

A CONVERGENT ADAPTIVE METHOD FOR ELLIPTIC EIGENVALUE PROBLEMS

S. GIANI* AND I.G. GRAHAM†

Abstract. We prove the convergence of an adaptive linear finite element method for computing eigenvalues and eigenfunctions of second order symmetric elliptic partial differential operators. The weak form is assumed to yield a bilinear form which is bounded and coercive in H^1 . Each step of the adaptive procedure refines elements in which a standard a posteriori error estimator is large and also refines elements in which the computed eigenfunction has high oscillation. The error analysis extends the theory of convergence of adaptive methods for linear elliptic source problems to elliptic eigenvalue problems, and in particular deals with various complications which arise essentially from the nonlinearity of the eigenvalue problem. Because of this nonlinearity, the convergence result holds under the assumption that the initial finite element mesh is sufficiently fine.

Keywords: Second-order Elliptic Problems, Eigenvalues, Adaptive finite element methods, convergence

Mathematics Subject Classification: 65N12, 65N25, 65N30, 65N50

Acknowledgement: We would like to thank Carsten Carstensen for kind support and very useful discussions.

1. Introduction. In the last decades, mesh adaptivity has been widely used to improve the accuracy of numerical solutions to many scientific problems. The basic idea is to refine the mesh only where the error is high, with the aim of achieving an accurate solution using an optimal number of degrees of freedom. There is a large numerical analysis literature on adaptivity, in particular on reliable and efficient a posteriori error estimates (e.g. [1]). Recently the question of convergence of adaptive methods has received intensive interest and a number of convergence results for the adaptive solution of boundary value problems have appeared (e.g. [9, 22, 23, 8, 7, 28]).

We prove here the convergence of an adaptive linear finite element algorithm for computing eigenvalues and eigenvectors of scalar symmetric elliptic partial differential operators in bounded polygonal or polyhedral domains, subject to Dirichlet boundary data. Such problems arise in many applications, e.g. resonance problems, nuclear reactor criticality and the modelling of photonic band gap materials, to name but three.

Our refinement procedure is based on two locally defined quantities, firstly a standard a posteriori error estimator and secondly a measure of the variability (or “oscillation”) of the computed eigenfunction. (Measures of “data oscillation” appear in the theory of adaptivity for boundary value problems, e.g. [22]. In the eigenvalue problem the computed eigenvalue and eigenfunction on the present mesh plays the role of “data” for the next iteration of the adaptive procedure.) Our algorithm performs local refinement on all elements on which the minimum of these two local quantities is sufficiently large. We prove that the adaptive method converges provided the initial mesh is sufficiently fine. The latter condition, while absent for adaptive methods for linear symmetric elliptic boundary value problems, commonly appears for nonlinear problems and can be thought of as a manifestation of the nonlinearity of the eigenvalue problem.

We believe that the present paper is the first contribution to the topic of convergence of adaptive methods for eigenvalue problems. Since writing this paper substantial improvements in the theory have been made in [6], where the need to adapt on the oscillations of the eigenvalue is removed and, in addition, the general convergence of the adaptive scheme to a non-spurious eigenvalue of the continuous problem is established.

The outline of the paper is as follows. In Section 2 we briefly describe the model elliptic eigenvalue problem and the numerical method and in Section 3 we describe a priori estimates, most of which are classical. Section 4 describes the a posteriori estimates and the adaptive algorithm. Section 5 proves that proceeding from one mesh to another ensures error reduction (up to oscillation of the computed eigenfunction) while the convergence result is presented in Section 6. Numerical experiments illustrating the theory are presented in Section 7.

2. Eigenvalue problem and numerical method. Throughout, Ω will denote a bounded domain in \mathbb{R}^d ($d = 2$ or 3). In fact Ω will be assumed to be a polygon ($d = 2$) or polyhedron ($d = 3$). We will be concerned with the problem of finding an eigenvalue $\lambda \in \mathbb{R}$ and eigenfunction $0 \neq u \in H_0^1(\Omega)$ satisfying

$$a(u, v) := \lambda b(u, v), \quad \text{for all } v \in H_0^1(\Omega), \quad (2.1)$$

*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK. Stefano.Giani@nottingham.ac.uk

†Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK. I.G.Graham@bath.ac.uk

where, for real valued functions u and v ,

$$a(u, v) = \int_{\Omega} \nabla u(x)^T \mathcal{A}(x) \nabla v(x) dx \quad \text{and} \quad b(u, v) = \int_{\Omega} \mathcal{B}(x) u(x) v(x) dx . \quad (2.2)$$

Here, the matrix-valued function \mathcal{A} is required to be uniformly positive definite, i.e.

$$0 < \underline{a} \leq \xi^T \mathcal{A}(x) \xi \leq \bar{a} \quad \text{for all} \quad \xi \in \mathbb{R}^d \quad \text{with} \quad |\xi| = 1 \quad \text{and all} \quad x \in \Omega. \quad (2.3)$$

The scalar function \mathcal{B} is required to be bounded above and below by positive constants for all $x \in \Omega$, i.e.

$$0 < \underline{b} \leq \mathcal{B}(x) \leq \bar{b} \quad \text{for all} \quad x \in \Omega. \quad (2.4)$$

We will assume that \mathcal{A} and \mathcal{B} are both piecewise constant on Ω and that any jumps in \mathcal{A} and \mathcal{B} are aligned with the meshes \mathcal{T}_n (introduced below), for all n .

Throughout the paper, for any polygonal (polyhedral) subdomain of $D \subset \Omega$, and any $s \in [0, 1]$, $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$ will denote the standard norm and seminorm in the Sobolev space $H^s(D)$. Also $(\cdot, \cdot)_{0,D}$ denotes the $L_2(D)$ inner product. We also define the energy norm induced by the bilinear form a :

$$\|u\|_{\Omega}^2 := a(u, u) \quad \text{for all} \quad u \in H_0^1(\Omega) ,$$

which, by (2.3), is equivalent to the $H^1(\Omega)$ seminorm. (The equivalence constant depends on the contrast \bar{a}/\underline{a} , but we are not concerned with this dependence in the present paper.) We also introduce the weighted L_2 norm:

$$\|u\|_{0,\mathcal{B},\Omega}^2 = b(u, u) = \int_{\Omega} \mathcal{B}(x) |u(x)|^2 dx .$$

and note the norm equivalence

$$\sqrt{\underline{b}} \|v\|_{0,\Omega} \leq \|v\|_{0,\mathcal{B},\Omega} \leq \sqrt{\bar{b}} \|v\|_{0,\Omega} . \quad (2.5)$$

Rewriting the eigenvalue problem (2.1) in standard normalised form, we seek $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ such that

$$\left. \begin{aligned} a(u, v) &= \lambda b(u, v), \quad \text{for all} \quad v \in H_0^1(\Omega) \\ \|u\|_{0,\mathcal{B},\Omega} &= 1 \end{aligned} \right\} \quad (2.6)$$

By the continuity of a and b and the coercivity of a on $H_0^1(\Omega)$ it is a standard result that (2.6) has a countable sequence of non-decreasing positive eigenvalues λ_j , $j = 1, 2, \dots$ with corresponding eigenfunctions $u_j \in H_0^1(\Omega)$ [4, 13, 29].

In this paper we will need some additional regularity for the eigenfunctions u_j , which will be achieved by making the following regularity assumption for the elliptic problem induced by a :

ASSUMPTION 2.1. *We assume that there exists a constant $C_{\text{ell}} > 0$ and $s \in [0, 1]$ with the following property. For $f \in L_2(\Omega)$, if $v \in H_0^1(\Omega)$ solves the problem $a(v, w) = (f, w)_{0,\Omega}$ for all $w \in H_0^1(\Omega)$, then $\|v\|_{1+s,\Omega} \leq C_{\text{ell}} \|f\|_{0,\Omega}$.*

Assumption 2.1 is satisfied with $s = 1$ when \mathcal{A} is constant (or smooth) and Ω is has a smooth boundary or is a convex polygon. In a range of other practical cases $s \in (0, 1)$, for example Ω non-convex (see [5]), or \mathcal{A} having a discontinuity across an interior interface (see [3]). Under Assumption 2.1 it follows that the eigenfunctions u_j of the problem (2.6) satisfy $\|u_j\|_{1+s,\Omega} \leq C_{\text{ell}} \lambda_j \sqrt{\bar{b}}$.

To approximate problem (2.6) we use the piecewise linear finite element method. Accordingly, let \mathcal{T}_n , $n = 1, 2, \dots$ denote a family of conforming triangular ($d = 2$) or tetrahedral ($d = 3$) meshes on Ω . Each mesh consists of elements denoted $\tau \in \mathcal{T}_n$. We assume that for each n , \mathcal{T}_{n+1} is a refinement of \mathcal{T}_n . For a typical element τ of any mesh, its diameter is denoted H_{τ} and the diameter of its largest inscribed ball is denoted ρ_{τ} . For each n , let H_n denote the piecewise constant mesh function on Ω , whose value on each element $\tau \in \mathcal{T}_n$ is H_{τ} and let $H_n^{\max} = \max_{\tau \in \mathcal{T}_n} H_{\tau}$. Throughout we will assume that the family of meshes \mathcal{T}_n is shape regular, i.e. there exists a constant C_{reg} such that

$$H_{\tau} \leq C_{\text{reg}} \rho_{\tau}, \quad \text{for all} \quad \tau \in \mathcal{T}_n \quad \text{and all} \quad n = 1, 2, \dots . \quad (2.7)$$

In the later sections of the paper the \mathcal{T}_n will be produced by an adaptive process which ensures shape regularity.

We let V_n denote the usual finite dimensional subspace of $H_0^1(\Omega)$, consisting of all continuous piecewise linear functions with respect to the mesh \mathcal{T}_n . Then the discrete formulation of problem (2.6) is to seek the eigenpairs $(\lambda_n, u_n) \in \mathbb{R} \times V_n$ such that

$$\left. \begin{aligned} a(u_n, v_n) &= \lambda_n b(u_n, v_n), \quad \text{for all } v_n \in V_n \\ \|u_n\|_{0, \mathcal{B}, \Omega} &= 1. \end{aligned} \right\} \quad (2.8)$$

The problem (2.8) has $N = \dim V_n$ positive eigenvalues (counted according to multiplicity) which we denote in non-decreasing order as $\lambda_{n,1} \leq \lambda_{n,2} \leq \dots \leq \lambda_{n,N}$. It is well-known (see [29, §6.3]) that for any j , $\lambda_{n,j} \rightarrow \lambda_j$ as $H_n^{\max} \rightarrow 0$ and (by the minimax principle - see e.g. [29, §6.1]) the convergence of the $\lambda_{n,j}$ is monotone decreasing i.e.

$$\lambda_{n,j} \geq \lambda_{m,j} \geq \lambda_j, \quad \text{for all } j = 1, \dots, N, \quad \text{and all } m \geq n. \quad (2.9)$$

Thus it is clear that there exists a *separation constant* $\rho > 0$ (depending on the spectrum of (2.6)) with the following property: If $\lambda_j = \lambda_{j+1} = \dots = \lambda_{j+R-1}$ is any eigenvalue of (2.6) of multiplicity $R \geq 1$, then

$$\frac{\lambda_j}{|\lambda_{n,\ell} - \lambda_j|} \leq \rho, \quad \ell \neq j, j+1, \dots, j+R-1, \quad (2.10)$$

provided H_n^{\max} is sufficiently small. (Note that for $\ell \neq j, j+1, \dots, j+R-1$, $\lambda_{n,\ell} \rightarrow \lambda_\ell \neq \lambda_j$.)

The a priori error analysis for our eigenvalue problem is classical (see, e.g. [4], [13] and [29]). In the next section we briefly recall some of the main known results and also prove a non-classical result (Theorem 3.2) which is essential to the proof of convergence of our adaptive scheme.

3. A priori analysis. In this section we shall assume that λ_j is an eigenvalue of (2.6) and $\lambda_{n,j}$ is its approximation as described above. Let u_j and $u_{n,j}$ be any corresponding normalised eigenvectors as defined in (2.6) and (2.8). From these we obtain the important basic identity:

$$\begin{aligned} a(u_j - u_{n,j}, u_j - u_{n,j}) &= a(u_j, u_j) + a(u_{n,j}, u_{n,j}) - 2a(u_j, u_{n,j}) \\ &= \lambda_j + \lambda_{n,j} - 2\lambda_j b(u_j, u_{n,j}) \\ &= \lambda_{n,j} - \lambda_j + \lambda_j (2 - 2b(u_j, u_{n,j})) \\ &= \lambda_{n,j} - \lambda_j + \lambda_j b(u_j - u_{n,j}, u_j - u_{n,j}). \end{aligned} \quad (3.1)$$

Using this and (2.9), we obtain

$$\|u_j - u_{n,j}\|_{\Omega}^2 = |\lambda_j - \lambda_{n,j}| + \lambda_j \|u_j - u_{n,j}\|_{0, \mathcal{B}, \Omega}^2. \quad (3.2)$$

The following theorem investigates the convergence of discrete eigenpairs. Although parts of it are very well-known, we do not know a suitable reference for all the results given below, so a brief proof is given for completeness. In the proof we make use of the orthogonal projection Q_n of $H_0^1(\Omega)$ onto V_n with respect to the inner product induced by $a(\cdot, \cdot)$, which has the property:

$$a(Q_n u, v_n) = \lambda b(u, v_n) \quad \text{for all } v_n \in V_n, \quad (3.3)$$

In the main result of this paper we prove convergence for adaptive approximations to eigenvalues and eigenvectors *assuming for simplicity a simple eigenvalue*. The following preliminary theorem is stated for a simple eigenvalue. However this result is known for multiple eigenvalues (see, e.g. [29]). More details are given in [11].

THEOREM 3.1. *Let λ_j be a simple eigenvalue of (2.6), let $\lambda_{n,j}$ be its associated approximation from solving (2.8) and let u_j and $u_{n,j}$ be any corresponding normalised eigenvectors. Then for all $1 \leq j \leq N$,*

(i)

$$|\lambda_j - \lambda_{n,j}| \leq \|u_j - u_{n,j}\|_{\Omega}^2; \quad (3.4)$$

(ii) There are constants $C_1, C_2 > 0$ and scalars $\alpha_{n,j} \in \{\pm 1\}$ such that

$$\begin{aligned} \|u_j - \alpha_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega} &\leq C_1(H_n^{\max})^s \|u_j - Q_n u_j\|_{\Omega} \\ &\leq C_1(H_n^{\max})^s \|u_j - \alpha_{n,j}u_{n,j}\|_{\Omega}, \end{aligned} \quad (3.5)$$

where s is as in Assumption 2.1.

(iii) For sufficiently small H_n^{\max} there is a constant C_2 such that

$$\|u_j - \alpha_{n,j}u_{n,j}\|_{\Omega} \leq C_2(H_n^{\max})^s. \quad (3.6)$$

The constants C_1, C_2 depend on the spectral information λ_ℓ , u_ℓ , $\ell = 1, \dots, j$, the separation constant ρ , the constants $C_{\text{ell}}, C_{\text{reg}}$ in Assumption 2.1 and in (2.7) and on the bounds $\bar{a}, \underline{a}, \bar{b}, \underline{b}$ in (2.3), (2.4).

Proof. The estimate (3.4) follows directly from (3.2). Note that (3.4) holds even if $u_{n,j}$ is not close to u , which may occur due to the non-uniqueness of the eigenvectors.

The proof of (3.5) is obtained by a reworking of the results in [29]. By the symmetry of a and b there exists a basis $\{u_{n,\ell} : \ell = 1, \dots, N\}$ of V_n (containing $u_{n,j}$) which is orthonormal with respect to inner product b , and each $u_{n,\ell}$ is an eigenvector of (2.8) corresponding to eigenvalue $\lambda_{n,\ell}$. Then with $\beta_{n,j} := b(Q_n u_j, u_{n,j})$, Parseval's equality yields

$$\|Q_n u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega}^2 = \sum_{\substack{\ell=1 \\ \ell \neq j}}^N b(Q_n u_j, u_{n,\ell})^2. \quad (3.7)$$

Then, since

$$\lambda_{n,\ell}b(Q_n u_j, u_{n,\ell}) = a(Q_n u_j, u_{n,\ell}) = a(u_j, u_{n,\ell}) = \lambda_j b(u_j, u_{n,\ell}),$$

we have $(\lambda_{n,\ell} - \lambda_j)b(Q_n u_j, u_{n,\ell}) = \lambda_j b(u_j - Q_n u_j, u_{n,\ell})$, and so

$$\begin{aligned} \|Q_n u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega}^2 &= \sum_{\substack{\ell=1 \\ \ell \neq j}}^N \left(\frac{\lambda_j}{\lambda_{n,\ell} - \lambda_j} \right)^2 b(u_j - Q_n u_j, u_{n,\ell})^2 \\ &\leq \rho^2 \sum_{\substack{\ell=1 \\ \ell \neq j}}^N b(u_j - Q_n u_j, u_{n,\ell})^2 \leq \rho^2 \|u_j - Q_n u_j\|_{0,\mathcal{B},\Omega}^2, \end{aligned}$$

with the last step again by Parseval's equality. Hence

$$\|u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega} \leq (1 + \rho) \|u_j - Q_n u_j\|_{0,\mathcal{B},\Omega}. \quad (3.8)$$

Moreover

$$\|u_j\|_{0,\mathcal{B},\Omega} - \|u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega} \leq \|\beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega} \leq \|u_j\|_{0,\mathcal{B},\Omega} + \|u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega}.$$

Since the u_j and the $u_{n,j}$ are normalised, this implies

$$1 - \|u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega} \leq |\beta_{n,j}| \leq 1 + \|u_j - \beta_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega}$$

and, combining these with (3.8), we have

$$|\beta_{n,j}| - 1 \leq (1 + \rho) \|u_j - Q_n u_j\|_{0,\mathcal{B},\Omega}.$$

Thus with $\alpha_{n,j} := \text{sign}(\beta_{n,j})$, we have $|\beta_{n,j} - \alpha_{n,j}| \leq (1 + \rho) \|u_j - Q_n u_j\|_{0,\mathcal{B},\Omega}$, and

$$\|u_j - \alpha_{n,j}u_{n,j}\|_{0,\mathcal{B},\Omega} \leq 2(1 + \rho) \|u_j - Q_n u_j\|_{0,\mathcal{B},\Omega}.$$

The first inequality in (3.5) now follows from an application of the standard Aubin-Nitsche duality argument, while the second is just the best approximation of Q_n in the energy norm.

The proof of (3.6) is a slight modification of that given in [29, Theorem 6.2]. The argument consists of obtaining an $\mathcal{O}((H_n^{\max})^{2s})$ estimate for the eigenvalue error $|\lambda_j - \lambda_{n,j}|$ and then combining this with (3.2) and (3.5). \square

The next theorem is a generalisation to eigenvalue problems of the standard monotone convergence property for linear symmetric elliptic PDEs, namely that if one enriches the finite dimensional space, then the error is bound to decrease. This result fails to hold for eigenvalue problems (even for symmetric elliptic partial differential operators), because of the nonlinearity of such problems. The best that we can do is to show that if the finite dimensional space is enriched, then the error will not increase very much. This is the subject of Theorem 3.2.

THEOREM 3.2. *For any $1 \leq j \leq N$, there exists a constant $q > 1$ such that, for $m \geq n$, the corresponding computed eigenpair $(\lambda_{m,j}, u_{m,j})$ satisfies:*

$$\|u_j - \alpha_{m,j} u_{m,j}\|_{\Omega} \leq q \|u_j - \alpha_{n,j} u_{n,j}\|_{\Omega}. \quad (3.9)$$

Proof. From Theorem 3.1 (ii), we obtain

$$\|u_j - \alpha_{m,j} u_{m,j}\|_{0,\mathcal{B},\Omega} \leq C_1 (H_m^{\max})^s \|u_j - Q_m u_j\|_{\Omega}. \quad (3.10)$$

Since \mathcal{T}_m is a refinement of \mathcal{T}_n , it follows that $V_n \subset V_m$ and so the best approximation property of Q_m ensures that

$$\|u_j - Q_m u_j\|_{\Omega} \leq \|u_j - Q_n u_j\|_{\Omega}.$$

Hence from (3.10) and using the fact that $H_m^{\max} \leq H_n^{\max}$, we have

$$\|u_j - \alpha_{m,j} u_{m,j}\|_{0,\mathcal{B},\Omega} \leq C_1 (H_n^{\max})^s \|u_j - Q_n u_j\|_{\Omega}. \quad (3.11)$$

Recalling that (3.2) holds for all eigenfunctions, and using (3.11) and then (2.9), we obtain

$$\begin{aligned} \|u_j - \alpha_{m,j} u_{m,j}\|_{\Omega}^2 &\leq |\lambda_j - \lambda_{m,j}| + \lambda_j \|u_j - \alpha_{m,j} u_{m,j}\|_{0,\mathcal{B},\Omega}^2 \\ &\leq |\lambda_j - \lambda_{m,j}| + \lambda_j C_1^2 (H_n^{\max})^{2s} \|u_j - Q_n u_j\|_{\Omega}^2 \\ &\leq |\lambda_j - \lambda_{n,j}| + \lambda_j C_1^2 (H_n^{\max})^{2s} \|u_j - Q_n u_j\|_{\Omega}^2. \end{aligned} \quad (3.12)$$

Hence, from (3.4) we obtain

$$\|u_j - u_{m,j}\|_{\Omega}^2 \leq \|u_j - u_{n,j}\|_{\Omega}^2 + \lambda_j C_1^2 (H_n^{\max})^{2s} \|u_j - Q_n u_j\|_{\Omega}^2. \quad (3.13)$$

But since Q_n yields the best approximation from V_n in the energy norm, we have

$$\|u_j - u_{m,j}\|_{\Omega}^2 \leq (1 + \lambda_j C_1^2 (H_n^{\max})^{2s}) \|u_j - u_{n,j}\|_{\Omega}^2, \quad (3.14)$$

which is in the required form. \square

REMARK 3.3. *From now on we will be concerned with a true eigenpair (λ_j, u_j) and its computed approximation $(\lambda_{j,n}, u_{j,n})$ on the mesh \mathcal{T}_n . Theorem 3.1 tells us that a priori $\lambda_{n,j}$ is “close” to λ_j and that the spaces spanned by u_j and $u_{n,j}$ are close. From now on we drop the subscript j and we simply write (λ, u) for the eigenpair of (2.6) (λ_n, u_n) for a corresponding eigenpair of (2.8) and the scalar $\alpha_{n,j}$ is abbreviated α_n .*

4. A posteriori analysis. This section contains our a posteriori error estimator and the definition of the adaptive algorithm for which convergence will be proved in the following sections.

Recalling the mesh sequence \mathcal{T}_n defined above, we let \mathcal{S}_n denote the set of all the interior edges (or the set of interior faces in 3D) of the elements of the mesh \mathcal{T}_n . For each $S \in \mathcal{S}_n$, we denote by $\tau_1(S)$ and $\tau_2(S)$ the elements sharing S (i.e. $\tau_1(S) \cap \tau_2(S) = S$) and we write $\Omega(S) = \tau_1(S) \cup \tau_2(S)$. We let \vec{n}_S denote the unit normal vector to S , orientated from $\tau_1(S)$ to $\tau_2(S)$. All elements, faces and edges are considered to be closed sets. Furthermore we denote the diameter of S by H_S . Note that by mesh regularity, $\text{diam}(\Omega(S)) \sim H_{\tau_i(S)}$, $i = 1, 2$.

NOTATION 4.1. We write $A \lesssim B$ when A/B is bounded by a constant which may depend on the functions \mathcal{A} and \mathcal{B} in (2.2), on \underline{a} , \bar{a} , \underline{b} and \bar{b} , on C_{ell} in Assumption 2.1, C_{reg} in (2.7). The notation $A \cong B$ means $A \lesssim B$ and $A \gtrsim B$.

All the constants depending on the spectrum, namely ρ in (2.10), q in (3.9), C_1 and C_2 in (3.5) and (3.6) are handled explicitly. Similarly all mesh size dependencies are explicit. Note that all eigenvalues of (2.8) satisfy $\lambda_n \gtrsim 1$, since $\lambda_n \geq \lambda_1 = a(u_1, u_1) \gtrsim |u_1|_{1,\Omega}^2 \gtrsim \|u_1\|_{0,\Omega}^2 \gtrsim \|u_1\|_{0,\mathcal{B},\Omega}^2 = 1$.

Our error estimator is obtained by adapting standard estimates for source problems to the eigenvalue problem. Analogous eigenvalue estimates can be found in [10] (for the Laplace problem) and [30] (for linear elasticity) and related results are in [18].

For a function g , which is piecewise continuous on the mesh \mathcal{T}_n , we introduce its jump across an edge (face) $S \in \mathcal{S}_n$ by:

$$[g]_S(x) := \left(\lim_{\tilde{x} \in \tau_1(S)} g(\tilde{x}) - \lim_{\tilde{x} \in \tau_2(S)} g(\tilde{x}) \right), \quad \text{for } x \in \text{int}(S).$$

Then for any function v with piecewise continuous gradient on \mathcal{T}_n we define, for $S \in \mathcal{S}_n$

$$J_S(v)(x) := [\vec{n}_S \cdot \mathcal{A}\nabla v]_S(x), \quad \text{for } x \in \text{int}(S).$$

The error estimator η_n on the mesh \mathcal{T}_n is defined as

$$\eta_n^2 := \sum_{S \in \mathcal{S}_n} \eta_{S,n}^2, \quad (4.1)$$

where, for each $S \in \mathcal{S}_n$,

$$\eta_{S,n}^2 := \|H_n \lambda_n u_n\|_{0,\mathcal{B},\Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0,S}^2. \quad (4.2)$$

The following lemma is proved, in a standard way, by adapting the usual arguments for linear source problems. Note again that λ is an eigenvalue of (2.6), λ_n is a nearby eigenvalue of (2.8) and u, u_n are any corresponding normalised eigenfunctions which are only ‘‘near’’ in the sense of Theorem 3.1.

LEMMA 4.2 (Reliability).

$$\| \|u - u_n\| \|_{\Omega} \lesssim \eta_n + G_n, \quad (4.3)$$

and

$$G_n := \frac{1}{2}(\lambda + \lambda_n) \frac{\|u - u_n\|_{0,\mathcal{B},\Omega}^2}{\| \|u - u_n\| \|_{\Omega}}. \quad (4.4)$$

REMARK 4.3. Recalling Remark 3.3, u_n in Lemma 4.2 is any normalised eigenvector of (2.8) corresponding to the simple eigenvalue λ , i.e. its sign is not unique. However the error estimators $\eta_{S,n}$ are independent of the sign of u_n . This is not a contradiction: we shall see that only one choice of eigenfunction will guarantee that the second term on the right-hand side of (4.3) is small, and only in this case is the left-hand side also guaranteed to be small.

A similar result to Lemma 4.2 was proved in [30, Proposition 5].

Proof. To ease readability we set $e_n = u - u_n$ in the proof. Note first that, since (λ, u) and (λ_n, u_n) respectively solve the eigenvalue problems (2.1) and (2.8), we have, for all $w_n \in V_n$,

$$\begin{aligned} \| \|e_n\| \|_{\Omega}^2 &= a(e_n, e_n) \\ &= a(e_n, e_n - w_n) + a(e_n, w_n) \\ &= a(e_n, e_n - w_n) + a(u, w_n) - a(u_n, w_n) \\ &= a(e_n, e_n - w_n) + b(\lambda u - \lambda_n u_n, w_n) \\ &= a(e_n, e_n - w_n) - b(\lambda u - \lambda_n u_n, e_n - w_n) + b(\lambda u - \lambda_n u_n, e_n). \end{aligned} \quad (4.5)$$

To estimate the first two terms on the right-hand side of (4.5), first note that, for all $v \in H_0^1(\Omega)$,

$$a(e_n, v) - b(\lambda u - \lambda_n u_n, v) = -a(u_n, v) + \lambda_n b(u_n, v)$$

Hence, using elementwise integration by parts (and the fact that $\mathcal{A}\nabla u_n$ is constant on each element and v vanishes on $\partial\Omega$), we obtain

$$\begin{aligned} a(e_n, v) - b(\lambda u - \lambda_n u_n, v) &= - \sum_{\tau \in \mathcal{T}_n} \int_{\tau} (\mathcal{A}\nabla u_n) \cdot \nabla v + \lambda_n b(u_n, v) \\ &= - \sum_{S \in \mathcal{S}_n} \int_S J_S(u_n) v + \lambda_n b(u_n, v), \end{aligned} \quad (4.6)$$

and hence, for all $w_n \in V_n$,

$$a(e_n, e_n - w_n) - b(\lambda u - \lambda_n u_n, e_n - w_n) = - \sum_{S \in \mathcal{S}_n} \int_S J_S(u_n)(e_n - w_n) + \lambda_n b(u_n, e_n - w_n). \quad (4.7)$$

Now recall the Scott-Zhang quasi-interpolation operator ([27]) which has the property that, for all $v \in H_0^1(\Omega)$, $I_n v \in V_n$ and

$$\|v - I_n v\|_{0,\tau} \lesssim H_{\tau} |v|_{1,\omega(\tau)}, \quad \|v - I_n v\|_{0,S} \lesssim H_S^{\frac{1}{2}} |v|_{1,\omega(S)}, \quad (4.8)$$

where $\omega(\tau)$ is the union of all elements sharing at least a point with τ , and $\omega(S)$ is the union of all elements sharing at least a point with S . (Note $\Omega(S) \subseteq \omega(S)$.) Substituting $w_n = I_n e_n$ in (4.7) and using the Cauchy-Schwarz inequality, together with estimates (4.8), we obtain:

$$a(e_n, e_n - w_n) - b(\lambda u - \lambda_n u_n, e_n - w_n) \lesssim \eta_n \|e_n\|_{\Omega}. \quad (4.9)$$

To estimate the third term on the right-hand side of (4.5), we simply observe that due to the normalisation in each of the eigenvalue problems (2.1) and (2.8) we have

$$b(\lambda u - \lambda_n u_n, e_n) = (\lambda + \lambda_n)(1 - b(u, u_n)) = \frac{1}{2}(\lambda + \lambda_n) \|e_n\|_{0,\mathcal{B},\Omega}^2. \quad (4.10)$$

Now, combine (4.9) and (4.10) with (4.5) and divide by $\|e_n\|_{\Omega}$ to obtain the result. \square

REMARK 4.4. *We shall see below that G_n defined above constitutes a “higher order term”.*

For mesh refinement based on the local contributions to η_n , we use the same marking strategy as in [9] and [22]. The idea is to refine a subset of the elements of \mathcal{T}_n whose side residuals sum up to a fixed proportion of the total residual η_n .

DEFINITION 4.5 (Marking Strategy 1). *Given a parameter $0 < \theta < 1$, the procedure is: mark the sides in a minimal subset $\hat{\mathcal{S}}_n$ of \mathcal{S}_n such that*

$$\left(\sum_{S \in \hat{\mathcal{S}}_n} \eta_{S,n}^2 \right)^{1/2} \geq \theta \eta_n. \quad (4.11)$$

To compute $\hat{\mathcal{S}}_n$, we compute all the “local residuals” $\eta_{S,n}$, then insert edges (faces) into $\hat{\mathcal{S}}_n$ in order of non-increasing magnitude of $\eta_{S,n}$, until (4.11) is satisfied. A minimal subset $\hat{\mathcal{S}}_n$ may not be unique. After this is done, we construct another set $\hat{\mathcal{T}}_n$, containing all the elements of \mathcal{T}_n which contain at least one edge (face) belonging to $\hat{\mathcal{S}}_n$.

In order to prove our convergence theory, we require an additional marking strategy based on oscillations (Definition 4.7 below). This also appears in some theories of adaptivity for source problems, e.g. [9], [22], [20], [8] and [7]), but to our knowledge has not yet been used in connection with eigenvalue problems.

The concept of “oscillation” is just a measure of how well a function may be approximated by piecewise constants on a particular mesh. For any function $v \in L_2(\Omega)$, and any mesh \mathcal{T}_n , we introduce its orthogonal projection $P_n v$ onto piecewise constants defined by:

$$(P_n v)|_\tau = \frac{1}{|\tau|} \int_\tau v_n, \quad \text{for all } \tau \in \mathcal{T}_n. \quad (4.12)$$

Then we make the definition:

DEFINITION 4.6 (Oscillations). *On a mesh \mathcal{T}_n , we define*

$$\text{osc}(v, \mathcal{T}_n) := \|H_n(v - P_n v)\|_{0, \mathcal{B}, \Omega}. \quad (4.13)$$

Note that

$$\text{osc}(v, \mathcal{T}_n) = \left(\sum_{\tau \in \mathcal{T}_n} H_\tau^2 \|v - P_n v\|_{0, \mathcal{B}, \tau}^2 \right)^{1/2}.$$

and that (by standard approximation theory and the ellipticity of $a(\cdot, \cdot)$),

$$\text{osc}(v, \mathcal{T}_n) \lesssim (H_n^{\max})^2 \|v\|_\Omega, \quad \text{for all } v \in H_0^1(\Omega). \quad (4.14)$$

The second marking strategy (introduced below) aims to reduce the oscillations corresponding to a particular approximate eigenfunction u_n .

DEFINITION 4.7 (Marking Strategy 2). *Given a parameter $0 < \tilde{\theta} < 1$: mark the elements in a minimal subset $\tilde{\mathcal{T}}_n$ of \mathcal{T}_n such that*

$$\text{osc}(u_n, \tilde{\mathcal{T}}_n) \geq \tilde{\theta} \text{osc}(u_n, \mathcal{T}_n). \quad (4.15)$$

Analogously to (4.11), we compute $\tilde{\mathcal{T}}_n$ by inserting elements τ into $\tilde{\mathcal{T}}_n$ according to non-increasing order of their local contributions $H_\tau^2 \|(u_n - P_n u_n)\|_{0, \mathcal{B}, \tau}^2$ until (4.15) is satisfied.

Our adaptive algorithm can then be stated:

Algorithm 1 Converging algorithm

Require: $0 < \theta < 1$

Require: $0 < \tilde{\theta} < 1$

loop

Solve the Problem (2.8) for (λ_n, u_n)

Mark the elements using the first marking strategy (Definition 4.5)

Mark any additional unmarked elements using the second marking strategy (Definition 4.7)

Refine the mesh \mathcal{T}_n and construct \mathcal{T}_{n+1}

end loop

In 2D at the n th iteration in Algorithm 1 each element in the set $\hat{\mathcal{T}}_n \cup \tilde{\mathcal{T}}_n$ is refined using the algorithm illustrated in Figure 4.1. This consists of three recursive applications of the newest node algorithm [21] to each marked triangle, first creating two sons, then four grandsons and finally bisecting two of the grandsons. This well-known algorithm is stated without name in [22, §5.1]), is called “bisection5” in [8] and is called “full refinement” in [28]. This technique creates of a new node in the middle of each marked side in $\hat{\mathcal{S}}_n$ and also a new node in the interior of each marked element. It follows from [21] that this algorithm yields shape regular conforming meshes in 2D.

In the 3D-case we use a suitable refinement that creates a new node on each marked face in $\hat{\mathcal{S}}_n$ and a node in the interior of each marked element.

In [22] and [20] it has been shown for linear source problems that the reduction of the error, as the mesh is refined, is triggered by the decay of oscillations of the source on the sequence of constructed meshes. For the eigenvalue problem (2.1) the quantity λu plays the role of data and in principle we have to ensure that oscillations of this quantity (or more precisely of its finite element approximation $\lambda_n u_n$), are sufficiently small. However $\lambda_n u_n$ may change if the mesh changes and so the proof of error reduction for eigenvalue problems is not as simple as it is for linear source problems. This is the essence of the theoretical difficulty dealt with in this paper.

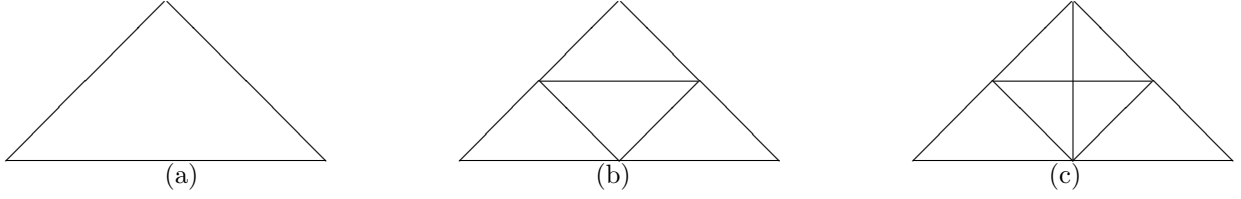


FIG. 4.1. The refinement procedure applied to an element of the mesh. In (a) the element before the refinement, in (b) after the three sides as been refined and in (c) after the bisection of one of the three new segments.

5. Error Reduction. In this section we give the proof of error reduction for Algorithm 1. The proof has been inspired by the corresponding theory for source problems in [22]. However the nonlinearity of the eigenvalue problem introduces new complications and there are several lemmas before the main theorem (Theorem 5.6). For the rest of the section let (λ_n, u_n) be an approximate eigenpair on a mesh \mathcal{T}_n , let \mathcal{T}_{n+1} be the mesh obtained by one iteration of Algorithm 1 and let (λ_{n+1}, u_{n+1}) be the corresponding eigenpair in the sense made precise in Remark 3.3.

The first lemma uses ideas from [22, Lemma 4.2] for the 2D case. The extension of this lemma to the 3D case is treated in Remark 5.2.

LEMMA 5.1. Consider the 2D case. Let $\hat{\mathcal{S}}_n$ be as defined in Definition 4.5 and let P_n be as defined in (4.12). For any $S \in \hat{\mathcal{S}}_n$, there exists a function $\Phi_S \in V_{n+1}$ such that $\text{supp}(\Phi_S) = \Omega(S)$ and also

$$\lambda_n \int_{\Omega(S)} \mathcal{B}(P_n u_n) \Phi_S - \int_S J_S(u_n) \Phi_S = \|H_n \lambda_n P_n u_n\|_{0, \mathcal{B}, \Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0, S}^2, \quad (5.1)$$

and

$$\|\Phi_S\|_{\Omega(S)}^2 \lesssim \|H_n \lambda_n P_n u_n\|_{0, \mathcal{B}, \Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0, S}^2, \quad (5.2)$$

where $\|v\|_{\Omega(S)}^2 := \int_{\Omega(S)} \nabla v^T \mathcal{A} \nabla v$.

Proof. Figure 5.1 illustrates two possible configurations of the domain $\Omega(S)$.



FIG. 5.1. Two cases of refined couples of elements .

We then define

$$\Phi_S := \alpha_S \varphi_S + \beta_1 \varphi_1 + \beta_2 \varphi_2, \quad (5.3)$$

where φ_S and φ_i are the nodal basis functions associated with the points x_S and x_i on \mathcal{T}_{n+1} , and α_S, β_i are defined by

$$\alpha_S = \begin{cases} -\frac{\|H_S^{1/2} J_S(u_n)\|_{0, S}^2}{\int_S J_S(u_n) \varphi_S} & \text{if } J_S(u_n) \neq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (5.4)$$

and

$$\beta_i = \begin{cases} \frac{\|H_n \lambda_n P_n u_n\|_{0,\mathcal{B},\tau_i(S)}^2 - \alpha_S \int_{\tau_i(S)} \mathcal{B} \lambda_n(P_n u_n) \varphi_S}{\int_{\tau_i(S)} \mathcal{B} \lambda_n(P_n u_n) \varphi_i} & \text{if } P_n u_n|_{\tau_i(S)} \neq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (5.5)$$

for $i = 1, 2$.

Note that $J_S(u_n)$ and $P_n u_n$ are constant on each element τ . Using the fact that $\text{supp}(\varphi_i) = \tau_i(S)$, for $i = 1, 2$ we can easily see that the above formulae imply

$$\alpha_S \int_S J_S(u_n) \varphi_S = -\|H_S^{1/2} J_S(u_n)\|_{0,S}^2, \quad (5.6)$$

$$\int_{\Omega(S)} \mathcal{B} \lambda_n(P_n u_n) (\alpha_S \varphi_S + \beta_1 \varphi_1 + \beta_2 \varphi_2) = \|H_n \lambda_n P_n u_n\|_{0,\mathcal{B},\Omega(S)}^2, \quad (5.7)$$

(and that these formulae remain true even if $J_S(u_n)$ or $P_n u_n|_{\tau_i(S)}$ vanish). Hence

$$\lambda_n \int_{\Omega(S)} \mathcal{B}(P_n u_n) \Phi_S - \int_S J_S(u_n) \Phi_S = \lambda_n \int_{\Omega(S)} \mathcal{B}(P_n u_n) (\alpha_S \varphi_S + \beta_1 \varphi_1 + \beta_2 \varphi_2) - \alpha_S \int_S J_S(u_n) \varphi_S$$

and (5.1) follows immediately on using (5.6) and (5.7).

To proceed from here note that by the shape-regularity of the mesh and the standard inverse estimate,

$$\|\phi_S\|_{\Omega(S)} \lesssim H_S^{-1} \|\phi_S\|_{0,\Omega(S)}.$$

Also for all elements $\tau \in \mathcal{T}_{n+1}$ with $\tau \subset \text{supp } \phi_S$, there exists an affine map $\chi : \hat{\tau} \rightarrow \tau$, where $\hat{\tau}$ is the unit simplex in \mathbb{R}^2 and $\hat{\phi}_S := \phi_S \circ \chi$ is a nodal basis function on $\hat{\tau}$. The Jacobian J_χ of χ is constant and is proportional to the area of τ . Hence

$$\|\phi_S\|_{0,\tau}^2 = \int_\tau |\phi_S|^2 = \int_{\hat{\tau}} |\hat{\phi}_S|^2 J_\chi \sim H_S^2,$$

which ensures at $\|\phi_S\|_{\Omega(S)} \lesssim 1$ and, similarly, $\|\varphi_i\|_{\Omega(S)} \lesssim 1$. Combining these with (5.3), we obtain

$$\|\Phi_S\|_{\Omega(S)}^2 \lesssim |\alpha_S|^2 + |\beta_1|^2 + |\beta_2|^2. \quad (5.8)$$

Now, note that by a simple change of variable, $\int_S \varphi_S$ is the integral over $[-H_S/2, H_S/2]$ of the one-dimensional hat function centred on 0 and so $\int_S \varphi_S \sim H_S$. Since $J_S(u_n)$ is constant on S , we have

$$|\alpha_S| \lesssim \frac{|J_S(u_n)| \|H_S^{1/2}\|_{0,S}^2}{H_S} \lesssim |J_S(u_n)| H_S \sim \|H_S^{1/2} J_S(u_n)\|_{0,S}. \quad (5.9)$$

Also since $P_n u_n$ is constant on each $\tau_i(S)$ and since $\int_{\tau_i(S)} \mathcal{B} \phi_i \sim H_{\tau_i(S)}^2$, we have

$$\begin{aligned} |\beta_i| &\lesssim \frac{\lambda_n |P_n u_n|_{\tau_i(S)} \|\mathcal{B}\|_{0,\mathcal{B},\tau_i(S)} + |\alpha_S| H_{\tau_i(S)}^2}{H_{\tau_i(S)}^2} \\ &\lesssim \lambda_n |P_n u_n|_{\tau_i(S)} \|\mathcal{B}\|_{0,\mathcal{B},\tau_i(S)} + |\alpha_S| \sim \lambda_n \|H_n P_n u_n\|_{0,\mathcal{B},\tau_i(S)} + |\alpha_S| \end{aligned}$$

This implies

$$|\beta_i|^2 \lesssim \|\lambda_n H_n P_n u_n\|_{0,\mathcal{B},\tau_i(S)}^2 + |\alpha_S|^2 \lesssim \|\lambda_n H_n P_n u_n\|_{0,\mathcal{B},\tau_i(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0,S}^2, \quad (5.10)$$

and the proof is completed by combining (5.8) with (5.9) and (5.10). \square

REMARK 5.2. *To extend the results in Lemma 5.1 to the 3D-case we need to use a refinement procedure for tetrahedra that creates a new node on each marked face in \hat{S}_n and a node in the interior of each marked*

element. The proof in the 3D-case is similar to the proof in the 2D-case: for each couple of refined elements we define

$$\Phi_S := \alpha_S \varphi_S + \beta_1 \varphi_1 + \beta_2 \varphi_2,$$

where φ_S is the nodal basis function associated to the new node on the shared face and φ_i are the nodal basis functions associated to the new nodes in the interior of the elements. The coefficients α_S , β_1 and β_2 can be chosen in the same way as in Lemma 5.1 and the rest of the proof proceeds similarly.

In the next lemma we bound the local error estimator above by the local difference of two discrete solutions coming from consecutive meshes, plus higher order terms. This kind of result is called “discrete local efficiency” by many authors.

Recall that \mathcal{T}_{n+1} is the refinement of \mathcal{T}_n obtained by applying Algorithm 1.

LEMMA 5.3. *For any $S \in \hat{\mathcal{S}}_n$, we have*

$$\begin{aligned} \eta_{\hat{\mathcal{S}},n}^2 &\lesssim \|u_{n+1} - u_n\|_{\Omega(S)}^2 + \|H_n(\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n)\|_{0,\mathcal{B},\Omega(S)}^2 \\ &\quad + \|H_n \lambda_n (u_n - P_n u_n)\|_{0,\mathcal{B},\Omega(S)}^2. \end{aligned} \quad (5.11)$$

Proof. Since the function Φ_S defined in Lemma 5.1 is in V_{n+1} and $\text{supp}(\Phi_S) = \Omega(S)$, we have

$$a(u_{n+1} - u_n, \Phi_S) = a(u_{n+1}, \Phi_S) - a(u_n, \Phi_S) = \lambda_{n+1} \int_{\Omega(S)} \mathcal{B}u_{n+1} \Phi_S - a(u_n, \Phi_S). \quad (5.12)$$

Now applying integration by parts to the last term on the right-hand side of (5.12), we obtain

$$a(u_{n+1} - u_n, \Phi_S) = \lambda_{n+1} \int_{\Omega(S)} \mathcal{B}u_{n+1} \Phi_S - \int_S J_S(u_n) \Phi_S. \quad (5.13)$$

Rewriting (5.13) and combining with (5.1), we obtain

$$\begin{aligned} a(u_{n+1} - u_n, \Phi_S) - \int_{\Omega(S)} \mathcal{B}(\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n) \Phi_S \\ = \lambda_n \int_{\Omega(S)} \mathcal{B}(P_n u_n) \Phi_S - \int_S J_S(u_n) \Phi_S \\ = \|H_n \lambda_n P_n u_n\|_{0,\mathcal{B},\Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0,S}^2. \end{aligned} \quad (5.14)$$

Rearranging this, and then applying the triangle and Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned} &\|H_n \lambda_n P_n u_n\|_{0,\mathcal{B},\Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0,S}^2 \\ &\leq |a(u_{n+1} - u_n, \Phi_S)| + \left| \int_{\Omega(S)} \mathcal{B}(\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n) \Phi_S \right| \\ &\leq \|u_{n+1} - u_n\|_{\Omega(S)} \|\Phi_S\|_{\Omega(S)} + \|\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n\|_{0,\mathcal{B},\Omega(S)} \|\Phi_S\|_{0,\mathcal{B},\Omega(S)} \\ &\lesssim \left(\|u_{n+1} - u_n\|_{\Omega(S)} + \|H_n(\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n)\|_{0,\mathcal{B},\Omega(S)} \right) \|\Phi_S\|_{\Omega(S)}. \end{aligned} \quad (5.15)$$

In the final step of (5.15) we made use of the Poincaré inequality $\|\Phi_S\|_{0,\mathcal{B},\Omega(S)} \lesssim H_S \|\Phi_S\|_{\Omega(S)}$ and also the shape-regularity of the meshes. In view of (5.2), the fact that $\lambda_n \geq \lambda_1 \gtrsim 1$ (see Notation 4.1) yields

$$\begin{aligned} &\|H_n \lambda_n P_n u_n\|_{0,\mathcal{B},\Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0,S}^2 \\ &\lesssim \left(\|u_{n+1} - u_n\|_{\Omega(S)} + \|H_n(\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n)\|_{0,\mathcal{B},\Omega(S)} \right)^2 \\ &\lesssim \|u_{n+1} - u_n\|_{\Omega(S)}^2 + \|H_n(\lambda_{n+1}u_{n+1} - \lambda_n P_n u_n)\|_{0,\mathcal{B},\Omega(S)}^2. \end{aligned} \quad (5.16)$$

Now, from the definition of $\eta_{S,n}$ in (4.2), and the triangle inequality, we have

$$\eta_{S,n}^2 \lesssim \|H_n \lambda_n P_n u_n\|_{0,\mathcal{B},\Omega(S)}^2 + \|H_S^{1/2} J_S(u_n)\|_{0,S}^2 + \|H_n \lambda_n (u_n - P_n u_n)\|_{0,\mathcal{B},\Omega(S)}^2. \quad (5.17)$$

The required inequality (5.11) now follows from (5.16) and (5.17). \square

In the main result of this section, Theorem 5.6 below, we will be interested in achieving an error reduction result of the form $\|u - \alpha_n u_{n+1}\|_{\Omega} \leq \rho \|u - \alpha_n u_n\|_{\Omega}$ for some $\rho < 1$. Note that we need to introduce the scalar α_n here to ensure nearness of the approximate eigenfunction to the true one.

To prove error reduction we exploit the identity

$$\begin{aligned} \|u - \alpha_n u_n\|_{\Omega}^2 &= \|u - \alpha_{n+1} u_{n+1} + \alpha_{n+1} u_{n+1} - \alpha_n u_n\|_{\Omega}^2 \\ &= \|u - \alpha_{n+1} u_{n+1}\|_{\Omega}^2 + \|\alpha_{n+1} u_{n+1} - \alpha_n u_n\|_{\Omega}^2 \\ &\quad + 2a(u - \alpha_{n+1} u_{n+1}, \alpha_{n+1} u_{n+1} - \alpha_n u_n). \end{aligned} \quad (5.18)$$

In the case of source problems (e.g. [22, 23]) the α_n is not needed and the last term on the right-hand side vanishes due to Galerkin orthogonality. However this approach is not available to us in the eigenvalue problem. Therefore a more technical approach is needed to bound the last two terms on the right-hand side of (5.18) from below. The main technical result is in the following lemma. Recall the convention in Notation 4.1.

LEMMA 5.4.

With u, u_n, α_n as in Remark 3.3,

$$\|\alpha_{n+1} u_{n+1} - \alpha_n u_n\|_{\Omega}^2 \gtrsim \theta^2 \|u - \alpha_n u_n\|_{\Omega}^2 - \text{osc}(\lambda_n u_n, \mathcal{T}_n)^2 - L_n^2, \quad (5.19)$$

where θ is defined in the marking strategy in Definition 4.5 and L_n satisfies the estimate:

$$L_n \leq \hat{C} (H_n^{\max})^s \|u - \alpha_n u_n\|_{\Omega}, \quad (5.20)$$

where \hat{C} depends on $\theta, \lambda, C_1, C_2$, and q .

REMARK 5.5. Note that the oscillation term in (5.19) is unaffected if we replace $\alpha_n u_n$ by u_n .

Proof. By Definition 4.5 and Lemma 5.3, we have

$$\begin{aligned} \theta^2 \eta_n^2 &\leq \sum_{S \in \mathcal{S}_n} \eta_{S,n}^2 \\ &\lesssim \|\alpha_{n+1} u_{n+1} - \alpha_n u_n\|_{\Omega}^2 + \|H_n (\lambda_{n+1} \alpha_{n+1} u_{n+1} - \lambda_n P_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega}^2 + \text{osc}(\lambda_n u_n, \mathcal{T}_n)^2. \end{aligned}$$

Hence, rearranging and making use of Lemma 4.2 and Remark 4.3, we have

$$\begin{aligned} \|\alpha_{n+1} u_{n+1} - \alpha_n u_n\|_{\Omega}^2 &\gtrsim \theta^2 \eta_n^2 - \|H_n (\lambda_{n+1} \alpha_{n+1} u_{n+1} - \lambda_n P_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega}^2 - \text{osc}(\lambda_n u_n, \mathcal{T}_n)^2 \\ &\gtrsim \theta^2 \|u - \alpha_n u_n\|_{\Omega}^2 - \text{osc}(\lambda_n u_n, \mathcal{T}_n)^2 \\ &\quad - \theta^2 \tilde{G}_n^2 - \|H_n (\lambda_{n+1} \alpha_{n+1} u_{n+1} - \lambda_n P_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega}^2, \end{aligned} \quad (5.21)$$

where \tilde{G}_n is the same as G_n in Lemma 4.2, but with u_n replaced by $\alpha_n u_n$.

Note that (5.21) is of the required form (5.19) with

$$L_n := \left(\theta^2 \tilde{G}_n^2 + \|H_n (\lambda_{n+1} \alpha_{n+1} u_{n+1} - \lambda_n P_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega}^2 \right)^{1/2}.$$

We now estimate the last two terms in (5.21) to obtain (5.20). To estimate \tilde{G}_n , we use Theorem 3.1(ii) to obtain

$$\begin{aligned} \tilde{G}_n &\lesssim \frac{1}{2} (\lambda + \lambda_n) C_1^2 (H_n^{\max})^{2s} \frac{\|u - Q_n u\|_{\Omega}^2}{\|u - \alpha_n u_n\|_{\Omega}} \\ &\leq \frac{1}{2} (\lambda + \lambda_n) C_1^2 (H_n^{\max})^{2s} \|u - \alpha_n u_n\|_{\Omega}. \end{aligned} \quad (5.22)$$

To estimate the last term in (5.21), we first use the triangle inequality to obtain

$$\|H_n(\lambda_{n+1}\alpha_{n+1}u_{n+1} - \lambda_n P_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega} \leq \|H_n(\lambda_{n+1}\alpha_{n+1}u_{n+1} - \lambda_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega} + \text{osc}(\lambda_n u_n, \mathcal{T}_n). \quad (5.23)$$

For the first term on the right-hand side of (5.23), we have

$$\|H_n(\lambda_{n+1}\alpha_{n+1}u_{n+1} - \lambda_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega} \leq H_n^{\max}(\|\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega} + \|\lambda u - \lambda_n \alpha_n u_n\|_{0,\mathcal{B},\Omega}). \quad (5.24)$$

Then, recalling (2.6) and Theorem 3.1, we obtain

$$\begin{aligned} \|\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega} &\leq |\lambda - \lambda_{n+1}|\|u\|_{0,\mathcal{B},\Omega} + \lambda_{n+1}\|u - \alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega} \\ &\leq \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 + \lambda_{n+1}C_1(H_n^{\max})^s \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}. \end{aligned} \quad (5.25)$$

Using Theorem 3.1 (iii) and then Theorem 3.2, this implies

$$\begin{aligned} \|\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega} &\lesssim (C_2 + \lambda_{n+1}C_1)(H_n^{\max})^s \|u - \alpha_{n+1}u_{n+1}\|_{\Omega} \\ &\leq q(C_2 + \lambda_{n+1}C_1)(H_n^{\max})^s \|u - \alpha_n u_n\|_{\Omega}. \end{aligned} \quad (5.26)$$

An identical argument shows

$$\|\lambda u - \lambda_n \alpha_n u_n\|_{0,\mathcal{B},\Omega} \lesssim (C_2 + \lambda_n C_1)(H_n^{\max})^s \|u - \alpha_n u_n\|_{\Omega}. \quad (5.27)$$

Combining (5.26) and (5.27) with (5.24), and using (2.9), we obtain

$$\|H_n(\lambda_{n+1}\alpha_{n+1}u_{n+1} - \lambda_n \alpha_n u_n)\|_{0,\mathcal{B},\Omega} \lesssim (1+q)(C_2 + \lambda_n C_1)(H_n^{\max})^{s+1} \|u - \alpha_n u_n\|_{\Omega}. \quad (5.28)$$

Now combining (5.28) with (5.21), (5.22) and (5.23) we obtain the result. \square

The next theorem contains the main result of this section. It shows that provided we start with a "fine enough" mesh \mathcal{T}_n , the mesh adaptivity algorithm will reduce the error in the energy norm.

THEOREM 5.6 (Error reduction). *For each $\theta \in (0, 1)$, there exists a sufficiently fine mesh threshold H_n^{\max} and constants $\mu > 0$ and $\rho \in (0, 1)$ (all of which may depend on θ and on the eigenvalue λ), with the following property. For any $\varepsilon > 0$ the inequality*

$$\text{osc}(\lambda_n u_n, \mathcal{T}_n) \leq \mu \varepsilon, \quad (5.29)$$

implies either $\|u - \alpha_n u_n\|_{\Omega} \leq \varepsilon$ or

$$\|u - \alpha_{n+1}u_{n+1}\|_{\Omega} \leq \rho \|u - \alpha_n u_n\|_{\Omega}.$$

Proof. In view of equation (5.18) and remembering that $\alpha_{n+1}u_{n+1} - \alpha_n u_n \in V_{n+1}$ we have

$$\begin{aligned} &\|u - \alpha_n u_n\|_{\Omega}^2 - \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 \\ &= \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{\Omega}^2 + 2a(u - \alpha_{n+1}u_{n+1}, \alpha_{n+1}u_{n+1} - \alpha_n u_n) \\ &= \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{\Omega}^2 + 2b(\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}, \alpha_{n+1}u_{n+1} - \alpha_n u_n). \end{aligned} \quad (5.30)$$

Before proceeding further, recall that by the assumptions (2.3) and (2.4), and the Poincaré inequality, there exists a constant C_P (depending on \mathcal{A} , \mathcal{B} and Ω) such that

$$\|v\|_{0,\mathcal{B},\Omega} \leq C_P \|v\|_{\Omega}, \quad \text{for all } v \in H_0^1(\Omega).$$

Now using Cauchy-Schwarz and then the Young inequality $2ab \leq \frac{1}{4C_P^2}a^2 + 4C_P^2b^2$ on the second term on the right-hand side of (5.30), we get

$$\begin{aligned} &\|u - \alpha_n u_n\|_{\Omega}^2 - \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 \\ &\geq \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{\Omega}^2 - 2\|\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega} \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{0,\mathcal{B},\Omega} \\ &\geq \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{\Omega}^2 - \frac{1}{4C_P^2} \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{0,\mathcal{B},\Omega}^2 - 4C_P^2 \|\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega}^2 \\ &\geq \frac{3}{4} \|\alpha_{n+1}u_{n+1} - \alpha_n u_n\|_{\Omega}^2 - 4C_P^2 \|\lambda u - \lambda_{n+1}\alpha_{n+1}u_{n+1}\|_{0,\mathcal{B},\Omega}^2. \end{aligned} \quad (5.31)$$

Hence

$$\| \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 \leq \| \|u - \alpha_n u_n\|_{\Omega}^2 - \frac{3}{4} \| \alpha_{n+1}u_{n+1} - \alpha_n u_n \|_{\Omega}^2 + 4C_P^2 \| \lambda u - \lambda_{n+1} \alpha_{n+1} u_{n+1} \|_{0,\mathcal{B},\Omega}^2 .$$

Applying Lemma 5.4, we see that there exist constants C, \hat{C} such that

$$\begin{aligned} \| \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 &\leq \left(1 - \frac{3}{4}C\theta^2 + \frac{3}{4}C\hat{C}^2(H_n^{\max})^{2s}\right) \| \|u - \alpha_n u_n\|_{\Omega}^2 \\ &\quad + 4C_P^2 \| \lambda u - \lambda_{n+1} \alpha_{n+1} u_{n+1} \|_{0,\mathcal{B},\Omega}^2 \\ &\quad + \frac{3}{4}C \operatorname{osc}(\lambda_n u_n, \mathcal{T}_n)^2 . \end{aligned}$$

Then making use of (5.26) we have

$$\| \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 \leq \gamma_n \| \|u - \alpha_n u_n\|_{\Omega}^2 + \frac{3}{4}C \operatorname{osc}(\lambda_n u_n, \mathcal{T}_n)^2 . \quad (5.32)$$

with

$$\gamma_n := \left[1 - \frac{3}{4}C\theta^2 + C'(H_n^{\max})^{2s}\right] , \quad (5.33)$$

where C' is another constant independent of n . Note that H_n^{\max} can be chosen sufficiently small so that $\gamma_m \leq \gamma$ for some $\gamma \in (0, 1)$ and all $m \geq n$.

Consider now the consequences of the inequality (5.29). If $\| \|u - \alpha_n u_n\|_{\Omega} > \varepsilon$ then (5.32) implies

$$\| \|u - \alpha_{n+1}u_{n+1}\|_{\Omega}^2 \leq \left[\gamma + \frac{3}{4}C\mu^2\right] \| \|u - \alpha_n u_n\|_{\Omega}^2 .$$

Now choose μ small enough so that

$$\rho := \left(\gamma + \frac{3}{4}C\mu^2\right)^{1/2} < 1 \quad (5.34)$$

to complete the proof. \square

6. Proof of convergence. The main result of this paper is Theorem 6.2 below which proves convergence of the adaptive method and also demonstrates the decay of oscillations of the sequence of approximate eigenfunctions. Before proving this result we need a final lemma.

LEMMA 6.1. *There exists a constant $\tilde{\rho} \in (0, 1)$ such that*

$$\operatorname{osc}(u_{n+1}, \mathcal{T}_{n+1}) \leq \tilde{\rho} \operatorname{osc}(u_n, \mathcal{T}_n) + (1+q)(H_n^{\max})^2 \| \|u - \alpha_n u_n\|_{\Omega} . \quad (6.1)$$

Proof. First recall that one of the key results in [22], namely [22, Lemma 3.8], is the proof that the oscillations of any fixed function $v \in H_0^1(\Omega)$ are reduced by applying one refinement based on Marking Strategy 2 (Definition 4.7). Thus we have (in view of Algorithm 1):

$$\operatorname{osc}(u_n, \mathcal{T}_{n+1}) \leq \tilde{\rho} \operatorname{osc}(u_n, \mathcal{T}_n), \quad (6.2)$$

where $0 < \tilde{\rho} < 1$ is independent of u_n . Thus, a simple application of the triangle inequality combined with (6.2) yields

$$\begin{aligned} \operatorname{osc}(u_{n+1}, \mathcal{T}_{n+1}) &\leq \operatorname{osc}(u_n, \mathcal{T}_{n+1}) + \operatorname{osc}(\alpha_{n+1}u_{n+1} - \alpha_n u_n, \mathcal{T}_{n+1}) \\ &\leq \tilde{\rho} \operatorname{osc}(u_n, \mathcal{T}_n) + \operatorname{osc}(\alpha_{n+1}u_{n+1} - \alpha_n u_n, \mathcal{T}_{n+1}) \end{aligned} \quad (6.3)$$

(Recall again that $\operatorname{osc}(u_n, \mathcal{T}_n) = \operatorname{osc}(\alpha_n u_n, \mathcal{T}_n)$.) A further application of the triangle inequality and then (4.14) yields

$$\begin{aligned} \operatorname{osc}(\alpha_{n+1}u_{n+1} - \alpha_n u_n, \mathcal{T}_{n+1}) &\leq \operatorname{osc}(u - \alpha_{n+1}u_{n+1}, \mathcal{T}_{n+1}) + \operatorname{osc}(u - \alpha_n u_n, \mathcal{T}_{n+1}) \\ &\lesssim (H_n^{\max})^2 (\| \|u - \alpha_{n+1}u_{n+1}\|_{\Omega} + \| \|u - \alpha_n u_n\|_{\Omega}) \end{aligned} \quad (6.4)$$

and then combining (6.3) and (6.4) and applying Theorem 3.2 completes the proof. \square

THEOREM 6.2. *Provided the initial mesh \mathcal{T}_0 is chosen so that H_0^{\max} is small enough, there exists a constant $p \in (0, 1)$ such that the recursive application of Algorithm 1 yields a convergent sequence of approximate eigenvalues and eigenvectors, with the property:*

$$\| \|u - \alpha_n u_n \| \|_{\Omega} \leq B_0 q p^n, \quad (6.5)$$

and

$$\lambda_n \text{ osc}(u_n, \mathcal{T}_n) \leq B_1 p^n, \quad (6.6)$$

where B_0 and B_1 are constants and q is the constant defined in Theorem 3.2.

REMARK 6.3. *The initial mesh convergence threshold and the constants B_0 and B_1 may depend on θ , $\tilde{\theta}$ and λ .*

Proof. The proof of this theorem is by induction and the induction step contains an application of Theorem 5.6. In order to ensure the reduction of the error, we have to assume that the starting mesh \mathcal{T}_0 is fine enough and μ in Theorem 5.6 is small enough such that for the chosen value of θ , the quantity ρ in (5.34) satisfies $\rho < 1$.

Then with $\tilde{\rho}$ as in Lemma 6.1, choose p in the range

$$\max\{\rho, \tilde{\rho}\} < p < 1.$$

We also set

$$B_1 = \text{osc}(\lambda_0 u_0, \mathcal{T}_0) \quad \text{and} \quad B_0 = \max\{\mu^{-1} p^{-1} B_1, \| \|u - u_0 \| \|_{\Omega}\}.$$

To perform the inductive proof, first note that by the definition of B_0 and Theorem 3.2,

$$\| \|u - u_0 \| \|_{\Omega} \leq B_0 \leq B_0 q,$$

since $q > 1$. Combined with the definition of B_1 we have shown the result for $n = 0$.

Now, suppose that for some $n > 0$ the inequalities (6.5) and (6.6) hold.

Now let us consider the outcomes, depending on whether the inequality

$$\| \|u - \alpha_n u_n \| \|_{\Omega} \leq B_0 p^{n+1}, \quad (6.7)$$

holds or not. If (6.7) holds then we can apply Theorem 3.2 to conclude that

$$\| \|u - \alpha_{n+1} u_{n+1} \| \|_{\Omega} \leq q \| \|u - \alpha_n u_n \| \|_{\Omega} \leq q B_0 p^{n+1},$$

which proves (6.5) for $n + 1$.

On the other hand, if (6.7) does not hold then, by definition of B_0 ,

$$\| \|u - \alpha_n u_n \| \|_{\Omega} > B_0 p^{n+1} \geq \mu^{-1} B_1 p^n. \quad (6.8)$$

Also, since we have assumed (6.6) for n , we have

$$\lambda_n \text{ osc}(u_n, \mathcal{T}_n) \leq \mu \varepsilon \quad \text{with} \quad \varepsilon := \mu^{-1} B_1 p^n. \quad (6.9)$$

Then (6.8) and (6.9) combined with Theorem 5.6 yields

$$\| \|u - \alpha_{n+1} u_{n+1} \| \|_{\Omega} \leq \rho \| \|u - \alpha_n u_n \| \|_{\Omega}$$

and so using the inductive hypothesis (6.5) combined with the definition of p , we have

$$\| \|u - \alpha_{n+1} u_{n+1} \| \|_{\Omega} \leq \rho B_0 q p^n \leq q B_0 p^{n+1},$$

which again proves (6.5) for $n + 1$.

To conclude the proof, we have to show that also (6.6) holds for $n + 1$. Using Lemma 6.1, (2.9) and the inductive hypothesis, we have

$$\begin{aligned} \lambda_{n+1} \operatorname{osc}(u_{n+1}, \mathcal{T}_{n+1}) &\leq \tilde{\rho} B_1 p^n + (1+q)(H_n^{\max})^2 \lambda_n B_0 q p^n \\ &\leq (\tilde{\rho} B_1 + (1+q)(H_0^{\max})^2 \lambda_0 B_0 q) p^n. \end{aligned} \quad (6.10)$$

Now, (recalling that $\tilde{\rho} < p$), in addition to the condition already imposed on H_0^{\max} we can further require that

$$\tilde{\rho} B_1 + (1+q)(H_0^{\max})^2 \lambda_0 B_0 q \leq p B_1.$$

This ensures that

$$\lambda_{n+1} \operatorname{osc}(u_{n+1}, \mathcal{T}_{n+1}) \leq B_1 p^{n+1},$$

thus concluding the proof. \square

7. Numerical Experiments. We present numerical experiments to illustrate the convergence theory. Algorithm 1 has been implemented in FORTRAN95. The mesh refinement has been done using the toolbox ALBERTA [25]. We used the package ARPACK [19] to compute eigenpairs and the sparse direct linear solver ME27 from the HSL [26, 14] to carry out the shift-invert solves required by ARPACK. Additional numerical experiments on photonic crystal problems and on 3D problems are given in [11] and [12].

7.1. Example: Laplace operator. In the first set of simulations we have solved the Laplace eigenvalue problem (i.e. $\mathcal{A} = I$ and $\mathcal{B} = 1$ in (2.2)) on a unit square with Dirichlet boundary conditions. The exact eigenvalues are known explicitly.

We compare different runs of Algorithm 1 using different values for θ and $\tilde{\theta}$ in Table 7.1. Since the problem is smooth, from Theorem 3.1 it follows that using uniform refinement the rate of convergence for eigenvalues should be $\mathcal{O}(H_n^{\max})^2$, or equivalently the rate of convergence in the number of degrees of freedom (DOFs) N should be $\mathcal{O}(N^{-1})$. We measure the rate of convergence by conjecturing that $|\lambda - \lambda_n| = CN^{-\beta}$ and estimating β for each pair of computations from the formula $\beta = -\log(|\lambda - \lambda_n|/|\lambda - \lambda_{n-1}|)/\log(\text{DOFs}_n/\text{DOFs}_{n-1})$. Similarly Table 7.2 contains the same kind of information relative to the fourth smallest eigenvalue of the problem. Our results show a convergence rate close to $\mathcal{O}(N^{-1})$ for $\theta, \tilde{\theta}$ sufficiently large. However the rate of convergence is sensitive to the values of θ and $\tilde{\theta}$.

Iteration	$\theta = \tilde{\theta} = 0.2$			$\theta = \tilde{\theta} = 0.5$			$\theta = \tilde{\theta} = 0.8$		
	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β
1	0.1350	400	-	0.1350	400	-	0.1350	400	-
2	0.1327	498	0.0802	0.1177	954	0.1581	0.0529	1989	0.5839
3	0.1293	613	0.1228	0.0779	1564	0.8349	0.0176	5205	1.1407
4	0.1256	731	0.1645	0.0501	1977	1.8788	0.0073	15980	0.7877
5	0.1215	854	0.2138	0.0351	2634	1.2383	0.0024	48434	0.9836
6	0.1165	970	0.3340	0.0176	4004	0.7885	0.0009	122699	1.0673
7	0.1069	1097	0.6962	0.0121	6588	0.7217	0.0003	312591	1.0083

TABLE 7.1

Comparison of the reduction of the error and DOFs of the adaptive method for the smallest eigenvalue for the Laplace problem on the unit square.

In the theory presented in [29] it is shown that the error for eigenvalues for smooth problems is bounded in terms of the square of the considered eigenvalue, i.e.

$$|\lambda - \lambda_n| \leq C \lambda^2 (H_n^{\max})^2. \quad (7.1)$$

Also, we know that the first and the fourth eigenvalues are 19.7392089 and 78.9568352 and so $\lambda_4 = 4\lambda_1$. Comparing errors in Table 7.2 with those in Table 7.1, we see that the errors are roughly multiplied by a factor of 16, as predicted by (7.1).

Iteration	$\theta = \tilde{\theta} = 0.2$			$\theta = \tilde{\theta} = 0.5$			$\theta = \tilde{\theta} = 0.8$		
	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β
1	2.1439	400	-	2.1439	400	-	2.1439	400	-
2	2.0997	505	0.0895	1.8280	1016	0.1658	0.7603	2039	0.6365
3	2.0549	626	0.1004	1.0850	1636	1.1662	0.2439	6793	0.9447
4	1.9945	759	0.1548	0.7792	2254	1.0331	0.0917	18717	0.9652
5	1.9164	883	0.2638	0.4936	3067	1.4826	0.0331	54113	0.9583
6	1.7717	1017	0.5557	0.3484	4681	0.8240	0.0120	146056	1.0181
7	1.6463	1131	0.6911	0.2578	7321	0.6730	0.0046	382024	0.9970

TABLE 7.2

Comparison of the reduction of the error and DOFs of the adaptive method for the fourth smallest eigenvalue for the Laplace problem on the unit square.

Often h-adaptivity uses only a marking strategy based on an estimation of the error, as in Marking Strategy 1 and avoids refining based on oscillations as in Marking Strategy 2. (Convergence of an adaptive scheme for eigenvalue problems which does not use marking strategy 2 is recently proved in [6].) To investigate the effects of refinement based on oscillations, in Table 7.3 we have computed the smallest eigenvalue for the Laplace problem keeping θ fixed and varying $\tilde{\theta}$ only. Reducing $\tilde{\theta}$ towards 0 has the effect of turning off the refinement arising from Marking Strategy 2. The results in Table 7.3 seem to suggest that the rate of convergence slightly increases as $\tilde{\theta}$ increases.

Iteration	$\theta = 0.8, \tilde{\theta} = 0.1$			$\theta = 0.8, \tilde{\theta} = 0.3$			$\theta = 0.8, \tilde{\theta} = 0.5$		
	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β
1	0.1350	400	-	0.1350	400	-	0.1350	400	-
2	0.0704	1269	0.5646	0.0698	1372	0.5353	0.0673	1555	0.5131
3	0.0307	2660	1.1215	0.0300	2821	1.1700	0.0285	3229	1.1757
4	0.0137	7492	0.7770	0.0133	7846	0.7980	0.0115	9140	0.8731
5	0.0056	18853	0.9699	0.0052	20189	0.9918	0.0046	22793	0.9913
6	0.0021	52247	0.9587	0.0020	55640	0.9382	0.0018	61582	0.9310
7	0.0008	140049	0.9834	0.0008	145773	1.0011	0.0007	161928	1.0238

TABLE 7.3

Comparison of the reduction of the error and DOFs of the adaptive method for the smallest eigenvalue for the Laplace problem on the unit square for a fixed value of θ and varying $\tilde{\theta}$.

We investigate this further in Table 7.4, where we take iterations 5,6 and 7 from Table 7.3 and we present the quantity $C^* := N \times |\lambda - \lambda_n|$, where N denotes the number of DOFs. Then C^* gives an indication of the size of the unknown constant in the optimal error estimate $|\lambda - \lambda_n| = \mathcal{O}(N^{-1})$. The results suggest that C^* stays fairly constant independent of $\tilde{\theta}$.

Iteration	$\theta = 0.8, \tilde{\theta} = 0.1$	$\theta = 0.8, \tilde{\theta} = 0.3$	$\theta = 0.8, \tilde{\theta} = 0.5$
5	1.06×10^2	1.05×10^2	1.05×10^2
6	1.10×10^2	1.11×10^2	1.11×10^2
7	1.12×10^2	1.12×10^2	1.13×10^2

TABLE 7.4

Values of C^* computed from Table 7.3

In Table 7.5 we have set $\tilde{\theta} = 0$. Although the convergence result given in this paper does not hold any more, the method is still clearly convergent. Comparing Table 7.1, Table 7.3 and Table 7.5 we see that with the second marking strategy the number of degrees of freedom grows faster than without it. To illustrate this effect better, Table 7.6 tabulates the number of elements $\#\tilde{\mathcal{T}}_n$ (marked by Marking Strategy 1) with the extra number of elements $\#(\tilde{\mathcal{T}}_n \setminus \hat{\mathcal{T}}_n)$ (marked by Marking Strategy 2 alone). Note that the new DOFs

created by mesh refinement come only from the refinement of the marked elements, but also from the closures used to keep the meshes conforming. It is clear that the number of elements marked as a result of the oscillations continues to rise as refinement proceeds, although much more slowly than the number marked by the residual-based criterion (Marking Strategy 1).

Iteration	$\theta = 0.2$			$\theta = 0.5$			$\theta = 0.8$		
	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β
1	0.1350	400	-	0.1350	400	-	0.1350	400	-
2	0.1328	447	0.1525	0.1209	648	0.2289	0.0704	1253	0.5704
3	0.1299	503	0.1824	0.0859	1036	0.7283	0.0307	2646	1.1125
4	0.1271	565	0.1958	0.0627	1455	0.9301	0.0138	7490	0.7697
5	0.1238	637	0.2157	0.0458	1965	1.0429	0.0056	18847	0.9734
6	0.1189	712	0.3650	0.0323	3031	0.8066	0.0021	52239	0.9585
7	0.1113	795	0.6014	0.0228	4372	0.9531	0.0008	140194	0.9828

TABLE 7.5

Comparison of the reduction of the error and DOFs of the adaptive method for the smallest eigenvalue for the Laplace problem on the unit square using marking strategy 1 only.

Iteration	$\theta = \tilde{\theta} = 0.2$		$\theta = \tilde{\theta} = 0.5$		$\theta = \tilde{\theta} = 0.8$	
	$\#\hat{\mathcal{T}}_n$	$\#(\tilde{\mathcal{T}}_n \setminus \hat{\mathcal{T}}_n)$	$\#\hat{\mathcal{T}}_n$	$\#(\tilde{\mathcal{T}}_n \setminus \hat{\mathcal{T}}_n)$	$\#\hat{\mathcal{T}}_n$	$\#(\tilde{\mathcal{T}}_n \setminus \hat{\mathcal{T}}_n)$
1	12	15	85	99	299	285
2	13	15	102	85	953	19
3	14	15	100	25	3069	198
4	14	14	173	7	7965	2053
5	15	13	310	48	22426	1486
6	15	12	552	184	58075	3005

TABLE 7.6

Comparison between the number of marked elements by strategy 1 (i.e. $\#\hat{\mathcal{T}}_n$) and the number of marked elements by strategy 2 only (i.e. $\#(\tilde{\mathcal{T}}_n \setminus \hat{\mathcal{T}}_n)$) for different values of θ and $\tilde{\theta}$ for the smallest eigenvalue of the Laplace problem on the unit square.

In Figure 7.1 we compare the performance of the adaptive algorithm with uniform bisection5 refinement (see Fig. 4.1) for the first and fourth eigenvalues of the Laplace operator. We note that in this case both methods converge with a similar rate, as is expected since in this case the regularity of eigenfunctions is H^2 . To complete this section we give in Table 7.7 an example of the performance of the adaptive method for computing non-simple eigenvalues. In this case we considered the second smallest eigenvalue of the Laplace operator on the unit square which has multiplicity 2. We see that although the theory given above does not strictly hold, the method performs similarly to the case of simple eigenvalues.

7.2. Example: Elliptic operator with discontinuous coefficients. In this example we investigate how our method copes with discontinuous coefficients. In order to do that we modified the smooth problem from Example 7.1. We inserted a square subdomain of side 0.5 in the center of the unit square domain. In the bilinear form (2.2), we also chose the function \mathcal{A} to be the scalar piecewise constant function which assumes the value 100 inside the inner subdomain and the value 1 outside it. As before \mathcal{B} in (2.2) is chosen as $\mathcal{B} = 1$. The jump in the value of \mathcal{A} generally produces a jump in the gradient of the eigenfunctions all along the boundary of the subdomain and at the corners of the subdomain (from both inside and outside) the eigenfunction has infinite gradient, arising from the usual corner singularities. We choose our initial mesh to be aligned with the discontinuity in \mathcal{A} and so only the corner singularities are active here. We still have Assumption 2.1, but now $s < 1$ and, from Theorem 3.1, using uniform refinement, the rate of convergence for eigenvalues should be $\mathcal{O}(H_n^{\max})^{2s}$ or equivalently $\mathcal{O}(N^{-s})$, where N is the number of DOFs. The adaptive method yields the optimal order of $\mathcal{O}(N^{-1})$ (which holds for uniform meshes and smooth problems) for large

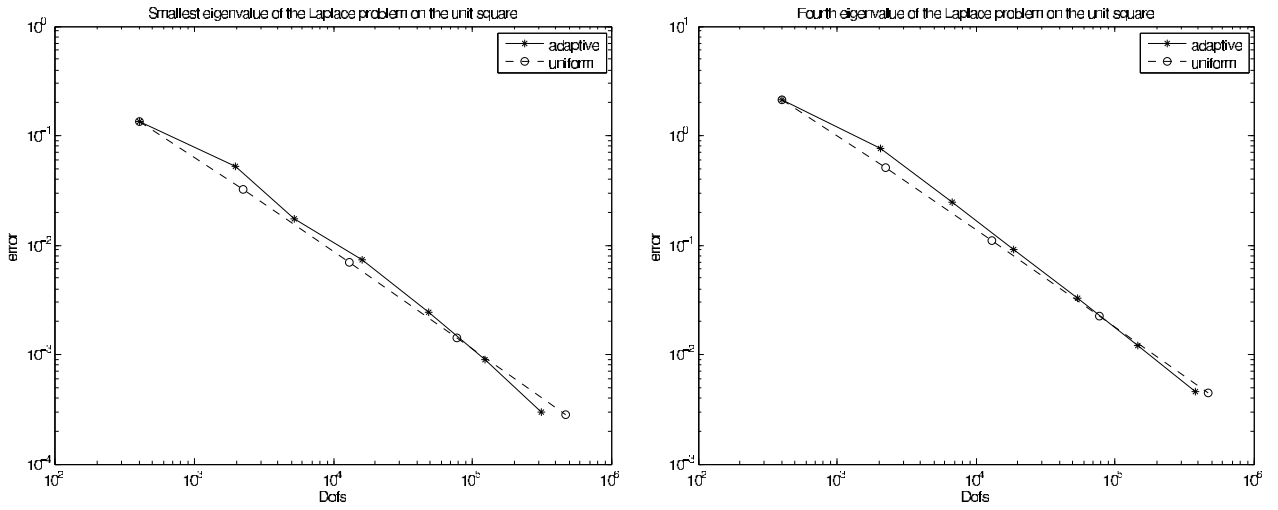


FIG. 7.1. Loglog plots of convergence of adaptive and uniform refinement for first eigenvalue of the Laplacian (left) and fourth eigenvalue of the Laplacian (right).

n	$\theta = \tilde{\theta} = 0.2$			$\theta = \tilde{\theta} = 0.5$			$\theta = \tilde{\theta} = 0.8$		
	$ \lambda - \lambda_n $	N	β	$ \lambda - \lambda_n $	N	β	$ \lambda - \lambda_n $	N	β
1	0.5802	400	-	0.5802	400	-	0.5802	400	-
2	0.5678	478	0.1212	0.4935	811	0.2291	0.2447	1533	0.6427
3	0.5514	562	0.1816	0.3201	1275	0.9564	0.0959	3640	1.0826
4	0.5329	646	0.2449	0.2295	1728	1.0953	0.0368	11747	0.8169
5	0.5111	735	0.3237	0.1521	2374	1.2950	0.0136	32881	0.9651
6	0.4758	829	0.5942	0.1078	3498	0.8875	0.0050	82968	1.0778
7	0.4392	918	0.7856	0.0782	5555	0.6938	0.0020	221521	0.9574

TABLE 7.7

Comparison of the reduction of the error and DOFs of the adaptive method for the second smallest eigenvalue for the Laplace problem on the unit square.

enough θ and $\tilde{\theta}$. (See Table 7.8.) Here we compute the “exact” λ using a mesh with about half a million of DOFs.

Iteration	$\theta = \tilde{\theta} = 0.2$			$\theta = \tilde{\theta} = 0.5$			$\theta = \tilde{\theta} = 0.8$		
	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β	$ \lambda - \lambda_n $	DOFs	β
1	1.1071	81	-	1.1071	81	-	1.1071	81	-
2	1.0200	103	0.3410	0.8738	199	0.2632	0.4834	356	0.5597
3	1.0105	129	0.0416	0.5848	314	0.8805	0.2244	799	0.9494
4	1.0039	147	0.0498	0.3983	491	0.8591	0.0990	2235	0.7957
5	0.8968	167	0.8843	0.2766	673	1.1564	0.0401	4764	1.1932
6	0.8076	194	0.6996	0.1933	975	0.9665	0.0180	12375	0.8372
7	0.8008	217	0.0747	0.1346	1476	0.8722	0.0065	29148	1.1888
8	0.7502	237	0.7401	0.0948	2080	1.0237	0.0020	65387	1.4482

TABLE 7.8

Comparison of the reduction of the error and DOFs of the adaptive method for the smallest eigenvalue for the 2D problem with discontinuous coefficient.

In Figure 7.2 we depict the mesh coming from the fourth iteration of Algorithm 1 with $\theta = \tilde{\theta} = 0.8$ for the smallest eigenvalue of this problem. This mesh is the result of multiple refinements using both marking

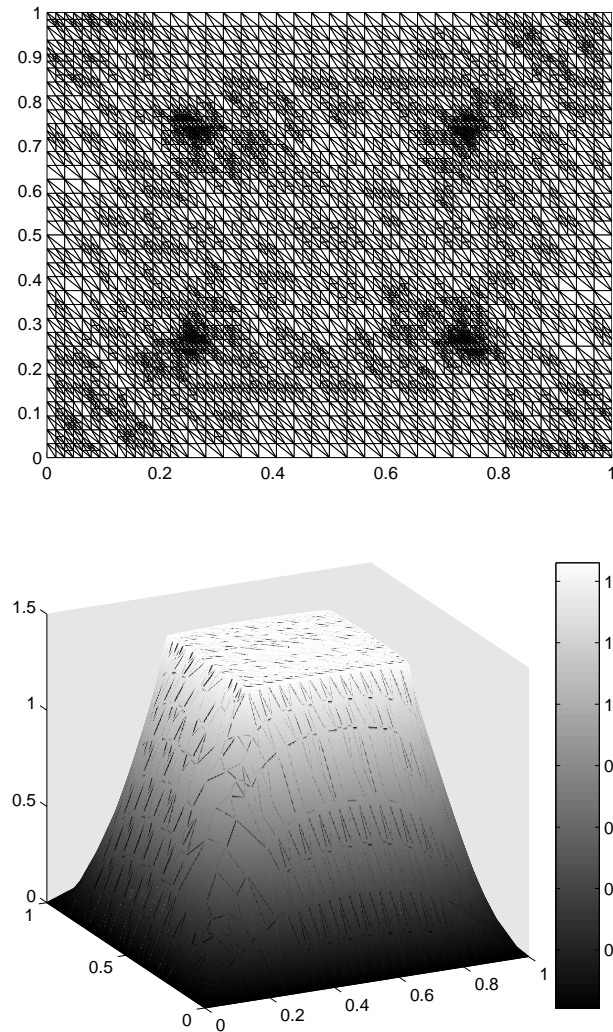


FIG. 7.2. A refined mesh from the adaptive method corresponding to the first eigenvalue of the 2D problem with discontinuous coefficient, and the corresponding eigenfunction .

strategies 1 and 2 each time. As can be seen the corners of the subdomain are much more refined than the rest of the mesh. This is clearly the effect of the first marking strategy, since the edge residuals have detected the singularity in the gradient of the eigenfunction at these points. In Figure 7.2 we also depict the corresponding eigenfunction.

In Figure 7.3, analogously to Figure 7.1, we compare the convergence of the adaptive method with uniform refinement for this example. Now, because of the lack of regularity, the superiority of the adaptive method is clearly visible.

REFERENCES

- [1] M. Ainsworth and J.T. Oden, *A Posterior Error Estimation in Finite Element Analysis*, Wiley, 2000.
- [2] H. Ammari and F. Santosa, Guided waves in a photonic bandgap structure with a line defect, *SIAM J. Appl. Math.* 64(6):2018-2033, 2004.
- [3] I. Babuška, The finite element method for elliptic equations with discontinuous coefficients *Computing* 5:207–213, 1970.
- [4] I. Babuška and J. Osborn. *Eigenvalue Problems*, in Handbook of Numerical Analysis Vol II, eds P.G. Cairlet and J.L. Lions, North Holland, 641-787, 1991.

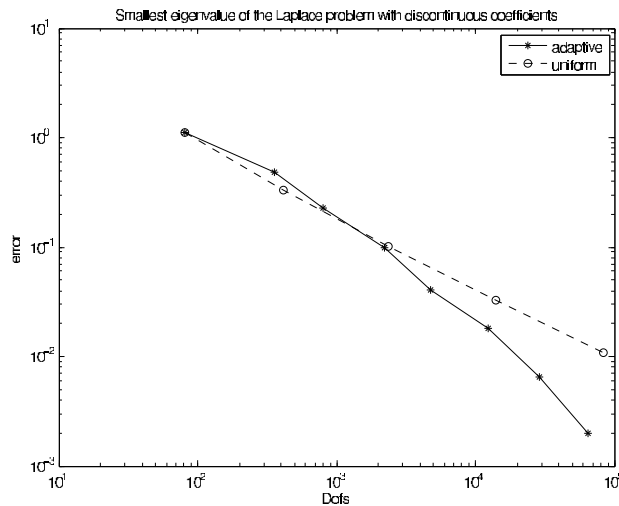


FIG. 7.3. Loglog plot of convergence of adaptive and uniform refinement for first eigenvalue of the problem with discontinuous coefficient.

- [5] M. Bourland, M. Dauge, M.-S. Lubuma and S. Niçaise, Coefficients of the singularities for elliptic boundary value problems on domains with conical points. III: Finite element methods on polygonal domains, *SIAM J. Numer. Anal.* 29: 136-155, 1992.
- [6] C. Carstensen and J. Gedicke, An oscillation-free adaptive FEM for symmetric eigenvalue problems, preprint, 2008.
- [7] C. Carstensen and R. H. W. Hoppe. Convergence analysis of an adaptive nonconforming finite element method. *Numer. Math.*, (103):251–266, 2006.
- [8] C. Carstensen and R. H. W. Hoppe. Error reduction and convergence for an adaptive mixed finite element method. *Math. Comp.*, 75(255):1033–1042, 2006.
- [9] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.
- [10] R.G. Durán, C. Padra and R. Rodríguez, A posteriori estimates for the finite element approximation of eigenvalue problems, *Math. Mod. Meth. Appl. Sci.* 13(8):1219-1229, 2003.
- [11] S. Giani. Convergence of Adaptive Finite Element Methods for Elliptic Eigenvalue Problems with Application to Photonic Crystals, PhD Thesis, University of Bath, 2008.
- [12] S. Giani and I.G. Graham, A convergent adaptive method for elliptic eigenvalue problems and numerical experiments, Research Report 14/08, Bath Institute for Complex Systems, 2008. <http://www.bath.ac.uk/math-sci/BICS/>
- [13] W. Hackbusch. *Elliptic Differential Equations*. Springer, 1992.
- [14] HSL archive, <http://hsl.rl.ac.uk/archive/hslarchive.html>
- [15] S. G. Johnson and J. D. Joannopoulos, Block-iterative frequency-domain methods for Maxwell’s equations in a planewave basis, *Optics Express* 8:173-190, 2001.
- [16] S. G. Johnson and J. D. Joannopoulos, Photonic Crystals. The Road from Theory to Practise, Kluwer Acad. Publ., 2002.
- [17] P. Kuchment, The mathematics of photonic crystals, *SIAM, Frontiers Appl. Math.* 22:207-272, 2001.
- [18] M.G. Larson, A posteriori and a priori analysis for finite element approximations of self-adjoint elliptic eigenvalue problems *SIAM J. Numer. Anal.* 38:608-625, 2000.
- [19] R. B. Lehoucq, D. C. Sorensen, and C. Yang, ARPACK Users’ Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods, SIAM, 1998
- [20] K. Mekchay and R. H. Nochetto, Convergence of adaptive finite element methods for general second order linear elliptic pdes. *SIAM J. Numer. Anal.*, 43:1803–1827, 2005.
- [21] W. Mitchell, Optimal multilevel iterative methods for adaptive grids. *SIAM J. Sci. Stat. Comput.* 13:146-167, 1992.
- [22] P. Morin, R. H. Nochetto and K. G. Siebert. Data oscillation and convergence of adaptive fem. *SIAM J. Numer. Anal.*, 38:466–488, 2000
- [23] P. Morin, R. H. Nochetto and K. G. Siebert. Convergence of adaptive finite element methods *SIAM Review*, 44(4): 631-658, 2002.
- [24] K. Sakoda, Optical Properties of Photonic Crystals, Springer-Verlag, 2001.
- [25] A. Schmidt, K. G. Siebert, ALBERT: An adaptive hierarchical finite element toolbox, Manual, 244 p., Preprint 06/2000 Freiburg.
- [26] J. A. Scott, Sparse Direct Methods: An Introduction. *Lecture Notes in Physics*, 535, 401, 2000
- [27] R. L. Scott, S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math Comp*, 54:483–493,1990.
- [28] R. Stevenson, Optimality of a standard adaptive finite element method, *Found. Comput. Math.* 7(2): 245-269, 2007.
- [29] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, 1973.
- [30] T.F. Walsh, G.M. Reese and U.L. Hetmaniuk, Explicit a posteriori error estimates for eigenvalue analysis of heterogeneous elastic structures *Comput. Methods Appl. Mech. Engrg.* 196: 3614-3623, 2007.