtained by any admissible mesh in $\mathbb{T}$ is characterized in terms of the following seminorm

$$|W|_{\mathcal{A}_s} = \sup_{m\in\mathbb{N}} m^s \inf_{\mathcal{T}\in\mathbb{T}(m)} \delta(W, W_{\mathcal{T}}).$$

In particular, we have $|W|_{\mathcal{A}_s} < \infty$ if the rate of convergence $\delta(W, W_{\mathcal{T}}) = O(m^{-s})$ holds true for the optimal triangulations $\mathcal{T}$ in $\mathbb{T}(m)$.

The main results of this section, stated in Theorem 4.1, shows that the same optimal rate of convergence is reached by the error quantity $\delta(W, W_{\mathcal{T}_\ell})$ associated with the mesh sequence $\{\mathcal{T}_\ell\}$ obtained from the adaptive algorithm presented in Section 3.

Theorem 4.1. *Provided the initial mesh-size and the bulk parameter θ are small enough, if for the eigenvalue cluster W it holds $|W|_{\mathcal{A}_s} < \infty$, then the sequence of discrete clusters W_ℓ computed on the mesh $\mathcal{T}_\ell$ satisfies the optimal estimate*

$$\delta(W, W_\ell)(\mathrm{card}(\mathcal{T}_\ell) - \mathrm{card}(\mathcal{T}_0))^s \lesssim |W|_{\mathcal{A}_s}.$$

Proof. We follow the lines of the proof of Theorem 3.1 in [27]. The main arguments are the same as in [17].

Given a positive β, we consider the quantity

$$\xi_\ell^2 = \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2 + \beta \sum_{j\in J} d(u_j, \Lambda_\ell u_j)^2$$

which will be used in the contraction argument of Proposition 6.11. We do not consider the trivial case $\xi_0 = 0$. Choose $0 < \tau \le |W|^2_{\mathcal{A}_s}/\xi_0^2$, and set $\varepsilon(\ell) = \sqrt{\tau}\,\xi_\ell$. Let $N(\ell) \in \mathbb{N}$ be minimal with the property

$$|W|^2_{\mathcal{A}_s} \le \varepsilon(\ell)^2\, N(\ell)^{2s}.$$

It can be easily seen that $N(\ell) > 1$, otherwise

$$|W|_{\mathcal{A}_s} \le \varepsilon(\ell)$$

but this, together with the definition of $\varepsilon(\ell)$, would violate the contraction property of Proposition 6.11.

From the minimality of $N(\ell)$ it turns out that

$$N(\ell) \le 2|W|^{1/s}_{\mathcal{A}_s} \varepsilon(\ell)^{-1/s} \quad \text{for all } \ell \in \mathbb{N}_0. \tag{4.1}$$

Let $\widetilde{\mathcal{T}}_\ell \in \mathbb{T}$ denote the optimal triangulation of cardinality

$$\mathrm{card}(\widetilde{\mathcal{T}}_\ell) \le \mathrm{card}(\mathcal{T}_0) + N(\ell)$$

in the sense that the operator $\widetilde{\Lambda} = \Lambda_{\widetilde{\mathcal{T}}_\ell}$ of Definition 4 with respect to the mesh $\widetilde{\mathcal{T}}_\ell$ satisfies

$$\sum_{j\in J} d(u_j, \widetilde{\Lambda} u_j)^2 \le N(\ell)^{-2s} |W|^2_{\mathcal{A}_s} \le \varepsilon(\ell)^2. \tag{4.2}$$

Let us consider the overlay $\widehat{\mathcal{T}}_\ell$, that is the smallest common refinement of $\mathcal{T}_\ell$ and $\widetilde{\mathcal{T}}_\ell$, which is known [17] to satisfy

$$\mathrm{card}(\mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell) \le \mathrm{card}(\widehat{\mathcal{T}}_\ell) - \mathrm{card}(\mathcal{T}_\ell) \le \mathrm{card}(\widetilde{\mathcal{T}}_\ell) - \mathrm{card}(\mathcal{T}_0) \le N(\ell). \tag{4.3}$$

This relation and (4.1)–(4.3) lead to

$$\mathrm{card}(\mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell) \le N(\ell) \le 2|W|^{1/s}_{\mathcal{A}_s} \varepsilon(\ell)^{-1/s}. \tag{4.4}$$

Let $\widehat{\Lambda}$ denote the operator $\Lambda_{\widehat{\mathcal{T}}_\ell}$ with respect to the mesh $\widehat{\mathcal{T}}_\ell$.

The following estimate

$$\sum_{j\in J} d(u_j, \widehat{\Lambda} u_j)^2 \le 3\varepsilon(\ell)^2 \tag{4.5}$$

follows from the quasi-orthogonality (see Proposition 6.9) applied to $\mathcal{T}_h = \widehat{\mathcal{T}}_\ell$ and $\mathcal{T}_H = \widetilde{\mathcal{T}}_\ell$. Indeed

$$(1 - C_{\mathrm{qo}}\rho(h_0)) \sum_{j\in J} d(u_j, \widehat{\Lambda} u_j)^2 \le (1 + C_{\mathrm{qo}}\rho(h_0)) \sum_{j\in J} d(u_j, \widetilde{\Lambda} u_j)^2.$$

Estimate (4.5) follows from the mesh-size condition $C_{\mathrm{qo}}\rho(h_0) \le 1/2$ and (4.2).

We now show the existence of a constant C_1 such that

$$\sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2 \le C_1 \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell)^2. \tag{4.6}$$

From the triangle inequality and the discrete reliability (see Proposition 6.7) we obtain for any $j \in J$

$$\begin{aligned} d(u_j, \Lambda_\ell u_j)^2 &\le 2d(u_j, \widehat{\Lambda}_\ell u_j)^2 + 2d(\widehat{\Lambda}_\ell u_j, \Lambda_\ell u_j)^2 \\ &\le 2d(u_j, \widehat{\Lambda}_\ell u_j)^2 + 2C_{\mathrm{drel}}^2 \mu_\ell(\mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell)^2 \\ &\quad + C\rho(h_0)^2 (d(u_j, \Lambda_\ell u_j) + d(u_j, \widehat{\Lambda}_\ell u_j))^2. \end{aligned}$$

Provided the initial mesh-size is sufficiently small, this leads to some constant C_2 such that with (4.5) it follows

$$\sum_{j\in J} d(u_j, \Lambda_\ell u_j)^2 \le C_2 \varepsilon(\ell)^2 + C_2 C_{\mathrm{drel}}^2 \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell)^2.$$

Let C_{eq} denote the constant of $C_2 \xi_\ell^2 \le C_{\mathrm{eq}} \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2$ (which exists by reliability). The efficiency (6.2), the definition of $\varepsilon(\ell)$, and the preceding estimates prove

$$\begin{aligned} C_{\mathrm{eff}}^{-2} \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2 &\le C_2 \varepsilon(\ell)^2 + C_2 C_{\mathrm{drel}}^2 \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell)^2 \\ &\le \tau C_{\mathrm{eq}} \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2 + C_2 C_{\mathrm{drel}}^2 \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell)^2. \end{aligned}$$

Defining $C_1 = (C_{\mathrm{eff}}^{-2} - \tau C_{\mathrm{eq}})^{-1} C_2 C_{\mathrm{drel}}^2$, which is positive for a sufficiently small choice of τ, we obtain (4.6).

In order to conclude the proof, we now make the following choice for the parameter θ:

$$0 < \theta \le 1 \big/ \big(C_1 (\mathsf{B}/\mathsf{A})^2 (2\mathsf{N}^2 + 4\mathsf{N}^3)\big)\ .$$

The marking step in the adaptive algorithm selects $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ with minimal cardinality such that

$$\theta \sum_{j\in J} \eta_{\ell,j}(\mathcal{T}_\ell)^2 \le \sum_{j\in J} \eta_{\ell,j}(\mathcal{M}_\ell)^2.$$

Estimate (4.6) and the definition of θ imply together with Lemma 3.2 that also $\mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell$ satisfies the bulk criterion, that is

$$\theta \sum_{j\in J} \eta_{\ell,j}(\mathcal{T}_\ell)^2 \le \sum_{j\in J} \eta_{\ell,j}(\mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell)^2.$$

The minimality of $\mathcal{M}_\ell$ and (4.4) show that

$$\operatorname{card}(\mathcal{M}_\ell) \leq \operatorname{card}(\mathcal{T}_\ell \setminus \widehat{\mathcal{T}}_\ell) \leq 2|W|_{\mathcal{A}_s}^{1/s} \tau^{-1/(2s)} \xi_\ell^{-1/s}. \tag{4.7}$$

It is proved in [5, 42] that there exists a constant C_{BDV} such that

$$\begin{aligned}\operatorname{card}(\mathcal{T}_\ell) - \operatorname{card}(\mathcal{T}_0) &\leq C_{\mathrm{BDV}} \sum_{k=0}^{\ell-1} \operatorname{card}(\mathcal{M}_k) \\ &\leq 2C_{\mathrm{BDV}} |W|_{\mathcal{A}_s}^{1/s} \tau^{-1/(2s)} \sum_{k=0}^{\ell-1} \xi_k^{-1/s}.\end{aligned}$$

The contraction property from Proposition 6.11 implies $\xi_\ell^2 \leq \rho_2^{\ell-k} \xi_k^2$ for $k = 0, \ldots, \ell$. Since $\rho_2 < 1$, a geometric series argument leads to

$$\sum_{k=0}^{\ell-1} \xi_k^{-1/s} \leq \xi_\ell^{-1/s} \sum_{k=0}^{\ell-1} \rho_2^{(\ell-k)/(2s)} \leq \xi_\ell^{-1/s} \rho_2^{1/(2s)} \Big/ \left(1 - \rho_2^{1/(2s)}\right).$$

The combination of the above estimates results in

$$\begin{aligned}&\operatorname{card}(\mathcal{T}_\ell) - \operatorname{card}(\mathcal{T}_0) \\ &\qquad \leq 2C_{\mathrm{BDV}} |W|_{\mathcal{A}_s}^{1/s} \tau^{-1/(2s)} \xi_\ell^{-1/s} \rho_2^{1/(2s)} \Big/ \left(1 - \rho_2^{1/(2s)}\right).\end{aligned}$$

The equivalence of ξ_ℓ^2 with the error $\sum_{j \in J} d(u_j, \Lambda_\ell u_j)^2$ (reliability and efficiency, see Section 6) concludes the proof.

□

5 Convergence of eigenvalues

The previous analysis shows that the adaptive procedure leads to the convergence of the quantity $\delta(W, W_\ell)$ which is related to the eigenfunctions belonging to the cluster. In this section we show that this estimate actually implies the optimal convergence of the eigenvalues.

The next discussion has been inspired by [21]. However, we do not make use explicitly of the spectral projections and follow a somehow more natural argument (at least for symmetric problems).

As usual, we consider the eigenvalues $\mu_i = 1/\lambda_i$ $(i = 1, \ldots)$ of T and $\mu_{\ell,i} = 1/\lambda_{\ell,i}$ $(i = 1, \ldots, \dim(M_\ell))$ of T_ℓ and discuss the convergence of $\mu_{\ell,j}$ to μ_j for $j \in J$. This standard notation conflicts with our theoretical error indicator; nevertheless, we believe that this overlap is not a source of confusion, since it is limited to this section where the error indicator is not mentioned.

Let $E : \mathcal{H} \to \mathcal{H}$ denote the $\mathcal{H}$ projection onto W and $E_\ell : \mathcal{H} \to \mathcal{H}$ the $\mathcal{H}$ projection onto W_ℓ. We denote by F_ℓ the restriction of E_ℓ to W

$$F_\ell = E_\ell|_W.$$

The following proposition shows that for ℓ large enough the operator F_ℓ is a bijection from W to W_ℓ (which have the same dimension N).

Proposition 5.1. *For ℓ large enough the operator F_ℓ is injective. Moreover, F_ℓ^{-1} is uniformly bounded in $\mathcal{L}(W_\ell, W)$ and*

$$\sup_{\substack{x\in W_\ell\\ \|x\|_{\mathcal{H}}=1}} \|F_\ell^{-1}x - x\|_{\mathcal{H}} \le C\delta(W, W_\ell).$$

Proof. It is enough to show that for ℓ sufficiently large $\|F_\ell y - y\|_{\mathcal{H}} \le (1/2)\|y\|_{\mathcal{H}}$ for all $y \in W$ (see also [21, Lemma 2]). Indeed, from the definition of F_ℓ it is immediate to get

$$\|F_\ell y - y\|_{\mathcal{H}} \le \|y - y_\ell\|_{\mathcal{H}} \quad \forall y_\ell \in W_\ell$$

which implies

$$\|F_\ell y - y\|_{\mathcal{H}} \le \delta(W, W_\ell)\|y\|_{\mathcal{H}}.$$

We can then conclude our proof from Theorem 4.1 observing that $\delta(W, W_\ell)$ tends to zero.

□

Let us define the following operators from W into itself:

$$\hat{T} = T|_W, \quad \hat{T}_\ell = F_\ell^{-1}T_\ell F_\ell.$$

It is clear that the eigenvalues of $\hat{T}$ ($\hat{T}_\ell$, resp.) are equal to μ_j ($\mu_{\ell,j}$ resp.), $j \in J$.

Lemma 5.2. *The following estimates hold true for all $x \in W$*

$$\begin{aligned} \|(T - T_\ell)x\|_{\mathcal{H}} &\le C\delta(W, W_\ell),\\ |(A - A_\ell)x|_a &\le C\delta(W, W_\ell),\\ \|(A - A_\ell)x\|_\Sigma &\le C\delta(W, W_\ell). \end{aligned} \tag{5.1}$$

Proof. Let us denote $u = Tx$, $u_\ell = T_\ell x$, $\sigma = \mathbf{G}(u) = Ax$, and $\sigma_\ell = \mathbf{G}_\ell(u_\ell) = A_\ell x$.

In order to prove the first estimate, we use a standard duality argument and introduce the following auxiliary problem: find $\zeta \in \Sigma$ and $w \in M$ such that

$$\begin{cases} a(\zeta, \tau) + b(\tau, w) = 0 & \forall \tau \in \Sigma\\ b(\zeta, v) = -(u - u_\ell, v)_{\mathcal{H}} & \forall v \in M. \end{cases}$$

We clearly have $\|\zeta\|_\Sigma + \|w\|_M \le C\|u - u_\ell\|_{\mathcal{H}}$. By standard arguments we get

$$\begin{aligned} \|u - u_\ell\|^2_{\mathcal{H}} &= (u - u_\ell, u - u_\ell)_{\mathcal{H}} = -b(\zeta, u - u_\ell)\\ &= -b(\zeta - \Pi_{F,\ell}\zeta, u) - b(\Pi_{F,\ell}\zeta, u - u_\ell)\\ &= a(\sigma, \zeta - \Pi_{F,\ell}\zeta) + a(\mathbf{G}(u) - \mathbf{G}_\ell(u_\ell), \Pi_{F,\ell}\zeta). \end{aligned} \tag{5.2}$$

For all $v_\ell \in M_\ell$, the first term can be estimated as follows:

$$\begin{aligned} |a(\sigma, \zeta - \Pi_{F,\ell}\zeta)| &= |a(\sigma - \mathbf{G}_\ell(v_\ell), \zeta - \Pi_{F,\ell}\zeta) + a(\mathbf{G}_\ell(v_\ell), \zeta - \Pi_{F,\ell}\zeta)|\\ &= |a(\sigma - \mathbf{G}_\ell(v_\ell), \zeta - \Pi_{F,\ell}\zeta) - b(\zeta - \Pi_{F,\ell}\zeta, v_\ell)|\\ &= |a(\sigma - \mathbf{G}_\ell(v_\ell), \zeta - \Pi_{F,\ell}\zeta)|\\ &\le C|\sigma - \mathbf{G}_\ell(v_\ell)|_a\|\zeta\|_\Sigma \le C|\sigma - \mathbf{G}_\ell(v_\ell)|_a\|u - u_\ell\|_{\mathcal{H}}. \end{aligned}$$

The second term in the last line of (5.2) can be estimated as follows:

$$|a(\mathbf{G}(u) - \mathbf{G}_\ell(u_\ell), \Pi_{F,\ell}\zeta)| \le C|\mathbf{G}(u) - \mathbf{G}_\ell(u_\ell)|_a \|\zeta\|_\Sigma \le C|(A - A_\ell)x|_a \|u - u_\ell\|_{\mathcal{H}}.$$

Hence

$$\|Tx - T_\ell x\|_{\mathcal{H}} \le C\left(|\sigma - \mathbf{G}_\ell(v_\ell)|_a + |(A - A_\ell)x|_a\right) \quad \forall v_\ell \in M_\ell.$$

Since the first term is bounded by $\delta(W, W_\ell)$, the final estimate will follow from the second estimate in (5.1).

Let us prove the second estimate in (5.1).

From the definition of W we have

$$x = \sum_{j \in J} \alpha_j u_j,$$

where we recall that $(\lambda_j, \sigma_j, u_j)$ is the generic eigensolution belonging to the cluster W and the coefficients are given by $\alpha_j = (x, u_j)$.

Hence, $Ax = \mathbf{G}(u)$ with $u = Tx$ and

$$Ax = \sum_{j \in J} \frac{1}{\lambda_j} \alpha_j \sigma_j.$$

Analogously, from (2.7),

$$A_\ell x = \sum_{j \in J} \frac{1}{\lambda_j} \alpha_j \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j).$$

We then obtain

$$|Ax - A_\ell x|_a = \left| \sum_{j \in J} \frac{1}{\lambda_j} \alpha_j (\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j)) \right|_a.$$

We now show that $|\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j)|_a$ can be bounded by $\delta(W, W_\ell)$. For all $v_\ell \in M_\ell$ we have

$$\begin{aligned}
|\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j)|_a^2 &= a(\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j), \sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j)) \\
&= a(\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j), \sigma_j - \mathbf{G}_\ell(v_\ell)) \\
&\quad + a(\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j), \mathbf{G}_\ell(v_\ell) - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j)).
\end{aligned}$$

Since the last term is vanishing for the properties of σ_j and the definitions of $T_\ell^{\lambda_j}$ and $\mathbf{G}_\ell$, we obtain

$$|\sigma_j - \mathbf{G}_\ell(T_\ell^{\lambda_j} u_j)|_a \le \inf_{v_\ell \in M_\ell} |\sigma_j - \mathbf{G}_\ell(v_\ell)|_a \le C\delta(W, W_\ell).$$

From [7, Prop. 4.3.4] and the definitions of A and A_ℓ it follows that

$$\|(A - A_\ell)x\|_\Sigma \le C|(A - A_\ell)x|_a + C\|x - x_\ell\|_{\mathcal{H}},$$

where $x_\ell \in M_\ell$ is the $\mathcal{H}$ projection of x. The first term is readily bounded by $\delta(W, W_\ell)$, while the second one is smaller than $\|x - F_\ell x\|_{\mathcal{H}}$ which has been already estimated in the proof of Proposition 5.1.

□

The following proposition is a crucial step for the bound of the eigenvalues.

Proposition 5.3. *The following estimate holds true*

$$\|\hat{T} - \hat{T}_\ell\|_{\mathcal{L}(W)} \leq C\delta(W, W_\ell)^2 \tag{5.3}$$

Proof. Let us define $\mathcal{S}_\ell = F_\ell^{-1}E_\ell - I : \mathcal{H} \to \mathcal{H}$. From the boundedness of the involved operators, it is immediate to observe that $\mathcal{S}_\ell$ is uniformly bounded.

For all $x \in W$ we have

$$(\hat{T} - \hat{T}_\ell)x = (T - T_\ell)x + \mathcal{S}_\ell(T - T_\ell)x \tag{5.4}$$

since $E_\ell \mathcal{S}_\ell = 0$. Let us estimate the first term. For all $x, y \in W$ with $\|x\|_\mathcal{H} = \|y\|_\mathcal{H} = 1$

$$\begin{aligned}((T - T_\ell)x, y)_\mathcal{H} &= -b(Ay, (T - T_\ell)x) + a((A - A_\ell)x, A_\ell y) + b(A_\ell y, (T - T_\ell)x)\\ &= -b((A - A_\ell)y, (T - T_\ell)x) + a((A - A_\ell)x, A_\ell y).\end{aligned}$$

The first term is bounded by a constant times $\delta(W, W_\ell)^2$, while the second term can be estimated as follows.

$$\begin{aligned}a((A - A_\ell)x, A_\ell y) &= a((A - A_\ell)x, (A_\ell - A)y) + a((A - A_\ell)x, Ay)\\ &= a((A - A_\ell)x, (A_\ell - A)y) - b((A - A_\ell)x, Ty)\\ &= a((A - A_\ell)x, (A_\ell - A)y) - b((A - A_\ell)x, (T - T_\ell)y)\\ &\leq C\delta(W, W_\ell)^2.\end{aligned}$$

The second term in (5.4) can be estimated using the following identity

$$(\mathcal{S}_\ell(T - T_\ell)x, y)_\mathcal{H} = (\mathcal{S}_\ell(T - T_\ell)x, y - F_\ell y)_\mathcal{H}$$

which finally leads to

$$|(\mathcal{S}_\ell(T - T_\ell)x, y - E_\ell y)_\mathcal{H}| \leq \|\mathcal{S}_\ell\|_{\mathcal{L}(\mathcal{H})}\|T - T_\ell\|_{\mathcal{L}(\mathcal{H})}\|I - F_\ell\|_{\mathcal{L}(\mathcal{H})}.$$

□

The operators $\hat{T}$ and $\hat{T}_\ell$ can be represented by symmetric positive definite matrices of dimension $\mathsf{N} \times \mathsf{N}$ (N being the dimension of W). The following theorem is then a standard consequence of matrix perturbation theory (see, for instance, [21, Theorem 3, items c) and d)]) and to the equivalences $\lambda_i = 1/\mu_i$ and $\lambda_{\ell,i} = 1/\mu_{\ell,i}$.

Theorem 5.4. *Let J denote the set of indices corresponding to the eigenvalues in the cluster W. Then*

$$\sup_{i \in J} \inf_{j \in J} |\lambda_i - \lambda_{\ell,j}| \leq C\delta(W, W_\ell)^2.$$

6 Auxiliary results

This section contains all technical results which have been used for the proof of Theorem 4.1. We arrange the presentation in three subsections: in the first one a superconvergence result is proved; in the second one we collect the results which hold for all refinements $\mathcal{T}_h$ of a given mesh $\mathcal{T}_H$; finally, in the third one we include the results which have been proved for the sequence of meshes $\{\mathcal{T}_\ell\}$ generated by our adaptive procedure.

6.1 A superconvergence result and other useful estimates

Let Π_h denote the orthogonal projection onto M_h.

Lemma 6.1 (Superconvergence for the source problem)**.** *There exist $\rho(h)$ tending to zero as h goes to zero such that*

$$\|\Pi_h u - T_h^\lambda u\| \lesssim \rho(h)\|\sigma - \mathsf{G}_h(T_h^\lambda u)\|_\Sigma.$$

Proof. This result has been proved in [23] and can be found in [29] or [7, §7.4] as well. □

Let $J^C = \{1, \ldots, N(h)\} \setminus J$ denote the indices of the discrete eigenvalues not belonging to the cluster and assume the initial mesh-size is small enough such that

$$K = \sup_{\mathcal{T}_h} \sup_{k\in J^C} \sup_{j\in J} \frac{\lambda_j}{|\lambda_j - \lambda_{h,k}|} < \infty.$$

Lemma 6.2. *For all $j \in J^C$ we have*

$$(u_{h,j}, T_h^\lambda u) = \frac{\lambda}{\lambda - \lambda_{h,j}}(T_h^\lambda u - \Pi_h u, u_{h,j}).$$

Proof. We have

$$\begin{aligned}-\lambda_{h,j}(u_{h,j}, T_h^\lambda u) &= b(\sigma_{h,j}, T_h^\lambda u) = -a(\sigma_{h,j}, \mathsf{G}_h(T_h^\lambda u)) = b(\mathsf{G}_h(T_h^\lambda u), u_{h,j})\\ &= -\lambda(u, u_{h,j}) = -\lambda(\Pi_h u, u_{h,j}).\end{aligned}$$

Adding $\lambda(u_{h,j}, T_h^\lambda u)$ on both sides of this identity leads to

$$(\lambda - \lambda_{h,j})(u_{h,j}, T_h^\lambda u) = \lambda(T_h^\lambda u - \Pi_h u, u_{h,j}).$$

□

Lemma 6.3 (cf. [43])**.** *Any eigensolution $(\lambda, \sigma, u) \in \mathbb{R} \times \Sigma \times W$ in the cluster satisfies*

$$\|T_h^\lambda u - \Lambda_h u\| \leq K\|\Pi_h u - T_h^\lambda u\|.$$

Proof. Let us define $e_h = T_h^\lambda u - \Lambda_h u$. The expansion in terms of the orthonormal basis $\{u_{h,j} \mid j = 1, \ldots, N(h)\}$ reads

$$e_h = \sum_{j\in J^C} \alpha_j u_{h,j} \quad \text{with} \sum_{j\in J^C} \alpha_j^2 = \|e_h\|^2.$$

This relation, Lemma 6.2, and Parceval's identity lead to

$$\begin{aligned}\|e_h\|^2 &= \sum_{j\in J^C} \alpha_j (T_h^\lambda u, u_{h,j}) = \sum_{j\in J^C} \alpha_j \frac{\lambda}{\lambda - \lambda_{h,j}}(T_h^\lambda u - \Pi_h u, u_{h,j})\\ &\leq K\Big(\sum_{j\in J^C} \alpha_j^2\Big)^{1/2} \|T_h^\lambda u - \Pi_h u\|.\end{aligned}$$

□

We are now ready to prove the superconvergence result for the eigenvalue problem.

Proposition 6.4 (Superconvergence for the eigenvalue problem)**.** *Any eigensolution* $(\lambda, \sigma, u) \in \mathbb{R} \times \Sigma \times W$ *in the cluster satisfies*

$$\|\Pi_h u - \Lambda_h u\| \lesssim \rho(h)\|\sigma - \mathbf{G}_h(T_h^\lambda u)\|_\Sigma.$$

Proof. The triangle inequality and Lemma 6.3 give

$$\|\Pi_h u - \Lambda_h u\| \leq \|\Pi_h u - T_h^\lambda u\| + \|T_h^\lambda u - \Lambda_h u\| \leq (1+K)\|\Pi_h u - T_h^\lambda u\|.$$

The right-hand side has been estimated in Lemma 6.1. □

The following result contains a useful bound of the norm of the error in Σ in terms of our error quantity.

Lemma 6.5 (Bound for the Σ norm)**.** *Any eigensolution* $(\lambda, \sigma, u) \in \mathbb{R} \times \Sigma \times M$ *satisfies*

$$\|\sigma - \mathbf{G}_h(\Lambda_h u)\|_\Sigma \lesssim |\sigma - \mathbf{G}_h(\Lambda_h u)|_a + (1+\lambda)\|u - \Lambda_h u\|. \tag{6.1}$$

Proof. The stability of the continuous problem implies

$$\begin{aligned}\|\sigma - \mathbf{G}_h(\Lambda_h u)\|_\Sigma \\ \lesssim \sup_{\substack{(\tau,v)\in\Sigma\times M\\ \|\tau\|_\Sigma+\|v\|=1}} \big(a(\sigma - \mathbf{G}_h(\Lambda_h u), \tau) + b(\sigma - \mathbf{G}_h(\Lambda_h u), v) + b(\tau, u - \Lambda_h u)\big).\end{aligned}$$

The identity (3.1) together with the continuous and discrete eigenvalue problems imply

$$\begin{aligned}b(\sigma - \mathbf{G}_h(\Lambda_h u), v) = b(\sigma, v) - b(\mathbf{G}_h(\Lambda_h u), \Pi_h v) &= \lambda\big((P_h^W u, \Pi_h v) - (u, v)\big)\\ &= \lambda(P_h^W u - u, v).\end{aligned}$$

Estimate (6.1) then follows from the continuity of a and b together with the elementary estimate $\|u - P_h^W u\| \leq \|u - \Lambda_h u\|$.

□

6.2 Properties valid for all refinements $\mathcal{T}_h$ of $\mathcal{T}_H$

We start this section by proving the efficiency of our theoretical error estimator on the generic mesh $\mathcal{T}_h$.

Proposition 6.6 (Efficiency)**.** *Let* (σ, u) *be an eigenpair associated to the eigenvalue* λ*, then there exists a positive constant* C_{eff}*, independent of* h*, such that*

$$\mu_h(u; \mathcal{T}_h) \leq C_{\mathrm{eff}}\, d(u, \Lambda_h u). \tag{6.2}$$

Proof. For the reader's convenience, we recall the definition of the error indicator $\mu_h(u;T)$ for a given element $T \in \mathcal{T}_h$:

$$\begin{aligned}\mu_h(u;T)^2 = \|h_T(\mathbf{G}_h(\Lambda_h u) - \nabla\Lambda_h u)\|_T^2 + \|h_T \operatorname{curl} \mathbf{G}_h(\Lambda_h u)\|_T^2\\ + \sum_{E\in\mathcal{E}(T)} h_E\|[\mathbf{G}_h(\Lambda_h u)]_E \cdot t_E\|_E^2.\end{aligned} \tag{6.3}$$

Following the same arguments as in [10, Lemma 6.3], we can prove that

$$h_T^2 \|G_h(\Lambda_h u) - \nabla \Lambda_h u\|_T^2 \lesssim d(u, \Lambda_h u)^2. \tag{6.4}$$

Finally, arguing as in the proof of Theorem 3.1 in [1], we can bound the remaining terms of the error indicator as follows:

$$\|h_T \operatorname{curl} \mathbf{G}_h(\Lambda_h u)\|_T^2 + \sum_{E \in \mathcal{E}(T)} h_E \|[\mathbf{G}_h(\Lambda_h u)]_E \cdot t_E\|_E^2 \lesssim \|\sigma - G_h(\Lambda_h u)\|_{\tilde{T}}^2, \tag{6.5}$$

where $\tilde{T}$ denotes the union of T and its neighboring elements.
We then obtain the desired result by summing equations (6.4) and (6.5) over each elements $T \in \mathcal{T}_h$.

□

The next result shows a uniform discrete reliability of the theoretical error estimator when evaluated on the mesh $\mathcal{T}_h$, refinement of $\mathcal{T}_H$.

First of all, we recall the well-known discrete Helmholtz decomposition which is valid for the finite element spaces we are considering. Suitable references for this result are [2] in the framework of discrete exterior calculus or [33]. In our setting the discrete Helmholtz decomposition reads (see [34, Lemma 2.5]): for any $\zeta_h \in \Sigma_h$ there exist $\alpha_h \in M_h$ and $\beta_h \in \mathcal{P}_{k+1}(\mathcal{T}_h)$ (the space of continuous piecewise polynomial of degree $k+1$) such that

$$\zeta_h = \mathbf{G}_h(\alpha_h) + \mathbf{curl}\, \beta_h. \tag{6.6}$$

In particular, $\alpha_h \in M_h$ is such that

$$\begin{aligned} a(\mathbf{G}_h(\alpha_h), \tau_h) + b(\tau_h, \alpha_h) &= 0 \quad && \forall \tau_h \in \Sigma_h \\ b(\mathbf{G}_h(\alpha_h), v_h) &= b(\zeta_h, v_h) && \forall v_h \in M_h. \end{aligned} \tag{6.7}$$

By definition of the bilinear form and the fact that $\operatorname{div} \Sigma_h = M_h$, we have that $\operatorname{div}(\mathbf{G}_h(\alpha_h) - \zeta_h) = 0$, hence $\mathbf{G}_h(\alpha_h) - \zeta_h = \mathbf{curl}\, \beta_h$. Using again (3.1) there exists $\hat{\tau}_h \in \Sigma_h$ such that $\operatorname{div} \hat{\tau}_h = \alpha_h$. From the discrete inf-sup condition we have $\|\hat{\tau}_h\| \le C\|\alpha_h\|$. Hence

$$\|\alpha_h\|^2 = b(\hat{\tau}_h, \alpha_h) = a(\mathbf{G}_h(\alpha_h), \hat{\tau}_h) \le |\mathbf{G}_h(\alpha_h)|_a \|\hat{\tau}_h\| \le C|\mathbf{G}_h(\alpha_h)|_a \|\alpha_h\|,$$

from which we obtain

$$\|\alpha_h\| \le C|\mathbf{G}_h(\alpha_h)|_a. \tag{6.8}$$

Proposition 6.7 (Discrete reliability)**.** *Provided the mesh-size of $\mathcal{T}_H$ is sufficiently small, we have*

$$\begin{aligned} &|\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u)|_a + \|\Lambda_h u - \Lambda_H u\| \\ &\qquad \le C_{\mathrm{drel}} \mu_H(u; \mathcal{T}_H \setminus \mathcal{T}_h) + C\rho(H)(d(u, \Lambda_h u) + d(u, \Lambda_H u)). \end{aligned}$$

Proof. From the discrete Helmholtz decomposition (6.6) there exist $\alpha_h \in M_h$ and $\beta_h \in \mathcal{P}_{k+1}(\mathcal{T}_h)$ such that

$$\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u) = \mathbf{G}_h(\alpha_h) + \mathbf{curl}\, \beta_h.$$

The term $\|\operatorname{\mathbf{curl}}\beta_h\|$ can be bounded by using standard arguments as in [24, 3, 34]. Actually, taking $\overline{\beta}_H$ as the Scott-Zhang interpolant [40] of β_h on the mesh $\mathcal{T}_H$,

$$\begin{aligned}|\operatorname{\mathbf{curl}}\beta_h|_a^2 &= a(\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u), \operatorname{\mathbf{curl}}\beta_h)\\ &= -a(\mathbf{G}_H(\Lambda_H u), \operatorname{\mathbf{curl}}(\beta_h - \overline{\beta}_H))\\ &= \sum_{T\in\mathcal{T}_H\setminus\mathcal{T}_h}\Big(\int_T(\beta_h - \overline{\beta}_H)\operatorname{curl}\mathbf{G}_H(\Lambda_H u)\,dx\\ &\qquad - \int_{\partial T}(\beta_h - \overline{\beta}_H)\mathbf{G}_H(\Lambda_H u)\cdot t\,ds\Big).\end{aligned}$$

Standard estimates for the Scott-Zhang interpolant give

$$|\operatorname{\mathbf{curl}}\beta_h|_a \lesssim \mu_H(u;\mathcal{T}_H\setminus\mathcal{T}_h).$$

The integration by parts and some straightforward algebraic manipulations lead to

$$\begin{aligned}|\mathbf{G}_h(\alpha_h)|_a^2 &= a(\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u), \mathbf{G}_h(\alpha_h))\\ &= \lambda(P_h^W u - P_H^W u, \alpha_h)\\ &= \lambda\big((P_h^W u - \Pi_h u, \alpha_h) + (\Pi_h u - \Pi_H u, \alpha_h - \Pi_H\alpha_h)\\ &\quad + (\Pi_H u - P_H^W u, \alpha_h)\big).\end{aligned}$$

We observe that $\|P_h^W u - \Pi_h u\| \le \|\Lambda_h u - \Pi_h u\|$; indeed, $P_h^W u$ is the best H-approximation of u into W_h and is characterized by $(P_h^W u - u, v_h) = (P_h^W u - \Pi_h u, v_h) = 0$ for all $v_h \in W_h$. Hence, the estimate (6.8), Proposition 6.4, and Lemma 6.5 prove for the first and the last term that

$$\begin{aligned}&(P_h^W u - \Pi_h u, \alpha_h) + (\Pi_H u - P_H^W u, \alpha_h)\\ &\quad\lesssim \big(\|P_h^W u - \Pi_h u\| + \|\Pi_H u - P_H^W u\|\big)|\mathbf{G}_h(\alpha_h)|_a\\ &\quad\lesssim \rho(H)(d(u,\Lambda_h u) + d(u,\Lambda_H u))|\mathbf{G}_h(\alpha_h)|_a.\end{aligned}$$

For the analysis of the remaining term, set $\xi = \alpha_h - \Pi_H\alpha_h$. It is shown in [34, Lemma 2.8 and Equation (3.9)] that ξ satisfies $\|\xi\| \lesssim H|\mathbf{G}_h(\alpha_h)|_a$. Thus, we have with Proposition 6.4 that

$$\begin{aligned}(\Pi_h u - \Pi_H u, \xi) &= (\Pi_h u - \Lambda_H u, \xi)\\ &= (\Pi_h u - \Lambda_h u, \xi) + (\Lambda_h u - \Lambda_H u, \xi)\\ &\lesssim (\rho(H)d(u,\Lambda_h u) + H\|\Lambda_h u - \Lambda_H u\|)|\mathbf{G}_h(\alpha_h)|_a.\end{aligned}$$

Altogether we obtain for the error in the vector variable that

$$\begin{aligned}&|\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u)|_a\\ &\quad\lesssim \mu_H(u;\mathcal{T}_H\setminus\mathcal{T}_h) + \rho(H)(d(u,\Lambda_h u) + d(u,\Lambda_H u)) + H\|\Lambda_h u - \Lambda_H u\|.\end{aligned}$$

It remains to estimate the error in the scalar variable.

Let $\hat{z}$ be the gradient of the solution $\hat{\phi}$ of the problem

$$\begin{aligned}\Delta\hat{\phi} &= \Lambda_h u - \Lambda_H u &&\text{in } \Omega\\ \hat{\phi} &= 0 &&\text{on } \partial\Omega.\end{aligned}$$

Using a (non-orthogonal) stable decomposition like the ones adopted in [31, Lemma 3.3] or [39, Lemma 2.1], it is possible to find $z \in H^1(\Omega)$ such that

$$\hat{z} = z + \mathbf{curl}\, \psi.$$

In particular we have

$$\begin{aligned} &\operatorname{div} z = \Lambda_h u - \Lambda_H u \\ &\|z\| + \|\nabla z\| \lesssim \|\Lambda_h u - \Lambda_H u\|. \end{aligned}$$

It follows

$$\begin{aligned} \|\Lambda_h u - \Lambda_H u\|^2 &= b(z, \Lambda_h u - \Lambda_H u) \\ &= b(\Pi_h^F z, \Lambda_h u) - b(\Pi_H^F z, \Lambda_H u) \\ &= -a(\mathbf{G}_h(\Lambda_h u), \Pi_h^F z) + a(\mathbf{G}_H(\Lambda_H u), \Pi_H^F z) \\ &= a(\mathbf{G}_H(\Lambda_H u) - \mathbf{G}_h(\Lambda_h u), \Pi_h^F z) + a(\mathbf{G}_H(\Lambda_H u), (\Pi_H^F - \Pi_h^F) z) \\ &\le |\mathbf{G}_H(\Lambda_H u) - \mathbf{G}_h(\Lambda_h u)|_a \|\Pi_h^F z\| \\ &\quad + a(\mathbf{G}_H(\Lambda_H u) - \nabla_H(\Lambda_H u), (\Pi_H^F - \Pi_h^F) z), \end{aligned} \tag{6.9}$$

where we have used the definition of the Fortin operators Π_h^F, Π_H^F, of $\mathbf{G}_h$ and $\mathbf{G}_H$, and, in the last term, the fact that the quantity $a(\nabla_H(\Lambda_H u), (\Pi_H^F - \Pi_h^F) z)$ vanishes.

We observe furthermore that $\Pi_h^F z - \Pi_H^F z = 0$ on the unrefined elements $\mathcal{T}_H \cap \mathcal{T}_h$. Since z is smooth enough to allow for stability and first-order approximation of Π_h^F and Π_H^F, we conclude

$$\begin{aligned} \|\Lambda_h u - \Lambda_H u\|^2 &\le |\mathbf{G}_H(\Lambda_H u) - \mathbf{G}_h(\Lambda_h u)|_a \|\Pi_h^F z\| \\ &\quad + \|H(\mathbf{G}_H(\Lambda_H u) - \nabla_H(\Lambda_H u))\|_{\mathcal{T}_H \setminus \mathcal{T}_h} \|H^{-1}(\Pi_h^F z - \Pi_H^F z)\| \\ &\lesssim \|\Lambda_h u - \Lambda_H u\| \\ &\quad (|\mathbf{G}_H(\Lambda_H u) - \mathbf{G}_h(\Lambda_h u)|_a + \mu_H(\Lambda_H u; \mathcal{T}_H \setminus \mathcal{T}_h)). \end{aligned}$$

□

By passing to the limit in the statement of Proposition 6.7, and observing that for H small enough the second term on the right-hand side can be absorbed, we obtain the standard reliability estimate.

Corollary 6.8 (Reliability)**.** *Provided the initial mesh-size is sufficiently fine, we have*

$$\sum_{j \in J} d(u_j, \Lambda_h u_j)^2 \le C_{\text{rel}}^2 \sum_{j \in J} \mu_h(u_j, \mathcal{T}_h)^2.$$

We conclude this section with a quasi-orthogonality property.

Proposition 6.9 (Quasi-orthogonality)**.** *There exists a constant C_{qo} such that*

$$d(\Lambda_h u, \Lambda_H u)^2 \le d(u, \Lambda_H u)^2 - d(u, \Lambda_h u)^2 + C_{\text{qo}}\, \rho(h)(d(u, \Lambda_h u)^2 + d(u, \Lambda_H u)^2).$$

Proof. The proof departs with the following obvious algebraic identities

$$\begin{aligned} |\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u)|_a^2 &= |\sigma - \mathbf{G}_H(\Lambda_H u)|_a^2 - |\sigma - \mathbf{G}_h(\Lambda_h u)|_a^2 \\ &\quad - 2a(\sigma - \mathbf{G}_h(\Lambda_h u), \mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u)) \end{aligned}$$

$$\|\Lambda_h u - \Lambda_H u\|^2 = \|u - \Lambda_H u\|^2 - \|u - \Lambda_h u\|^2 - 2(\Pi_h u - \Lambda_h u, \Lambda_h u - \Lambda_H u).$$

The exact and discrete eigenvalue problems together with the inclusion $\operatorname{div} \Sigma_H \subseteq M_H$ imply

$$\begin{aligned} a(\sigma - \mathbf{G}_h(\Lambda_h u), \mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u)) &= -b(\mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u), u - \Lambda_h u) \\ &= \lambda(P_h^W u - P_H^W u, \Pi_h u - \Lambda_h u). \end{aligned}$$

Therefore it follows from Proposition 6.4, Lemma 6.5, and the Young inequality that

$$\begin{aligned} &|a(\sigma - \mathbf{G}_h(\Lambda_h u), \mathbf{G}_h(\Lambda_h u) - \mathbf{G}_H(\Lambda_H u))| + |(\Pi_h u - \Lambda_h u, \Lambda_h u - \Lambda_H u)| \\ &\quad \leq \|\Pi_h u - \Lambda_h u\| \left(\|\Lambda_h u - \Lambda_H u\| + \lambda \|P_h^W u - P_H^W u\|\right) \\ &\quad \lesssim \rho(h)(d(u, \Lambda_h u)^2 + d(u, \Lambda_H u)^2). \end{aligned}$$

□

6.3 Contraction property

While the properties of the previous subsection are valid for any refinement $\mathcal{T}_h$ of a mesh $\mathcal{T}_H$, in this section we deal with the mesh sequence $\mathcal{T}_\ell$ which is the output of the adaptive strategy described in Section 3.

The following property is quite standard and can be proved with the techniques of [17].

Lemma 6.10 (Error estimator reduction property). *Provided the initial mesh-size is sufficiently small such that the bulk criteria for μ_ℓ and η_ℓ are equivalent (see Lemma 3.2), there exist constants $\rho_1 \in (0,1)$ and $K \in (0,+\infty)$ such that $\mathcal{T}_\ell$ and its one-level refinement $\mathcal{T}_{\ell+1}$ generated by AFEM satisfy*

$$\sum_{j\in J} \mu_{\ell+1}(u_j, \mathcal{T}_{\ell+1})^2 \leq \rho_1 \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2 + K_1 \sum_{j\in J} d(\Lambda_{\ell+1} u_j, \Lambda_\ell u_j)^2.$$

The following proposition presents the main contraction property which is essential for the convergence proof of the adaptive strategy.

Proposition 6.11 (Contraction property). *Provided the initial mesh-size is sufficiently small, there exist $\rho_2 \in (0,1)$ and $\beta \in (0,+\infty)$ such that the term*

$$\xi_\ell^2 = \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2 + \beta \sum_{j\in J} d(u_j, \Lambda_\ell u_j)^2 \tag{6.10}$$

satisfies

$$\xi_{\ell+1}^2 \leq \rho_2 \xi_\ell^2 \quad \text{for all } \ell \in \mathbb{N}.$$

Proof. Throughout the proof, we use the following notation

$$e_\ell^2 = \sum_{j\in J} d(u_j, \Lambda_\ell u_j)^2 \qquad \mu_\ell^2 = \sum_{j\in J} \mu_\ell(u_j, \mathcal{T}_\ell)^2.$$

The error estimator reduction from Lemma 6.10 and the quasi-orthogonality from Lemma 6.9 imply the following bound

$$\mu_{\ell+1}^2 + K_1(1 - C_{\mathrm{qo}}\rho(h_0)) e_{\ell+1}^2 \leq \rho_1 \mu_\ell^2 + K_1(1 + C_{\mathrm{qo}}\rho(h_0)) e_\ell^2.$$

For any $\varepsilon \in (0,1)$, the last bound and the reliability (Corollary 6.8) give

$$\mu_{\ell+1}^2 + K_1(1 - C_{\mathrm{qo}}\rho(h_0))e_{\ell+1}^2 \\ \leq (\rho_1 + \varepsilon C_{\mathrm{rel}}^2 K_1)\mu_\ell^2 + K_1(1 - \varepsilon + C_{\mathrm{qo}}\rho(h_0))e_\ell^2.$$

We take $\beta = K_1(1 - C_{\mathrm{qo}}\rho(h_0))$ and

$$\rho_2 = \max\left\{\rho_1 + \varepsilon C_{\mathrm{rel}}^2 K_1, \frac{1 - \varepsilon + C_{\mathrm{qo}}\rho(h_0)}{1 - C_{\mathrm{qo}}\rho(h_0)}\right\},$$

so that

$$\mu_{\ell+1}^2 + \beta e_{\ell+1}^2 \leq \rho_2(\mu_\ell^2 + \beta e_\ell^2).$$

The choice of a sufficiently small ε and of a sufficiently small initial mesh-size h_0 leads to $\rho_2 < 1$. □

7 Extension to three space dimensions

The results presented in the previous sections hold true also in three dimensions, provided the definitions of the computable and theoretical error indicators are modified as follows.

Definition 9. *Let $\mathcal{T}_h$ be a simplicial decomposition of Ω and let $(\sigma_{h,j}, u_{h,j}) \in \Sigma_h \times M_h$ be a discrete eigensolution computed on the mesh $\mathcal{T}_h$. Then, for all $T \in \mathcal{T}_h$ we define*

$$\eta_{h,j}(T)^2 = \|h_T(\sigma_{h,j} - \nabla u_{h,j})\|_T^2 + \|h_T \,\mathbf{curl}\, \sigma_{h,j}\|_T^2 + \sum_{F \in \mathcal{F}(T)} h_F \|[\sigma_{h,j}]_F \times n_F\|_F^2,$$

where h_T is the diameter of T, $\mathcal{F}(T)$ denotes the set of faces of T, h_F is the area of the face F, and n_F is its unit normal vector. As usual, $[\sigma_h]_F \times n_F$ denotes the jump of the trace of $\sigma_h \times n_F$ for internal faces and the trace for boundary faces.

Definition 10. *Let $\mathcal{T}_h \in \mathbb{T}$ be a triangulation and let (σ, u) be an eigensolution associated to the eigenvalue λ (in particular, this is used in the definition of Λ_h). For all $T \in \mathcal{T}_h$ we define*

$$\mu_h^2(u;T) = \|h_T(\mathbf{G}(\Lambda_h u) - \nabla \Lambda_h u)\|_T^2 + \|h_T \,\mathbf{curl}\, \mathbf{G}(\Lambda_h u)\|_T^2 \\ + \sum_{F \in \mathcal{F}(T)} h_F \|[\mathbf{G}(\Lambda_h u)]_F \times n_F\|_F^2.$$

In the three-dimensional case, the only proof which needs to be modified is the one of the discrete reliability of Proposition 6.7 since it relies on the discrete Helmholtz decomposition which is different in two or three dimensions.

Let V_h denote the $H(\mathbf{curl})$-conforming edge elements of Nédélec (see [7]).

Then, in the three dimensional case, the discrete Helmholtz decomposition reads (see [34], Lemma 2.6): for any $\xi_h \in \Sigma_h$ there exist $\alpha_h \in M_h$ and $\beta_h \in V_h$ such that

$$\xi_h = \mathbf{G}_h(\alpha_h) + \mathbf{curl}\, \beta_h.$$

Proposition 7.1 (Discrete reliability)**.** *Provided the mesh-size of $\mathcal{T}_H$ is sufficiently small, we have*

$$|\mathsf{G}_h(\Lambda_h u) - \mathsf{G}_H(\Lambda_H u)|_a + \|\Lambda_h(u) - \Lambda_H(u)\| \\ \leq C_{\mathrm{drel}}\mu_H(u;\tilde{\mathcal{R}}) + C\rho(H)(d(u,\Lambda_h u) + d(u,\Lambda_H u)),$$

where $\tilde{\mathcal{R}} = \{T \in \mathcal{T}_H : \bar{T} \cap \bar{T}' \neq \emptyset \quad \forall T' \in (\mathcal{T}_H \setminus \mathcal{T}_h)\}$.

Proof. Using the discrete Helmholtz decomposition, we write the error in the vectorial variable as

$$\mathsf{G}_h(\Lambda_h u) - \mathsf{G}_H(\Lambda_H u) = \mathsf{G}_h(\alpha_h) + \mathbf{curl}\,\beta_h,$$

with $\alpha_h \in M_h$ and $\beta_h \in V_h$.

The term $|\mathsf{G}_h(\alpha_h)|_a$ can be treated without any modification as in the two dimensional case. Moreover, following the same argument as in [34, Lemma 3.1.], it can be proved that

$$|\,\mathbf{curl}\,\beta_h|_a \lesssim \mu_H(u;\tilde{\mathcal{R}}).$$

As in the proof of Proposition 6.7, the error in the scalar variable can be bounded by using the duality argument of [31, 15] and we can repeat the same arguments of the 2D case from Equation (6.9) onwards, concluding the proof. □

Remark 2. *Compared with the two-dimensional case, in the three-dimensional version of the discrete reliability, the set $\mathcal{T}_H \setminus \mathcal{T}_h$ is replaced with the slighliy larger set $\tilde{\mathcal{R}}$ which essentially is $\mathcal{T}_H \setminus \mathcal{T}_h$ plus one additional layer of simplices. The shape-regularity implies that there is a constant C such that*

$$\mathrm{card}(\tilde{\mathcal{R}}) \leq C\,\mathrm{card}(\mathcal{T}_H \setminus \mathcal{T}_h).$$

and therefore the estimate (4.7) remains valid at the expense of the multiplicative constant C, and with this modification the proof of Theorem 4.1 applies also to the three-dimensional case.

Acknowledgements

The second named author gratefully acknowledges the hospitality of the Dipartimento di Matematica “F. Casorati” (University of Pavia) during his stay in September 2014.

References

[1] A. Alonso. Error estimators for a mixed method. *Numer. Math.*, 74(4):385–395, 1996.

[2] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta Numer.*, 15:1–155, 2006.

[3] R. Becker and S. Mao. An optimally convergent adaptive mixed finite element method. *Numer. Math.*, 111(1):35–54, 2008.

[4] R. Becker, S. Mao, and Z. Shi. A convergent nonconforming adaptive finite element method with quasi-optimal complexity. *SIAM J. Numer. Anal.*, 47(6):4639–4659, 2010.

[5] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numer. Math.*, 97(2):219–268, 2004.

[6] D. Boffi. Finite element approximation of eigenvalue problems. *Acta Numer.*, 19:1–120, 2010.

[7] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2013.

[8] D. Boffi, F. Brezzi, and L. Gastaldi. On the convergence of eigenvalues for mixed formulations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 25(1-2):131–154 (1998), 1997. Dedicated to Ennio De Giorgi.

[9] D. Braess and R. Verfürth. A posteriori error estimators for the Raviart-Thomas element. *SIAM J. Numer. Anal.*, 33(6):2431–2444, 1996.

[10] C. Carstensen. A posteriori error estimate for the mixed finite element method. *Math. Comp.*, 66(218):465–476, 1997.

[11] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[12] C. Carstensen, D. Gallistl, and M. Schedensack. Adaptive nonconforming Crouzeix-Raviart FEM for eigenvalue problems. *Math. Comp.*, 84(293):1061–1087, 2015.

[13] C. Carstensen and J. Gedicke. An oscillation-free adaptive FEM for symmetric eigenvalue problems. *Numer. Math.*, 118(3):401–427, 2011.

[14] C. Carstensen and J. Gedicke. An adaptive finite element eigenvalue solver of quasi-optimal computational complexity. *SIAM J. Numer. Anal.*, 50(3):1029–1057, 2012.

[15] C. Carstensen, D. Peterseim, and A. Schröder. The norm of a discretized gradient in $H(div)^*$ for a posteriori finite element error analysis. *Numer. Math.*, 2015. To appear.

[16] C. Carstensen and H. Rabus. An optimal adaptive mixed finite element method. *Math. Comp.*, 80(274):649–667, 2011.

[17] J. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[18] L. Chen, M. Holst, and J. Xu. Convergence and optimality of adaptive mixed finite element methods. *Math. Comp.*, 78(265):35–53, 2009.

[19] X. Dai, L. He, and A. Zhou. Convergence and quasi-optimal complexity of adaptive finite element computations for multiple eigenvalues. *IMA J. Numer. Anal.*, 2014. To appear, DOI 10.1093/imanum/dru059.

[20] X. Dai, J. Xu, and A. Zhou. Convergence and optimal complexity of adaptive finite element eigenvalue computations. *Numer. Math.*, 110(3):313–355, 2008.

[21] J. Descloux, N. Nassif, and J. Rappaz. On spectral approximation. II. Error estimates for the Galerkin method. *RAIRO Anal. Numér.*, 12(2):113–119, iii, 1978.

[22] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, 33(3):1106–1124, 1996.

[23] J. Douglas, Jr. and J. E. Roberts. Mixed finite element methods for second order elliptic problems. *Mat. Apl. Comput.*, 1(1):91–103, 1982.

[24] R. G. Durán, L. Gastaldi, and C. Padra. A posteriori error estimators for mixed approximations of eigenvalue problems. *Math. Models Methods Appl. Sci.*, 9(8):1165–1178, 1999.

[25] D. Gallistl. *Adaptive Finite Element Computation of Eigenvalues.* Doctoral dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Mathematik, 2014.

[26] D. Gallistl. Adaptive nonconforming finite element approximation of eigenvalue clusters. *Comput. Methods Appl. Math.*, 14(4):509–535, 2014.

[27] D. Gallistl. An optimal adaptive FEM for eigenvalue clusters. *Numer. Math.*, 2014. Published online, doi: 10.1007/s00211-014-0671-8.

[28] E. M. Garau, P. Morin, and C. Zuppa. Convergence of adaptive finite element methods for eigenvalue problems. *Math. Models Methods Appl. Sci.*, 19(5):721–747, 2009.

[29] F. Gardini. Mixed approximation of eigenvalue problems: a superconvergence result. *M2AN Math. Model. Numer. Anal.*, 43(5):853–865, 2009.

[30] G. N. Gatica and M. Maischak. A posteriori error estimates for the mixed finite element method with Lagrange multipliers. *Numer. Methods Partial Differential Equations*, 21(3):421–450, 2005.

[31] G. N. Gatica, R. Oyarzúa, and F.-J. Sayas. A residual-based a posteriori error estimator for a fully-mixed formulation of the Stokes-Darcy coupled problem. *Comput. Methods Appl. Mech. Engrg.*, 200(21-22):1877–1891, 2011.

[32] S. Giani and I. G. Graham. A convergent adaptive method for elliptic eigenvalue problems. *SIAM J. Numer. Anal.*, 47(2):1067–1091, 2009.

[33] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1986. Theory and algorithms.

[34] J. Huang and Y. Xu. Convergence and complexity of arbitrary order adaptive mixed element methods for the Poisson equation. *Sci. China Math.*, 55(5):1083–1098, 2012.

[35] M. G. Larson and A. Målqvist. A posteriori error estimates for mixed finite element approximations of elliptic problems. *Numer. Math.*, 108(3):487–500, 2008.

[36] C. Lovadina and R. Stenberg. Energy norm a posteriori error estimates for mixed finite element methods. *Math. Comp.*, 75(256):1659–1674 (electronic), 2006.

[37] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, Berlin, 2009.

[38] R. H. Nochetto and A. Veeser. Primer of adaptive finite element methods. In *Multiscale and adaptivity: modeling, numerics and applications*, volume 2040 of *Lecture Notes in Math.*, pages 125–225. Springer, Heidelberg, 2012.

[39] J. E. Pasciak and J. Zhao. Overlapping Schwarz methods in $H(\text{curl})$ on polyhedral domains. *J. Numer. Math.*, 10(3):221–234, 2002.

[40] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[41] R. Stevenson. Optimality of a standard adaptive finite element method. *Foundations of Computational Mathematics*, 7(2):245–269, 2007.

[42] R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.

[43] G. Strang and G. J. Fix. *An analysis of the finite element method*. Prentice-Hall Series in Automatic Computation. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1973.

[44] B. I. Wohlmuth and R. H. W. Hoppe. A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements. *Math. Comp.*, 68(228):1347–1378, 1999.