# FUNCTIONAL PERTURBATION RESULTS AND THE BALANCED AFEM ALGORITHM FOR SELF-ADJOINT PDE EIGENVALUE PROBLEMS

AGNIESZKA MIEDLAR*

**Abstract.** We introduce functional perturbation results for PDE eigenvalue problems including the functional backward error and the functional condition number. These results are used to establish a combined a posteriori error estimator embodying the discretization and the approximation error for the simple eigenpair. Based on known perturbation results in $H^1(\Omega)$ and $H^{-1}(\Omega)$ norms and a standard residual a posteriori error estimator, a balancing AFEM algorithm is proposed. The stopping criterion for the eigensolver is based on the equilibrating strategy, i.e., iterations proceed as long as the discrete part of the error estimator dominates the continuous part. All our statements are illustrated with several numerical examples.

**Key words.** functional backward error, functional condition number, self-adjoint eigenvalue problem, adaptive finite element method (AFEM), balanced AFEM algorithm

**AMS subject classifications.** 65F15, 65N15, 65N25, 65N30

**1. Introduction.** Although, the need of designing the adaptive finite element methods concerning the iterative errors was already noticed in [20] not much attention was dedicated to this issue. In [15] fully computable a posteriori error estimates which take into account an inexact solution of the linear algebraic system are derived together with efficient stopping criteria for iterative solvers in the context of a second-order elliptic pure diffusion model boundary value problem discretized in the lowest-order Raviart-Thomas-Nédélec space [5]. The convergence of inexact adaptive finite element solvers for elliptic self-adjoint boundary value problems was recently proved in [1] using the quasi-optimality result introduced in [9]. Moreover, a practical stopping criterion for the Conjugate Gradient method with the Hestenes-Stiefel and the Golub-Meurant error estimates is proposed there.

Most of the existing convergence results for AFEM (Adaptive Finite Element Method) both for boundary and eigenvalue problems are based on perturbation arguments assuming that the a posteriori error estimator evaluated at an approximate solution does not differ much from the estimator evaluated at the exact solution for which the convergence was proved. Moreover, very few results in the literature analyze the impact of the algebraic eigenvalue solver on the total computational cost of the AFEM algorithm or design an appropriate eigensolver stopping criterion. An AFEM algorithm without Galerkin orthogonality is introduced in [13]. The residual type estimator is constructed directly with the inexact solution of the boundary problem and controls an algebraic error in terms of the BPX preconditioner [4].

Research for PDE eigenvalue problems have increased during the last few years, however there are still many challenges. A combined adaptive finite element method with an iterative algebraic eigenvalue of quasi-optimal computational complexity (AFEMES) was introduced in [7]. In [21] the convergence and a quasi-optimality of the inexact inverse iteration coupled with adaptive finite element methods exploiting the well-known results from boundary value problems are studied. In general

* Institut für Mathematik MA 4-5, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, FRG, miedlar@math.tu-berlin.de. Supported by *Berlin Mathematical School* and *Deutsche Forschungsgemeinschaft*, through the DFG Research Center MATHEON *Mathematics for Key Technologies* in Berlin, Germany.

the non-linear nature of eigenvalue problems does not allow to simply extend results from boundary value problems. Also defining the right norms to measure the corresponding errors is more complicated than in the boundary value problem case, where the Conjugate Gradient method naturally minimizes the energy norm of the error. Efficient and reliable adaptive algorithms should take into consideration not only the discretization errors, but also iteration errors and especially for the non-symmetric problems the conditioning of the eigenvalues.

In [2], Arioli et al. introduce *functional backward error and condition number* and the so-called *Compatibility Theorem* for boundary value problems. Functional backward errors and condition numbers are used to analyze the continuous dependence of the inexact solution on the data, in particular to analyze the approximation error and the backward stability of the algebraic eigenvalue problem.

Here, we are interested in applying this theoretical framework to analyze the influence of the accuracy of the algebraic eigenvalue approximation on the adaptivity process.

In [8] the authors proved that the eigenfunction error in the $H^1(\Omega)$-norm is equivalent to the $H^{-1}(\Omega)$-norm of the corresponding residual. However, this result uses the assumption that the corresponding algebraic eigenvalue problem is solved exactly. Using the theoretical framework of [2] we show the relation between the functional backward error and the dual norm of the residual. Our proofs do not require any special assumptions on the inexact eigenpair $(\widetilde{\lambda}, \widetilde{u})$, but we are particularly interested in the eigenpair approximation resulting from an iterative eigensolver applied to the self-adjoint PDE eigenvalue problem discretized with linear finite elements.

Moreover, with the shift-invert Lanczos method and the discrete equivalence of the $H^1(\Omega)$- and $H^{-1}(\Omega)$-norm introduced in [14] we derive a combined residual a posteriori error estimator which allows to incorporate the approximation error into the adaptation process. With this new a posteriori error estimator we present and analyze a new balanced AFEM algorithm which significantly reduces the number of eigensolver iterations.

The paper is organized as follows. We present the theoretical framework of functional perturbation analysis and establish the *Eigenvalue Functional Compatibility Theorem* and functional condition numbers in Section 2. In Section 3 we introduce the combined residual a posteriori error estimator which we use to derive the balanced AFEM algorithm in Section 4. Section 5 presents some numerical results obtained with our algorithm on a model Laplace eigenvalue problem and comparisons with some well established estimators.

## 2. Functional backward error and condition number.

We follow the notation of [2]. Let $H$ be a Hilbert space and $V$ a closed subspace, i.e., $V \subset H$. Let $\mathcal{BL}(V)$, $\mathcal{BL}(H)$ define the space of continuous bilinear forms $V \times V \to \mathbb{R}$ and $H \times H \to \mathbb{R}$, respectively. We first describe the functional backward error analysis and the functional condition number for self-adjoint PDE eigenvalue problems in variational formulation:

*For $a(\cdot,\cdot) \in \mathcal{BL}(V)$ and $b(\cdot,\cdot) \in \mathcal{BL}(H)$ find $\lambda \in \mathbb{R}$ such that there exist $u \in V$, $u \neq 0$ satisfying*

$$a(u,v) = \lambda b(u,v), \ \textit{for all } v \in V. \tag{2.1}$$

For the sake of simplicity, since $V \subset H$ and $b(\cdot,\cdot)$ is also a bilinear form on $V$, we restrict ourself to the space $V$. We assume that $a(\cdot,\cdot)$, $b(\cdot,\cdot)$ are both continuous

(bounded) on $V$ and $a(\cdot,\cdot)$ is coercive ($V$-elliptic), namely

$$\begin{aligned} a(u,v) &\le C_1\|u\|_V\|v\|_V, && \text{for all } u,v\in V, && \text{with some constant } C_1>0,\\ b(u,v) &\le C_2\|u\|_V\|v\|_V, && \text{for all } u,v\in V, && \text{with some constant } C_2>0,\\ a(v,v) &\ge C_3\|v\|_V^2, && \text{for all } v\in V, && \text{with some constant } C_3>0. \end{aligned} \tag{2.2}$$

The space of bounded linear functionals $\ell : V \to \mathbb{R}$ on $V$ is the *dual space* and denoted by $V'$. $V'$ is itself a normed vector space with a *dual norm* defined by

$$\|\ell(v)\|_{V'} := \sup_{v\in V}\frac{|\ell(v)|}{\|v\|_V}, \quad v\neq 0. \tag{2.3}$$

We can also define a norm of the bilinear form $a(u,v)\in\mathcal{BL}(V)$ via

$$\|a(u,v)\|_{\mathcal{BL}(V)} := \sup_{u,v\in V}\frac{|a(u,v)|}{\|u\|_V\|v\|_V}, \quad u,v\neq 0. \tag{2.4}$$

The underlying theoretical framework of the backward error analysis [11] in functional spaces has the following form.

Definition 2.1 (The functional condition number and backward error). *Let $a$, $b\in\mathcal{BL}(V)$ be bilinear forms on $V$. Consider a simple eigenpair $(\lambda,u)$ (i.e., an eigenvalue of algebraic multiplicity one and its corresponding eigenvector) of the problem (2.1) and let $\varphi$ be the mapping*

$$\varphi: \quad (a,b) \quad \to \quad (\lambda,u).$$

*Furthermore, let $(\widetilde{\lambda},\widetilde{u})\in\mathbb{R}\times V$ be an approximation of the eigenpair $(\lambda,u)$. Then the functional condition number of $\varphi$ is*

$$\mathcal{C} = \limsup_{\substack{\delta a\to 0\\ \delta b\to 0}}\frac{\|(\widetilde{\lambda},\widetilde{u})-(\lambda,u)\|_V}{\|(\delta a,\delta b)\|_{\mathcal{BL}(V)}},$$

*and the normwise functional backward error associated with $(\widetilde{\lambda},\widetilde{u})$ is*

$$\eta = \min\{\varepsilon>0;\, \|\delta a\|_{\mathcal{BL}(V)}\le\varepsilon\|a\|_{\mathcal{BL}(V)},\ \|\delta b\|_{\mathcal{BL}(V)}\le\varepsilon\|b\|_{\mathcal{BL}(V)}$$

$$\text{such that} \quad (a+\delta a)(\widetilde{u},v) = \widetilde{\lambda}(b+\delta b)(\widetilde{u},v)\},$$

*where $\delta a,\delta b\in\mathcal{BL}(V)$ are perturbations of the bilinear form $a$ and $b$, respectively, and $\varepsilon\in\mathbb{R}^+$.*

The mapping $\varphi$ describes how the eigenpair $(\lambda,u)$ is determined by the bilinear forms $a,b\in\mathcal{BL}(V)$. Similarly, as in the algebraic eigenvalue problem, the condition number is the smallest upper bound for the ratio between the error in the output, i.e., eigenpair, and the error in the input, i.e., perturbations in both bilinear forms [11]. The normwise functional backward error determines the size of the smallest perturbation in the bilinear forms, such that the approximate eigenpair is the exact eigenpair of the perturbed variational equation. However, here the situation is much more complicated. Since, usually, the approximate eigenpair $(\widetilde{\lambda},\widetilde{u})$ is a finite element

solution obtained after discretizing the PDE $\varphi_{disc}$ and applying the iterative eigensolver $\varphi_{iter}$ to the finite dimensional problem, the mapping $\varphi$ in fact is a composite of two maps, i.e.,

$$\varphi((a,b)) = \varphi_{iter} \circ \varphi_{disc}((a,b)) = (\lambda, u).$$

Applying the perturbation arguments leads to

$$\varphi((a+\delta a, b+\delta b)) = \varphi_{iter} \circ \varphi_{disc}((a+\delta a, b+\delta b)) = (\widetilde{\lambda}, \widetilde{u}).$$

Therefore, the error between the exact continuous eigenpair $(\lambda, u)$ and its finite element approximation $(\widetilde{\lambda}, \widetilde{u})$ can be written as

$$\|(\widetilde{\lambda}, \widetilde{u}) - (\lambda, u)\|_V = \mathcal{C}_{iter}\varphi_{iter}\Big(\mathcal{C}_{disc}\varphi_{disc}((a+\delta a, b+\delta b)) + \epsilon\Big).$$

In the ideal situation, when the mapping $\varphi$ is linear, obtaining the condition number and the backward error which combine information from the discretization and the iteration process would be relatively easy. However, eigenvalue problems are nonlinear, i.e., $\varphi$ is a nonlinear mapping and the condition number of the problem does not depend linearly on the two easily determined condition numbers $\mathcal{C}_{disc}$ and $\mathcal{C}_{iter}$. Therefore, it seems natural to analyze the variational formulation without any assumptions about the exactness of the finite element solution.
As a starting point let us determine the *normwise functional backward error* as a generalization of [3, Theorem 3.1].

THEOREM 2.2. *Let $a$, $b$ be bilinear forms on $V$, i.e., $a, b \in \mathcal{BL}(\mathcal{V})$. Furthermore, let $\delta a \in \mathcal{BL}(\mathcal{V})$ and $\delta b \in \mathcal{BL}(\mathcal{V})$ be perturbations of the bilinear forms $a, b \in \mathcal{BL}(V)$, respectively, such that*

$$\|\delta a\|_{\mathcal{BL}(\mathcal{V})} \leq \varepsilon \|a\|_{\mathcal{BL}(\mathcal{V})}, \quad \|\delta b\|_{\mathcal{BL}(\mathcal{V})} \leq \varepsilon \|b\|_{\mathcal{BL}(\mathcal{V})},$$

*with $\varepsilon \in \mathbb{R}^+$. Then, the normwise functional backward error associated with the approximate eigenpair $(\widetilde{\lambda}, \widetilde{u})$ of (2.1) is*

$$\eta = \frac{\|R(v)\|_{V'}}{(|\widetilde{\lambda}|\|b\|_{\mathcal{BL}(V)} + \|a\|_{\mathcal{BL}(V)})\|\widetilde{u}\|_V}, \tag{2.5}$$

*with the residual $R(v) = a(\widetilde{u}, v) - \widetilde{\lambda} b(\widetilde{u}, v) \in V'$.*

*Proof.* From the backward error analysis we have the following identity

$$(a+\delta a)(\widetilde{u}, v) = \widetilde{\lambda}(b+\delta b)(\widetilde{u}, v),$$

which corresponds to

$$R(v) = -\delta a(\widetilde{u}, v) + \widetilde{\lambda}\delta b(\widetilde{u}, v).$$

Moreover, for all $v \in V$, the following relation holds

$$|R(v)| = |-\delta a(\widetilde{u}, v) + \widetilde{\lambda}\delta b(\widetilde{u}, v)| \leq \|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V\|v\|_V + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V\|v\|_V.$$

Hence, we have

$$\frac{|R(v)|}{\|v\|_V} \leq \|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V,$$

for all $v \in V$.
Therefore, it is obvious that

$$\sup_{v\in V\setminus\{0\}} \frac{|R(v)|}{\|v\|_V} \leq \|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V.$$

We note that from the definition of the dual norm (2.3), it follows that

$$\|R(v)\|_{V'} \leq \|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V.$$

With the assumptions on $\|\delta a\|_{\mathcal{BL}(V)}$, $\|\delta b\|_{\mathcal{BL}(V)}$ we have that

$$\|R(v)\|_{V'} \leq \varepsilon\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \varepsilon|\widetilde{\lambda}|\|b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V,$$

and therefore,

$$\frac{\|R(v)\|_{V'}}{\big(|\widetilde{\lambda}|\|b\|_{\mathcal{BL}(V)} + \|a\|_{\mathcal{BL}(V)}\big)\|\widetilde{u}\|_V} \leq \varepsilon.$$

Since this inequality holds for an arbitrarily chosen $\varepsilon$, it holds also for $\eta$, i.e.,

$$\frac{\|R(v)\|_{V'}}{\big(|\widetilde{\lambda}|\|b\|_{\mathcal{BL}(V)} + \|a\|_{\mathcal{BL}(V)}\big)\|\widetilde{u}\|_V} \leq \eta.$$

Let $\delta a$, $\delta b$ be defined as follows (see, e.g., [2, 3])

$$\delta a(u,v) = -\frac{\|a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \|b\|_{\mathcal{BL}(V)}\widetilde{\lambda}\|\widetilde{u}\|_V}\|u\|_V R(v)$$

and

$$\delta b(u,v) = \frac{\|b\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \|b\|_{\mathcal{BL}(V)}\widetilde{\lambda}\|\widetilde{u}\|_V}\|u\|_V R(v).$$

Since for the smallest eigenvalue $\widetilde{\lambda} > 0$, we have from the definition of the norm of the bilinear form (2.4) and the dual norm (2.3) that

$$\begin{aligned}\|\delta a\|_{\mathcal{BL}(V)} &= \sup_{u,v\in V}\frac{|\delta a(u,v)|}{\|u\|_V\|v\|_V} = \frac{\|a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \|b\|_{\mathcal{BL}(V)}|\widetilde{\lambda}|\|\widetilde{u}\|_V}\sup_{v\in V}\frac{|R(v)|}{\|v\|_V}\\ &= \frac{\|a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \|a\|_{\mathcal{BL}(V)}\widetilde{\lambda}\|\widetilde{u}\|_V}\|R(v)\|_{V'}\end{aligned}$$

and

$$\begin{aligned}\|\delta b\|_{\mathcal{BL}(V)} &= \sup_{u,v\in V}\frac{|\delta b(u,v)|}{\|u\|_V\|v\|_V} = \frac{\|b\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \|b\|_{\mathcal{BL}(V)}|\widetilde{\lambda}|\|\widetilde{u}\|_V}\sup_{v\in V}\frac{|R(v)|}{\|v\|_V}\\ &= \frac{\|b\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \|b\|_{\mathcal{BL}(V)}\widetilde{\lambda}\|\widetilde{u}\|_V}\|R(v)\|_{V'}.\end{aligned}$$

Therefore, the following inequalities hold

$$\|\delta a\|_{\mathcal{BL}(V)} \leq \varepsilon \|a\|_{\mathcal{BL}(V)} \quad \text{and} \quad \|\delta b\|_{\mathcal{BL}(V)} \leq \varepsilon \|b\|_{\mathcal{BL}(V)},$$

with

$$\varepsilon = \frac{\|R(v)\|_{V'}}{\big(|\widetilde{\lambda}|\|b\|_{\mathcal{BL}(V)} + \|a\|_{\mathcal{BL}(V)}\big)\|\widetilde{u}\|_V}.$$

Moreover,

$$\begin{aligned}
(a+\delta a)(\widetilde{u},v) &= a(\widetilde{u},v) + \delta a(\widetilde{u},v) \\
&= a(\widetilde{u},v) - \frac{\|a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + \widetilde{\lambda}\|b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V}\|\widetilde{u}\|_V R(v) \\
&= R(v) + \widetilde{\lambda} b(\widetilde{u},v) - \frac{\|a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\| + \widetilde{\lambda}\|b\|_{\mathcal{BL}(V)}} R(v) \\
&= \widetilde{\lambda} b(\widetilde{u},v) + \frac{R(v)\|a\|_{\mathcal{BL}(V)}\| + R(v)\widetilde{\lambda}\|b\|_{\mathcal{BL}(V)}\| - R(v)\|a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\| + \widetilde{\lambda}\|b\|_{\mathcal{BL}(V)}} \\
&= \widetilde{\lambda} b(\widetilde{u},v) + \frac{R(v)\widetilde{\lambda}\|b\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}\| + \widetilde{\lambda}\|b\|_{\mathcal{BL}(V)}} \\
&= \widetilde{\lambda} b(\widetilde{u},v) + \widetilde{\lambda}\delta b(\widetilde{u},v).
\end{aligned}$$

Thus,

$$(a+\delta a)(\widetilde{u},v) = \widetilde{\lambda}(b+\delta b)(\widetilde{u},v)$$

and

$$\eta \leq \varepsilon = \frac{\|R(v)\|_{V'}}{\big(|\widetilde{\lambda}|\|b\|_{\mathcal{BL}(V)} + \|a\|_{\mathcal{BL}(V)}\big)\|\widetilde{u}\|_V}.$$

□

Following [2], with the functional backward error analysis, we prove the existence of perturbations $\delta a$ and $\delta b$ such that the approximate eigenpair $(\widetilde{\lambda},\widetilde{u})$ of (2.1) is the exact solution of the same variational equation perturbed with $\delta a$ and $\delta b$. This result can be stated as the *Eigenvalue Functional Compatibility Theorem.*

THEOREM 2.3 (Eigenvalue Functional Compatibility Theorem). *Let $(\widetilde{\lambda},\widetilde{u})$ be an approximate eigenpair of the variational formulation (2.1) and let $R(v)$ the corresponding residual, i.e., $R(v) = a(\widetilde{u},v) - \widetilde{\lambda} b(\widetilde{u},v)$. Then the following are equivalent:*

*(i) There exist $\delta a \in \mathcal{BL}(\mathcal{V})$ and $\delta b \in \mathcal{BL}(\mathcal{V})$ such that*

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

*with*

$$\|\delta a\|_{\mathcal{BL}(\mathcal{V})} \leq \alpha, \quad \|\delta b\|_{\mathcal{BL}(\mathcal{V})} \leq \beta,$$

*(ii)*

$$\|R(v)\|_{V'} \leq \alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V, \tag{2.6}$$

*with* $\alpha, \beta \in \mathbb{R}^+$.

*Proof.* Following the proof of Theorem 2.2 we get

$$\|R(v)\|_{V'} \leq \|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V,$$

which together with the assumptions on $\|\delta a\|_{\mathcal{BL}(V)}$, $\|\delta b\|_{\mathcal{BL}(V)}$ yields (2.6).
Set

$$\delta a(u,v) = -\frac{\alpha}{\alpha\|\widetilde{u}\|_V + \beta\widetilde{\lambda}\|\widetilde{u}\|_V}\|u\|_V R(v)$$

and

$$\delta b(u,v) = \frac{\beta}{\alpha\|\widetilde{u}\|_V + \beta\widetilde{\lambda}\|\widetilde{u}\|_V}\|u\|_V R(v).$$

Then we have

$$\begin{aligned}
-\delta a(\widetilde{u},v) + \widetilde{\lambda}\delta b(\widetilde{u},v) &= \frac{\alpha}{\alpha\|\widetilde{u}\|_V + \beta\widetilde{\lambda}\|\widetilde{u}\|_V}\|\widetilde{u}\|_V R(v) \\
&\quad + \widetilde{\lambda}\frac{\beta}{\alpha\|\widetilde{u}\|_V + \beta\widetilde{\lambda}\|\widetilde{u}\|_V}\|\widetilde{u}\|_V R(v) \\
&= \Big(\frac{\alpha\|\widetilde{u}\|}{\alpha\|\widetilde{u}\|_V + \beta\widetilde{\lambda}\|\widetilde{u}\|_V} + \frac{\beta\widetilde{\lambda}\|\widetilde{u}\|}{\alpha\|\widetilde{u}\|_V + \beta\widetilde{\lambda}\|\widetilde{u}\|_V}\Big)R(v),
\end{aligned}$$

and therefore

$$-\delta a(\widetilde{u},v) + \widetilde{\lambda}\delta b(\widetilde{u},v) = R(v).$$

Assuming $\|R(v)\|_{V'} \leq \alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V$ and employing the definition of the norm of the bilinear form (2.4) and of the dual norm (2.3) we get

$$\begin{aligned}
\|\delta a\|_{\mathcal{BL}(V)} &= \sup_{u,v\in V}\frac{|\delta a(u,v)|}{\|u\|_V\|v\|_V} \leq \frac{\alpha}{\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V}\sup_{v\in V}\frac{|R(v)|}{\|v\|_V} \\
&= \frac{\alpha}{\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V}\|R(v)\|_{V'} \\
&\leq \frac{\alpha}{\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V}\big(\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V\big)
\end{aligned}$$

and

$$\begin{aligned}
\|\delta b\|_{\mathcal{BL}(V)} &= \sup_{u,v\in V}\frac{|\delta b(u,v)|}{\|u\|_V\|v\|_V} \leq \frac{\beta}{\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V}\sup_{v\in V}\frac{|R(v)|}{\|v\|_V} \\
&= \frac{\beta}{\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V}\|R(v)\|_{V'} \\
&\leq \frac{\beta}{\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V}\big(\alpha\|\widetilde{u}\|_V + \beta|\widetilde{\lambda}|\|\widetilde{u}\|_V\big).
\end{aligned}$$

Therefore,

$$\|\delta a\|_{\mathcal{BL}(V)} \leq \alpha \quad \text{and} \quad \|\delta b\|_{\mathcal{BL}(V)} \leq \beta,$$

which completes the proof. □

The normwise functional backward error determines the size of the smallest perturbation such that the approximate eigenpair is an exact solution of the perturbed variational equation. It provides a posteriori information which is dependent on the approximate solution and thus on the choice of the solution method. On the other hand, the condition number allows to analyze the sensitivity of the eigenpairs to the particular perturbations in the input data [11]. Together, they provide a possibility to estimate the total error, i.e., the forward error of the problem. In the context of finite element methods, the crucial element is to reduce the number of controlled parameters. Since, the discretization process and the iteration process interact, we would like to eliminate their influence whenever it is possible. The condition number can be viewed as a priori information about the problem which describes a property of the problem and we would like it to be independent of the applied solution method and in particular independent of the discretization parameters as for boundary value problems [2].

Definition 2.4 (Condition number of the elliptic eigenvalue problem). *Let $\delta a \in \mathcal{BL}(V)$, $\delta b \in \mathcal{BL}(V)$ be perturbations of bilinear forms $a, b \in \mathcal{BL}(V)$ defined as in the variational formulation (2.1), respectively, such that*

$$\|\delta a\|_{\mathcal{BL}(V)} \leq \varepsilon\alpha, \qquad \|\delta b\|_{\mathcal{BL}(V)} \leq \varepsilon\beta, \qquad \textit{with} \;\; \varepsilon, \alpha, \beta \in \mathbb{R}^+.$$

*The relative functional eigenvalue and eigenfunction condition number, $\mathcal{C}(\mathcal{EVP}_\lambda)$, $\mathcal{C}(\mathcal{EVP}_u)$, for the variational problem (2.1), are the smallest constants $C_\lambda$, $C_u$ for which the inequalities*

$$|\lambda - \widetilde{\lambda}| \leq \varepsilon C_\lambda |\lambda|,$$

$$\|u - \widetilde{u}\|_{V'} \leq \varepsilon C_u \|u\|_V,$$

*are satisfied. Moreover,*

$$\sin_V \angle(\widetilde{u}, u) \leq \varepsilon C_u.$$

Theorem 2.5 (The functional eigenvalue condition number $\mathcal{C}(\mathcal{EVP}_\lambda)$). *Let $\delta a \in \mathcal{BL}(V)$, $\delta b \in \mathcal{BL}(V)$ be perturbations of bilinear forms $a, b \in \mathcal{BL}(V)$ defined as in the variational formulation (2.1), respectively, such that*

$$\|\delta a\|_{\mathcal{BL}(V)} \leq \varepsilon\alpha, \qquad \|\delta b\|_{\mathcal{BL}(V)} \leq \varepsilon\beta,$$

*with $\varepsilon, \alpha, \beta \in \mathbb{R}^+$. Then,*

$$|\lambda - \widetilde{\lambda}| \leq \varepsilon(|\lambda|\beta + \alpha) + \;\; h.o.t.,$$

*where by h.o.t we denote the higher order terms, which here are given by* $\mathcal{O}((\varepsilon\beta)^2) + \mathcal{O}(\angle(\widetilde{u},u)^2)$.
*Moreover, the functional eigenvalue condition number* $\mathcal{C}(\mathcal{EVP}_\lambda)$ *satisfies the bound*

$$\mathcal{C}(\mathcal{EVP}_\lambda) \leq \frac{|\lambda|\beta+\alpha}{|\lambda|} + h.o.t..$$

*Proof.* The main idea of the proof is to estimate the eigenvalue error without assuming that the approximate eigenpair $(\widetilde{\lambda},\widetilde{u})$ fulfills the weak formulation (2.1). First, we assume that the approximate eigenpair $(\widetilde{\lambda},\widetilde{u})$ satisfy the equation

$$(a+\delta a)(\widetilde{u},v) = \widetilde{\lambda}(b+\delta b)(\widetilde{u},v),$$

which means that it fulfills the modified variational equation

$$\widetilde{a}(\widetilde{u},v) = \widetilde{\lambda}\widetilde{b}(\widetilde{u},v).$$

Now, due to the linearity of both bilinear forms we get

$$a(\widetilde{u},v) + \delta a(\widetilde{u},v) = \widetilde{\lambda}b(\widetilde{u},v) + \widetilde{\lambda}\delta b(\widetilde{u},v).$$

The essential observation is that $v$ can be chosen arbitrarily in $V$, so let $v=u$ then

$$a(\widetilde{u},u) + \delta a(\widetilde{u},u) = \widetilde{\lambda}b(\widetilde{u},u) + \widetilde{\lambda}\delta b(\widetilde{u},u).$$

With $a(\widetilde{u},u) = \lambda b(\widetilde{u},u)$ and writing $\widetilde{\lambda}$ as $(\lambda-(\lambda-\widetilde{\lambda}))$ we obtain that

$$(\lambda-\widetilde{\lambda})b(\widetilde{u},u) = (\lambda-(\lambda-\widetilde{\lambda}))\delta b(\widetilde{u},u) - \delta a(\widetilde{u},u)$$

and

$$|\lambda-\widetilde{\lambda}||b(\widetilde{u},u)| = |(\lambda-(\lambda-\widetilde{\lambda}))\delta b(\widetilde{u},u) - \delta a(\widetilde{u},u)|.$$

After applying the triangle inequality we get

$$|\lambda-\widetilde{\lambda}||b(\widetilde{u},u)| \leq |\lambda|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V\|u\|_V + |\lambda-\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V\|u\|_V + \|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V\|u\|_V.$$

As, $\|\delta a\|_{\mathcal{BL}(V)} \leq \varepsilon\alpha$ and $\|\delta b\|_{\mathcal{BL}(V)} \leq \varepsilon\beta$, we have

$$|\lambda-\widetilde{\lambda}||b(\widetilde{u},u)| \leq \varepsilon\beta|\lambda|\|\widetilde{u}\|_V\|u\|_V + \varepsilon\beta|\lambda-\widetilde{\lambda}|\|\widetilde{u}\|_V\|u\|_V + \varepsilon\alpha\|\widetilde{u}\|_V\|u\|_V.$$

Since $|b(\widetilde{u},u)| \neq 0$, the last inequality yields

$$|\lambda-\widetilde{\lambda}| \leq (\varepsilon\beta|\lambda| + \varepsilon\beta|\lambda-\widetilde{\lambda}| + \varepsilon\alpha)\frac{\|\widetilde{u}\|_V\|u|_V}{|b(\widetilde{u},u)|}.$$

Following the definition of the angle between functions, it is straightforward that

$$|\lambda-\widetilde{\lambda}| \leq (\varepsilon\beta|\lambda| + \varepsilon\beta|\lambda-\widetilde{\lambda}| + \varepsilon\alpha)\frac{1}{\cos\angle(\widetilde{u},u)}.$$

We are now in the position to apply Wilkinson's first order perturbation result [19], namely

$$|\lambda - \widetilde{\lambda}| \leq (|\lambda|\varepsilon\beta + |\lambda - \widetilde{\lambda}|\varepsilon\beta + \varepsilon\alpha) + \mathcal{O}(\angle(\widetilde{u}, u)^2).$$

If $\varepsilon\beta < 1$ then

$$|\lambda - \widetilde{\lambda}| \leq \varepsilon(1 - \varepsilon\beta)^{-1}(|\lambda|\beta + \alpha) + \mathcal{O}(\angle(\widetilde{u}, u)^2).$$

After neglecting higher order terms in the Taylor expansion of $(1 - \varepsilon\beta)^{-1}$, we finally get

$$|\lambda - \widetilde{\lambda}| \leq \varepsilon(|\lambda|\beta + \alpha) + \mathcal{O}((\varepsilon\beta)^2) + \mathcal{O}(\angle(\widetilde{u}, u)^2).$$

□

REMARK 2.6. *We note that if* $\delta b = 0$ *then*

$$|\lambda - \widetilde{\lambda}| \leq \frac{\|\delta a\|_{\mathcal{BL}(V)}\|\widetilde{u}\|_V\|u\|_V}{|b(\widetilde{u}, u)|} = \frac{\|\delta a\|_{\mathcal{BL}(V)}}{\cos\angle(\widetilde{u}, u)}$$

*and the relative eigenvalue error is given by*

$$\begin{aligned}\frac{|\lambda - \widetilde{\lambda}|}{|\lambda|} &\leq \frac{\|\delta a\|_{\mathcal{BL}(V)}}{|\lambda|\cos\angle(\widetilde{u}, u)} \leq \frac{\|\delta a\|_{\mathcal{BL}(V)}}{|\lambda|}\sec\angle(\widetilde{u}, u) = \frac{\|\delta a\|_{\mathcal{BL}(V)}}{|\lambda|} + \mathcal{O}(\angle(\widetilde{u}, u)^2) \\ &\leq \underbrace{\frac{\|a\|_{\mathcal{BL}(V)}}{|\lambda|}}_{\substack{\textit{relative eigenvalue}\\ \textit{condition number}}} \underbrace{\frac{\|\delta a\|_{\mathcal{BL}(V)}}{\|a\|_{\mathcal{BL}(V)}}}_{\textit{backward error}} + \mathcal{O}(\angle(\widetilde{u}, u)^2).\end{aligned}$$

*The essential observation is that the last inequality is a direct infinite dimensional analogue of the well-known result [3, Section 3.4] or [10].*

THEOREM 2.7 (The functional eigenfunction condition number $\mathcal{C}(\mathcal{EVP}_u)$). *Let* $(\lambda, u)$ *be an exact eigenpair of the variational formulation (2.1) and let* $(\widetilde{\lambda}, \widetilde{u})$ *be its approximation. Let bilinear forms* $a,\ b \in \mathcal{BL}(V)$ *satisfy (2.2) and let* $\delta a, \delta b \in \mathcal{BL}(V)$ *be perturbations of* $a, b$*, respectively, such that*

$$\|\delta a\|_{\mathcal{BL}(V)} \leq \varepsilon\alpha, \qquad \|\delta b\|_{\mathcal{BL}(V)} \leq \varepsilon\beta,$$

*with* $\varepsilon, \alpha, \delta \in \mathbb{R}^+$*. Let* $C_2, C_3$ *be the continuity and the coercivity constant as in (2.2) and* $C = (|\lambda|\beta + \alpha)(C_2 + \varepsilon\beta + 1)$*. If* $C_3 > \varepsilon C + C_2|\lambda|$ *then*

$$\|u - \widetilde{u}\|_V \leq \varepsilon\Big(1 - \frac{\varepsilon C + |\lambda|C_2}{C_3}\Big)^{-1}\frac{1}{C_3}C\|u\|_V.$$

*The functional eigenfunction condition number* $\mathcal{C}(\mathcal{EVP}_u)$ *satisfies the bound*

$$\mathcal{C}(\mathcal{EVP}_u) \leq \frac{C}{C_3} = \frac{(|\lambda|\beta + \alpha)(C_2 + \varepsilon\beta + 1)}{C_3}.$$

*Proof.* Let $(\widetilde{\lambda}, \widetilde{u})$ be the approximate solution of the variational equation (2.1). The inexact eigenpair satisfies a perturbed variational equation

$$(a + \delta a)(\widetilde{u}, v) = \widetilde{\lambda}(b + \delta b)(\widetilde{u}, v),$$

which can be written as

$$a(\widetilde{u}, v) + \delta a(\widetilde{u}, v) = \widetilde{\lambda} b(\widetilde{u}, v) + \widetilde{\lambda}\delta b(\widetilde{u}, v).$$

Setting $\widetilde{u} = u - (u - \widetilde{u})$ implies that

$$a(u - (u - \widetilde{u}), v) + \delta a(u - (u - \widetilde{u}), v) = \widetilde{\lambda} b(u - (u - \widetilde{u}), v) + \widetilde{\lambda}\delta b(u - (u - \widetilde{u}), v).$$

Then

$$a(u, v) - a(u - \widetilde{u}, v) + \delta a(u, v) - \delta a(u - \widetilde{u}, v) = \widetilde{\lambda} b(u, v) - \widetilde{\lambda} b(u - \widetilde{u}, v) + \widetilde{\lambda}\delta b(u, v) - \widetilde{\lambda}\delta b(u - \widetilde{u}, v).$$

Since $a(u, v) = \lambda b(u, v)$, we have

$$\lambda b(u, v) - a(u - \widetilde{u}, v) + \delta a(u, v) - \delta a(u - \widetilde{u}, v) = \widetilde{\lambda} b(u, v) - \widetilde{\lambda} b(u - \widetilde{u}, v) + \widetilde{\lambda}\delta b(u, v) - \widetilde{\lambda}\delta b(u - \widetilde{u}, v).$$

With the choice $v = u - \widetilde{u}$, we get

$$\begin{aligned} -a(u - \widetilde{u}, u - \widetilde{u}) = {} & \widetilde{\lambda} b(u, u - \widetilde{u}) - \lambda b(u, u - \widetilde{u}) - \widetilde{\lambda} b(u - \widetilde{u}, u - \widetilde{u}) + \widetilde{\lambda}\delta b(u, u - \widetilde{u}) \\ & - \widetilde{\lambda}\delta b(u - \widetilde{u}, u - \widetilde{u}) - \delta a(u, u - \widetilde{u}) + \delta a(u - \widetilde{u}, u - \widetilde{u}). \end{aligned}$$

With the coercivity condition (2.2), we can show that

$$\begin{aligned} C_3\|u - \widetilde{u}\|_V^2 \leq {} & |\widetilde{\lambda} - \lambda|\|b\|_{\mathcal{BL}(V)}\|u\|_V\|u - \widetilde{u}\|_V + |\widetilde{\lambda}|_V\|b\|_{\mathcal{BL}(V)}\|u - \widetilde{u}\|_V^2 \\ & + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|u\|_V\|u - \widetilde{u}\|_V + |\widetilde{\lambda}|\|\delta b\|_{\mathcal{BL}(V)}\|u - \widetilde{u}\|_V^2 \\ & + \|\delta a\|_{\mathcal{BL}(V)}\|u\|_V\|u - \widetilde{u}\|_V + \|\delta a\|_{\mathcal{BL}(V)}\|u - \widetilde{u}\|_V^2. \end{aligned}$$

The boundness of the bilinear form $b$ (2.2) and the assumptions on the size of perturbations $\|\delta a\|_{\mathcal{BL}(V)}$, $\|\delta b\|_{\mathcal{BL}(V)}$, yield

$$\begin{aligned} C_3\|u - \widetilde{u}\|_V^2 \leq {} & |\widetilde{\lambda} - \lambda|C_2\|u\|_V\|u - \widetilde{u}\|_V + |\widetilde{\lambda}|C_2\|u - \widetilde{u}\|_V^2 \\ & + |\widetilde{\lambda}|\varepsilon\beta\|u\|_V\|u - \widetilde{u}\|_V + |\widetilde{\lambda}|\varepsilon\beta\|u - \widetilde{u}\|_V^2 \\ & + \varepsilon\alpha\|u\|_V\|u - \widetilde{u}\|_V + \varepsilon\alpha\|u - \widetilde{u}\|_V^2. \end{aligned}$$

Since $\|u - \widetilde{u}\|_V \neq 0$, we get

$$\begin{aligned} C_3\|u - \widetilde{u}\|_V \leq {} & |\widetilde{\lambda} - \lambda|C_2\|u\|_V + |\widetilde{\lambda}|C_2\|u - \widetilde{u}\|_V \\ & + |\widetilde{\lambda}|\varepsilon\beta\|u\|_V + |\widetilde{\lambda}|\varepsilon\beta\|u - \widetilde{u}\|_V + \varepsilon\alpha\|u\|_V + \varepsilon\alpha\|u - \widetilde{u}\|_V. \end{aligned}$$

By setting $\widetilde{\lambda} = \lambda - (\lambda - \widetilde{\lambda})$, we obtain

$$\begin{aligned} C_3\|u-\widetilde{u}\|_V \leq |\widetilde{\lambda}-\lambda|C_2\|u\|_V + |\lambda-(\lambda-\widetilde{\lambda})|C_2\|u-\widetilde{u}\|_V \\ + |\lambda-(\lambda-\widetilde{\lambda})|\varepsilon\beta\|u\|_V + |\lambda-(\lambda-\widetilde{\lambda})|\varepsilon\beta\|u-\widetilde{u}\|_V \\ + \varepsilon\alpha\|u\|_V + \varepsilon\alpha\|u-\widetilde{u}\|_V, \end{aligned}$$

and

$$\begin{aligned} C_3\|u-\widetilde{u}\|_V \leq |\widetilde{\lambda}-\lambda|(C_2\|u\|_V + C_2\|u-\widetilde{u}\|_V + \varepsilon\beta\|u\| + \varepsilon\beta\|u-\widetilde{u}\|_V) \\ + |\lambda|C_2\|u-\widetilde{u}\|_V + |\lambda|\varepsilon\beta\|u\|_V + |\lambda|\varepsilon\beta\|u-\widetilde{u}\|_V \\ + \varepsilon\alpha\|u\| + \varepsilon\alpha\|u-\widetilde{u}\|_V. \end{aligned}$$

Using the estimate of the eigenvalue error from Theorem 2.5, we arrive at

$$\begin{aligned} C_3\|u-\widetilde{u}\|_V \leq \Big(\varepsilon(|\lambda|\beta+\alpha)\Big)\Big(C_2\|u\|_V + C_2\|u-\widetilde{u}\|_V + \varepsilon\beta\|u\|_V + \varepsilon\beta\|u-\widetilde{u}\|_V\Big) \\ + |\lambda|C_2\|u-\widetilde{u}\|_V + |\lambda|\varepsilon\beta\|u\|_V + |\lambda|\varepsilon\beta\|u-\widetilde{u}\|_V \\ + \varepsilon\alpha\|u\|_V + \varepsilon\alpha\|u-\widetilde{u}\|_V, \end{aligned}$$

which implies that

$$\begin{aligned} C_3\|u-\widetilde{u}\|_V \leq \varepsilon\Big(|\lambda|\beta+\alpha\Big)\Big(C_2+\varepsilon\beta+1\Big)\|u\|_V \\ + \varepsilon\Big(|\lambda|\beta+\alpha\Big)\Big(C_2+\varepsilon\beta+1\Big)\|u-\widetilde{u}\|_V + |\lambda|C_2\|u-\widetilde{u}\|_V. \end{aligned} \tag{2.7}$$

Let now $C = \Big(|\lambda|\beta+\alpha\Big)\Big(C_2+\varepsilon\beta+1\Big)$, then (2.7) reads

$$C_3\|u-\widetilde{u}\|_V \leq \varepsilon C\|u\|_V + \varepsilon C\|u-\widetilde{u}\|_V + |\lambda|C_2\|u-\widetilde{u}\|_V$$

and the inequality

$$\|u-\widetilde{u}\|_V(C_3 - \varepsilon C - |\lambda|C_2) \leq \varepsilon C_3\|u\|_V,$$

holds. If $(C_3 > \varepsilon C + |\lambda|C_2)$, then we have that

$$\|u-\widetilde{u}\|_V \leq \frac{1}{(C_3 - \varepsilon C - |\lambda|C_2)}\varepsilon C\|u\|_V$$

and readily

$$\|u-\widetilde{u}\|_V \leq \varepsilon\Big(1 - \frac{\varepsilon C + |\lambda|C_2}{C_3}\Big)^{-1}\frac{C}{C_3}\|u\|_V.$$

Therefore, the functional eigenfunction condition number $\mathcal{C}(\mathcal{EVP}_u)$ satisfies

$$\mathcal{C}(\mathcal{EVP}_u) \leq \frac{C}{C_3} = \frac{(|\lambda|\beta+\alpha)(C_2+\varepsilon\beta+1)}{C_3}.$$

□

**3. A combined residual a posteriori error estimator for self-adjoint eigenvalue problems.** Let us consider the Laplace eigenvalue problem:
*Determine a non-trivial eigenpair* $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$,

$$-\Delta u = \lambda u \quad in \quad \Omega, \quad u = 0 \quad on \quad \partial\Omega, \tag{3.1}$$

in its variational form (2.1) with $V := H_0^1(\Omega)$.

Let $(\lambda_h, u_h) \in \mathbb{R} \times V_h$ be the exact discrete eigenpair (the Galerkin solution), such that

$$a(u_h, v_h) = \lambda_h b(u_h, v_h) \quad \text{for all} \quad v_h \in V_h \subset V.$$

For the case of the Galerkin solution $(\lambda_h, u_h) \in \mathbb{R} \times V_h$ the equivalence between the discretization error $\|u - u_h\|_{H_0^1(\Omega)}$ and the dual norm of the residual, i.e., $\|R(v)\|_{H^{-1}(\Omega)} = \|a(u_h, v) - \lambda_h b(u_h, v)\|_{H^{-1}(\Omega)}$, up to the higher-order terms, is a well-known relation [8, 12], which allows to derive the standard residual error estimator $\eta(\lambda_h, u_h)$, i.e.,

$$\eta(\lambda_h, u_h) \lesssim \|u - u_h\|_{H_0^1(\Omega)} \lesssim \eta(\lambda_h, u_h),$$

with the norm equivalence

$$c\|u - u_h\|_{H^1(\Omega)} \leq \|u - u_h\|_{H_0^1(\Omega)} \leq C\|u - u_h\|_{H^1(\Omega)},$$

for some constants $c, C > 0$. Here $x \lesssim y$ means inequality up to constant independent on the mesh-size $h$.

If we consider the inexact finite element solution $(\widetilde{\lambda}_h, \widetilde{u}_h) \in \mathbb{R} \times V_h$, the corresponding equivalence relation between the complete error $\|u - \widetilde{u}_h\|_{H_0^1(\Omega)}$ and the dual norm of the residual, i.e., $\|R(v)\|_{H^{-1}(\Omega)} = \|a(\widetilde{u}_h, v) - \widetilde{\lambda}_h b(\widetilde{u}_h, v)\|_{H^{-1}(\Omega)}$ holds by choosing $\varepsilon = \eta$ in Theorem 2.7. However, the equivalence between the dual norm of the residual, i.e., $\|R(v)\|_{H^{-1}(\Omega)} = \|a(\widetilde{u}_h, v) - \widetilde{\lambda}_h b(\widetilde{u}_h, v)\|_{H^{-1}(\Omega)}$ and the standard residual error estimator evaluated at the inexact eigenpair $(\widetilde{\lambda}_h, \widetilde{u}_h)$, i.e., $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ does not follow and is hard to prove.

Nevertheless, it is still possible to control the adaptive finite element method with the standard residual error estimator calculated with a non-Galerkin solution $(\widetilde{\lambda}_h, \widetilde{u}_h)$ if the influence of the approximation error $\|u_h - \widetilde{u}_h\|_{H_0^1(\Omega)}$ can be analyzed and combined into the error bounds. In [13] a combined a posteriori error estimator for a non-Galerkin solution of the boundary value problem was designed where the $H^{-1}(\Omega)$-norm of the residual is split directly into the discrete part, the BPX preconditioner [4] of the algebraic residual and the continuous part, i.e., the standard residual error estimator $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$. The first part of the estimator controls the iteration error and reflects the influence of the quality of the approximate solution on the estimator, where the latter term measures the size of the discretization error.
Let us consider the following bound on the global eigenfunction error

$$\|u - \widetilde{u}_h\|_{H_0^1(\Omega)} \leq \|u - u_h\|_{H_0^1(\Omega)} + \|u_h - \widetilde{u}_h\|_{H_0^1(\Omega)},$$

where $\|u - u_h\|_{H_0^1(\Omega)}$ is the discretization error and $\|u_h - \widetilde{u}_h\|_{H_0^1(\Omega)}$ the error in the solution of the algebraic eigenvalue problem. We now exploit the idea of the combined a posteriori error estimator introduced in [13] and bound each term independently.

In order to control the discretization error, we can use the standard residual error estimator $\eta(\lambda_h, u_h)$, i.e.,

$$\eta(\lambda_h, u_h) \lesssim \|u - u_h\|_{H_0^1(\Omega)} \lesssim \eta(\lambda_h, u_h).$$

Let $\mathbf{u}_h$, $\widetilde{\mathbf{u}}_h$ be the representation vector for $u_h$ and $\widetilde{u}_h$, calculated as an approximate solution of the generalized eigenvalue problem

$$A_h \mathbf{u}_h = \lambda_h B_h \mathbf{u}_h,$$

with the matrices, $A_h$ and $B_h$ being symmetric and symmetric positive definite, respectively, and $\|\cdot\|_{\mathbf{H}}$ be a discrete $H_0^1(\Omega)$-norm. Then

$$\|u_h - \widetilde{u}_h\|_{H_0^1(\Omega)} = \|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{\mathbf{H}},$$

and the following upper bound holds

$$\|u - \widetilde{u}_h\|_{H^1(\Omega)} \leq \|u - u_h\|_{H^1(\Omega)} + \|u_h - \widetilde{u}_h\|_{H^1(\Omega)} \lesssim \eta(\lambda_h, u_h) + \|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{\mathbf{H}}. \quad (3.2)$$

The correspondence between this upper bound and the upper bound derived in [13] is straightforward. The discretization error is controlled by the standard residual a posteriori error estimator, while the approximation error is controlled by the difference between underlying representation vectors measured in the discrete $H_0^1(\Omega)$-norm. In contrast to [13], where all the bounds are derived directly with the inexact eigenpair $(\widetilde{\lambda}_h, \widetilde{u}_h)$, here we first use the residual a posteriori error estimator derived with the Galerkin solution $u_h$, i.e., $\eta(\lambda_h, u_h)$ and afterwards apply the perturbation argument where we assume that the difference between $\eta(\lambda_h, u_h)$ and $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ can be measured by $\|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{\mathbf{H}}$, which makes this result slightly weaker, i.e., the lack of the lower bound prevents to prove the efficiency of the estimator.

The major questions are now: How are the discrete $H_0^1(\Omega)$-norm and its dual norm defined, how can we estimate the size of the approximation error in the discrete $H_0^1(\Omega)$-norm and how is the dual norm $H^{-1}(\Omega)$ of the discrete residual taken into account.

In [14] a shift-invert Lanczos method with a specially defined inner product was introduced. It was shown that applying the $\mathbf{H} := (A + B)$-inner product in the shift-invert Lanczos method for the matrix pencil $(A, B)$ enables to measure the size of the residual vector in the discrete $H^{-1}(\Omega)$-norm and that the $\mathbf{H}$-norm itself can be identified as a discrete $H_0^1(\Omega)$-norm. Let us recall the crucial perturbation result stated in [14].

PROPOSITION 3.1. *[14, Proposition 3.2]*

*Let $(A, B)$ be a symmetric definite pencil, $\mathbf{H} = A+B$ a symmetric positive definite matrix, and $\sigma$ a real number such that $A_\sigma := A - \sigma B$ is invertible. Let $\widetilde{x}$ be a nonzero vector in $\mathbb{R}^n$, $\widetilde{\lambda}$ a real number such that $\widetilde{\lambda} \neq \sigma$ and the residual vector*

$$r = A\widetilde{x} - \widetilde{\lambda} B\widetilde{x}.$$

*If*

$$\Big|\frac{1}{\lambda - \sigma} - \frac{1}{\widetilde{\lambda} - \sigma}\Big| = \min_{\lambda_i} \Big|\frac{1}{\lambda_i - \sigma} - \frac{1}{\widetilde{\lambda} - \sigma}\Big|,$$

*and* $Ax = \lambda Bx$, *where* $\|x\|_{\mathbf{H}} = 1$, *then*

$$\Big|\frac{\lambda - \widetilde{\lambda}}{\lambda - \sigma}\Big| \leq \frac{\|r\|_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\|\widetilde{x}\|_{\mathbf{H}}},$$

$$0 \leq |\sin_{\mathbf{H}} \angle(\widetilde{x}, x)| \leq \Big|\frac{\lambda_\gamma - \sigma}{\lambda_\gamma - \widetilde{\lambda}}\Big| \frac{\|r\|_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\|\widetilde{x}\|_{\mathbf{H}}},$$

*where*

$$\Big|\frac{1}{\lambda_\gamma - \sigma} - \frac{1}{\widetilde{\lambda} - \sigma}\Big| = \min_{\lambda_i \neq \lambda} \Big|\frac{1}{\lambda_i - \sigma} - \frac{1}{\widetilde{\lambda} - \sigma}\Big|.$$

*Proof.* See [14]. □

Following Proposition 3.1, the $\mathbf{H}$-norm can be identified as a discrete $H_0^1(\Omega)$-norm, whereas the $A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}$-norm can be identified as a discrete $H^{-1}(\Omega)$-norm.

Our next task is to approximate the eigenvector error in a discrete $H_0^1(\Omega)$-norm. With the appropriate normalization, $\|u_h\|_{H_0^1(\Omega)} = 1$, the difference between the representation vectors is given by

$$\|\mathbf{u}_h - \widetilde{\mathbf{u}}_h\|_{\mathbf{H}} = \sin_{\mathbf{H}} \angle(\mathbf{u}_h, \widetilde{\mathbf{u}}_h).$$

Let us now apply Proposition 3.1 to the pencil $(A_h, B_h)$. With $\mathbf{H} = A_h + B_h$ we have

$$\sin_{\mathbf{H}} \angle(\mathbf{u}_h, \widetilde{\mathbf{u}}_h) \lesssim \frac{\|\mathbf{r}\|_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\|\widetilde{\mathbf{u}}_h\|_{\mathbf{H}}}. \tag{3.3}$$

which together with (3.2) gives a final upper bound

$$\|u - \widetilde{u}_h\|_{H^1(\Omega)} \lesssim \eta(\lambda_h, u_h) + \frac{\|\mathbf{r}\|_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\|\widetilde{\mathbf{u}}_h\|_{\mathbf{H}}} := \eta_{new}(\lambda_h, u_h). \tag{3.4}$$

This theoretical upper bound has to be slightly modified to be used in practice. At first we are not able to compute $\eta(\lambda_h, u_h)$. However, since we know that the second term in (3.4) measures the approximation error, we can use $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ instead. On the other hand, we want to get the right hand side of (3.3) as a by product of the shift-invert Lanczos method. Therefore, as shown in [14], performing $m$ iterations of the shift-inverse Lanczos algorithm with the pencil $(A_h, B_h)$ and $\mathbf{H} = A_h + B_h$ inner product, gives an approximate eigenpair $(\widetilde{\lambda}_h, \widetilde{\mathbf{u}}_h)$ such that

$$\frac{\|\mathbf{r}\|_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\|\widetilde{\mathbf{u}}_h\|_{\mathbf{H}}} = \frac{\beta_m |e_m^T s|}{\theta},$$

where $(\theta, s)$ is an approximate eigenpair of the tridiagonal Lanczos matrix $H_m$ from the Lanczos factorization $V_m H_m + \beta_m v_{m+1} e_m^T$. For the complete analysis we refer to [14].

Consequently, after $m$ iterations of the shift-invert Lanczos method we have the following global upper bound for the eigenfunction error

$$\|u - \widetilde{u}_h\|_{H^1(\Omega)} \lesssim \eta(\widetilde{\lambda}_h, \widetilde{u}_h) + \frac{\beta_m |e_m^T s|}{\theta} := \eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h).$$

**4. The balanced AFEM algorithm.** With an appropriate norm and the new combined a posteriori error estimator, we can reduce the number of Lanczos iterations preformed during the AFEM algorithm for solving self-adjoint eigenvalue problems, e.g., the Laplace eigenvalue problem. We consider stopping criteria for the iterative eigensolver and the adaptive mesh refinement based on the new estimator $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$. As we have mentioned before, the upper bound (3.4) is a global upper bound. The only part of the new estimator which can be used to control the error locally, is a standard residual type error estimator, i.e., $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ or precisely elementwise refinement indicators $\eta_T(\widetilde{\lambda}_h, \widetilde{u}_h)$ extracted from it. Although, the discrete $H^{-1}(\Omega)$-norm of the algebraic residual contains the global information about the discrete problem, we can still define an appropriate stopping criterion and perform the adaptive mesh refinement.

Let us for a moment consider the ideal situation when after each Lanczos iteration step we check if the $H^{-1}(\Omega)$-norm of the algebraic residual, i.e., $\frac{\beta_m |e_m^T s|}{\theta}$, is smaller than a certain fraction $\omega$ of the corresponding continuous part of the error, i.e., $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$. Obviously, we can stop the iteration if this condition is satisfied or continue if not. This balancing strategy ensures that the iteration stops when the approximation error is of order of the discretization error. Therefore, the set of marked elements, determined by selecting only those elements for which local refinement indicators $\eta_T(\widetilde{\lambda}_h, \widetilde{u}_h)$ are satisfying a certain marking criterion, will guarantee a certain accuracy of the solution. Of course calculating all refinement indicators $\eta_T(\widetilde{\lambda}_h, \widetilde{u}_h)$ at every iteration of the eigensolver is not optimal. Therefore, the number $m$ of the Lanczos iterations performed before calculating $\eta_T(\widetilde{\lambda}_h, \widetilde{u}_h)$ and a proper parameter $\omega$ for equilibration have to be analyzed. In order to have an efficient estimator it is necessary to obtain also a lower bound, i.e.,

$$\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h) \lesssim \|u - \widetilde{u}_h\|_{H^1(\Omega)}.$$

Unfortunately, the proof for the efficiency of this new estimator is still under investigation. Instead, we will analyze the behavior of the estimator numerically. The pseudo-code of the balanced AFEM algorithm described in this section is presented below (Algorithm 1). In [14] the shift-invert Lanczos method is discussed, however, since we are interested is the smallest eigenvalue, we analyze the simplest version of the algorithm, i.e., with shift $\sigma = 0$. Of course calculating the error estimator $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ at each iteration step is expensive and not possible in practice, nevertheless, in order to formulate some hypothesis about the minimal number of required iterations we will make this effort here.

**5. Numerical experiments.** Throughout this section we investigate several aspects of our new balanced AFEM algorithm. As a model example, we consider the Laplace eigenvalue problem (3.1) on the L-shape domain, i.e., $\Omega = [-1,1] \times [0,1] \cup [-1,0] \times [-1,0]$. We dedicate all numerical experiments to approximating only the smallest eigenvalue which we compare with a reference value obtained in [18], i.e.,

$$\lambda_1 \approx 9.639723844.$$

At first we analyze the complexity and the accuracy dependence of our new algorithm on the choice of the balancing parameter $\omega$. Next we give some empirical information about the minimal number of Lanczos iterations required at each step of the balanced AFEM algorithm to determine an accurate final approximation. Since our combined error estimator $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$, contains a residual type a posteriori error estimator $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$, the last part of this section is dedicated to compare the behavior

**Algorithm 1** The balanced AFEM algorithm for computing the smallest eigenvalue of the Laplace eigenvalue problem (3.1)

**Input:** An initial regular triangulation $\mathcal{T}_h^i$, a balancing parameter $\omega$, a maximal number of degrees of freedom `maxDOF`

**Output:** Approximation $\widetilde{\lambda}_h$ to the smallest eigenvalue $\lambda$ of (3.1) together with the corresponding approximate eigenfunction $\widetilde{u}_h$.

**Solve:** Discretize problem (3.1) on $\mathcal{T}_h^i$ and obtain the matrix pencil $(A_h, B_h)$

Select shift $\sigma = 0$

**while** $\frac{\|\mathbf{r}\|_{A_h^{-1}\mathbf{H}A_h^{-1}}}{\|\widetilde{\mathbf{u}}_h\|_{\mathbf{H}}} \geq \omega\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ **do**

Perform one iteration of the **H**-Lanczos method with $(A_h, B_h)$

Compute the approximate eigenpair $(\widetilde{\lambda}_h, \widetilde{u}_h)$, the continuous and the discrete error estimators $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ and $\frac{\|\mathbf{r}\|_{A_h^{-1}\mathbf{H}A_h^{-1}}}{\|\widetilde{\mathbf{u}}_h\|_{\mathbf{H}}}$

**end while**

**Estimate:** Calculate the combined error estimator $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ and the standard error estimator $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$

**Mark:** Mark the elements based on $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ using the bulk criterion

**Refine:** Refine the coarse mesh $\mathcal{T}_h^i$ using the green, blue or red refinement to get $\mathcal{T}_h^{i+1}$

**if** #DOF $<$ `maxDOF` **then**

**return** $(\widetilde{\lambda}_H, \widetilde{\mathbf{u}}_H)$

**else**

Start the algorithm with $\mathcal{T}_h^{i+1}$

**end if**

of the combined error estimator $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ where $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ is chosen to be a standard residual type estimator $\eta_{DPR}(\widetilde{\lambda}_h, \widetilde{u}_h)$ [12] or an edge residual error estimator $\eta_{CG}(\widetilde{\lambda}_h, \widetilde{u}_h)$ [8].

All the experiments were realized with the OPENFFW [6] finite element framework, which contains implementations of both $\eta_{DPR}(\lambda_h, u_h)$ and $\eta_{CG}(\lambda_h, u_h)$ error estimators introduced in [12] and [8], respectively. For our new balanced AFEM algorithm we use our own implementation of the Lanczos method with the **H**-inner product, while the standard MATLAB *eigs* function is used for comparison.

**5.1. Balancing with different values of parameter $\omega$.** Let us first examine the behavior of our new balanced AFEM algorithm with respect to the choice of the parameter $\omega$. We consider here the ideal situation that after each Lanczos iteration step we can check whether the $H^{-1}(\Omega)$-norm of the algebraic residual, i.e., $\frac{\beta_m |e_m^T s|}{\theta}$, is smaller than a certain fraction $\omega$ of the corresponding $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$. Here we restrict ourselves to $\eta(\widetilde{\lambda}_h, \widetilde{u}_h) = \eta_{DPR}(\widetilde{\lambda}_h, \widetilde{u}_h)$ and we assume that the minimal number of required Lanczos iterations is set to $2k + 1$, where $k$ is the number of eigenvalues of interest, here $k = 1$.

Tables $1 - 3$ present the numerical results for $\omega = 0.5, 0.1, 0.9$, respectively. Except for the eigenvalue approximation and the number of degrees of freedom, information about the size of the continuous $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ and the discrete $\frac{\|\mathbf{r}\|_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\|\widetilde{\mathbf{u}}_h\|_{\mathbf{H}}}$ part of the

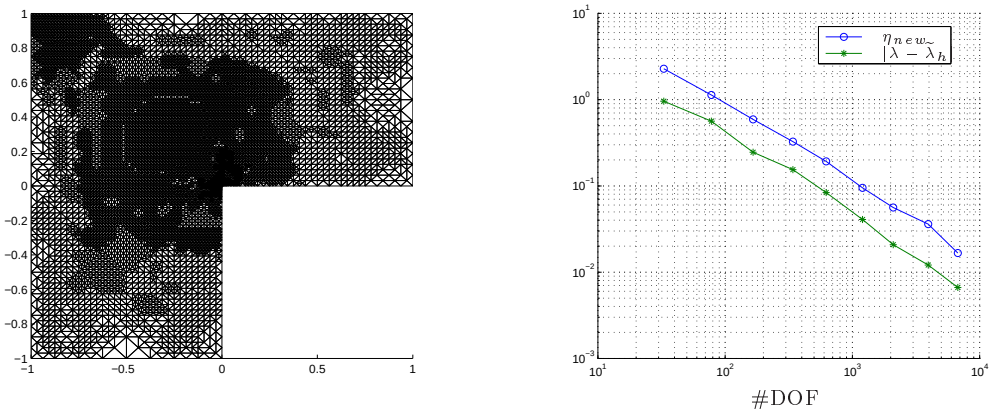


FIG. 1. *The final mesh with 6737 degrees of freedom and the convergence history for the smallest eigenvalue for (3.1) on the L-shape domain with* $\omega = 0.5$.

estimator is presented. We notice that for $\omega = 0.5$ three Lanczos iterations are enough to obtain a good error estimator to steer the adaptive mesh refinement.

Reducing $\omega$ to $\omega = 0.1$ forces the discrete residual to be smaller which leads to slightly more Lanczos iterations, however, the final accuracy is reached with a smaller amount of degrees of freedom. Obviously choosing $\omega = 0.9$ leads to meshes with more degrees of freedom but reduces the number of Lanczos iterations performed on each step of the adaptive algorithm. The corresponding final meshes and convergence history plots are depicted in Figures $1 - 3$.

These few examples show that determining the approximate solution of the same accuracy with a different balancing parameter is possible. Particularly chosen values of $\omega$ may lead to more Lanczos iterations or more degrees of freedom. Of course, performing one more Lanczos iteration on the coarse mesh is cheaper than dealing with finer problems, therefore, this decision has to be made by the user, depending on existing limitations.

TABLE 1
*Approximations of the smallest eigenvalue for (3.1) on the L-shape domain with* $\omega = 0.5$.

| ref. level | #DOF | # iterations | $\widetilde{\lambda}_1$ | $\lvert\lambda_1 - \widetilde{\lambda}_1\rvert$ | $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ | $\frac{\lVert \mathbf{r} \rVert_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ | $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ |
|---|---|---|---|---|---|---|---|
| 1 | 33 | 3 | 10.6008 | 0.9610 | 2.2454 | 0.0401 | 2.2853 |
| 2 | 78 | 3 | 10.2025 | 0.5628 | 1.0976 | 0.0349 | 1.1325 |
| 3 | 166 | 3 | 9.8854 | 0.2457 | 0.5476 | 0.0430 | 0.5906 |
| 4 | 341 | 3 | 9.7941 | 0.1544 | 0.2890 | 0.0373 | 0.3262 |
| 5 | 622 | 3 | 9.7234 | 0.0837 | 0.1540 | 0.0384 | 0.1924 |
| 6 | 1203 | 4 | 9.6805 | 0.0407 | 0.0866 | 0.0085 | 0.0950 |
| 7 | 2096 | 4 | 9.6606 | 0.0208 | 0.0472 | 0.0091 | 0.0562 |
| 8 | 3946 | 4 | 9.6518 | 0.0121 | 0.0273 | 0.0088 | 0.0361 |
| 9 | 6737 | 5 | 9.6464 | 0.0066 | 0.0154 | 0.0012 | 0.0167 |

TABLE 2
*Approximations of the smallest eigenvalue for (3.1) on the L-shape domain with* $\omega = 0.1$.

| ref. level | #DOF | # iteration | $\widetilde{\lambda}_1$ | $\lvert\lambda_1 - \widetilde{\lambda}_1\rvert$ | $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ | $\frac{\lVert \mathbf{r} \rVert_{A_{\sigma}^{-1}\mathbf{H}A_{\sigma}^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ | $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ |
|---|---|---|---|---|---|---|---|
| 1 | 33 | 3 | 10.5862 | 0.9465 | 2.2422 | 0.0284 | 2.2706 |
| 2 | 78 | 3 | 10.1976 | 0.5578 | 1.0908 | 0.0370 | 1.1278 |
| 3 | 168 | 3 | 9.8906 | 0.2509 | 0.5550 | 0.0448 | 0.5998 |
| 4 | 330 | 4 | 9.7850 | 0.1452 | 0.2964 | 0.0080 | 0.3044 |
| 5 | 601 | 4 | 9.7043 | 0.0646 | 0.1588 | 0.0075 | 0.1663 |
| 6 | 1174 | 4 | 9.6814 | 0.0416 | 0.0886 | 0.0087 | 0.0973 |
| 7 | 2048 | 5 | 9.6601 | 0.0204 | 0.0489 | 0.0013 | 0.0502 |
| 8 | 3839 | 5 | 9.6510 | 0.0113 | 0.0280 | 0.0012 | 0.0292 |
| 9 | 6585 | 5 | 9.6465 | 0.0068 | 0.0158 | 0.0013 | 0.0171 |

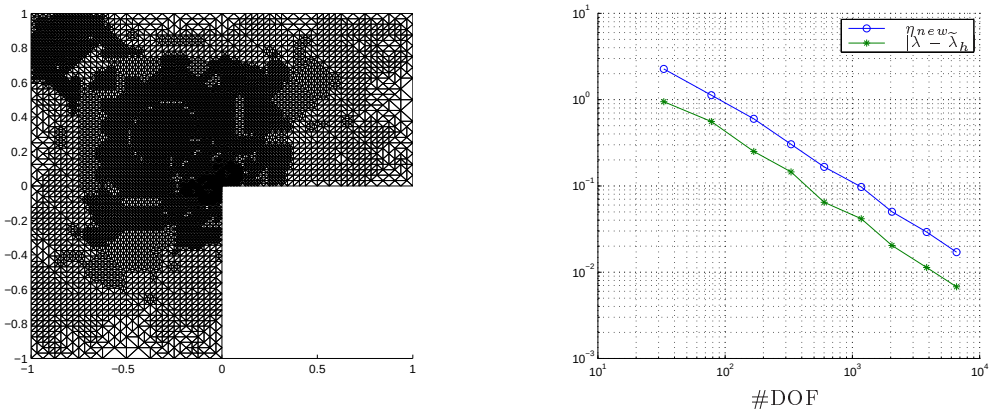


FIG. 2. *The final mesh with 6585 degrees of freedom and the convergence history for the smallest eigenvalue for (3.1) on the L-shape domain with* $\omega = 0.1$.

**5.2. Balancing with $\omega = 0.5$ and different minimal number of required Lanczos iteration.** Although, in [16] the minimal number of required Arnoldi/Lanczos iterations, as the well-known rule of thumb, is set to $2k+1$, where $k$ is the number of eigenvalues of interest, here we analyze the actual restrictions in this respect. Table 4 contains numerical results for the case where no restrictions are given. In first refinement steps we see that only one iteration was enough to assure that the iteration error is smaller than the discretization error, however, this one iteration was of course not enough to obtain a good approximation of the eigenvalue. Surprisingly, after few adaptive steps, the approximate solution has the same accuracy as the corresponding approximation (approximation at the same adaptive step) obtained by the AFEM algorithm starting with a much better approximation, see, e.g., ref. level 5 in Table 1 and 4. Of course the corresponding grid is much finer in the latter case. The resulting final mesh and the convergence history are presented in Figure 4.

We have already noticed, that performing only one Lanczos iteration may not be enough to assure the accuracy and the optimal complexity of our new balanced AFEM algorithm. Nevertheless, our next example shows that performing at least two

TABLE 3
*Approximations of the smallest eigenvalue for (3.1) on the L-shape domain with $\omega = 0.9$.*

| ref. level | #DOF | # iteration | $\widetilde{\lambda}_1$ | $\lvert\lambda_1 - \widetilde{\lambda}_1\rvert$ | $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ | $\frac{\lVert \mathbf{r} \rVert_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ | $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ |
|---|---|---|---|---|---|---|---|
| 1 | 33 | 3 | 10.5838 | 0.9441 | 2.2443 | 0.0244 | 2.2688 |
| 2 | 78 | 3 | 10.2022 | 0.5625 | 1.0891 | 0.0398 | 1.1289 |
| 3 | 169 | 3 | 9.8829 | 0.2431 | 0.5535 | 0.0385 | 0.5920 |
| 4 | 342 | 3 | 9.8110 | 0.1713 | 0.2888 | 0.0507 | 0.3395 |
| 5 | 638 | 3 | 9.7329 | 0.0932 | 0.1552 | 0.0469 | 0.2021 |
| 6 | 1209 | 3 | 9.7104 | 0.0707 | 0.0874 | 0.0457 | 0.1331 |
| 7 | 2200 | 4 | 9.6597 | 0.0200 | 0.0462 | 0.0088 | 0.0550 |
| 8 | 4028 | 4 | 9.6518 | 0.0121 | 0.0266 | 0.0088 | 0.0355 |
| 9 | 6916 | 4 | 9.6473 | 0.0075 | 0.0151 | 0.0088 | 0.0239 |

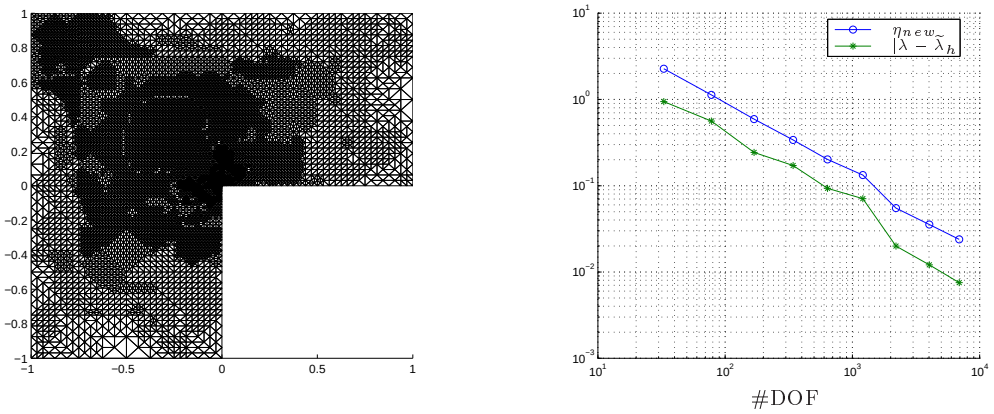


FIG. 3. *The final mesh with 6916 degrees of freedom and the convergence history for the smallest eigenvalue for (3.1) on the L-shape domain with $\omega = 0.9$.*

Lanczos iterations at every step of our AFEM algorithm seems to be enough. The corresponding numerical results and the convergence history are given in Table 5 and Figure 5.

For comparison Table 6 and Figure 6 present results obtained on the uniformly refined grid. The information about the size of the continuous $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ and the discrete $\frac{\lVert \mathbf{r} \rVert_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ part of the estimator is given only to illustrate the stopping criterion used for the Arnoldi/Lanczos process. None of this information is used during the grid generation process. To conclude, we point out that obtaining the solution with an accuracy of $10^{-3}$ for our balanced AFEM algorithm requires two times fewer degrees of freedom than determining the similar solution on the uniformly refined grid, with the same number of Arnoldi/Lanczos iterations.

**5.3. Comparison with different error estimators.** As we mentioned at the beginning of this section, our new combined error estimator consists of two parts: the continuous part $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ and the discrete part $\frac{\lVert \mathbf{r} \rVert_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$. So far we have analyzed

TABLE 4
*Approximations of the smallest eigenvalue for (3.1) on the L-shape domain with $\omega = 0.5$ and no restriction on the minimal number of iterations.*

| ref. level | #DOF | # iteration | $\widetilde{\lambda}_1$ | $\lvert\lambda_1 - \widetilde{\lambda}_1\rvert$ | $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ | $\frac{\lVert \mathbf{r} \rVert_{\mathbf{A}_\sigma^{-1}\mathbf{H}\mathbf{A}_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ | $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ |
|---|---|---|---|---|---|---|---|
| 1 | 33 | 1 | 37.9495 | 28.3098 | 9.3925 | 1.4206 | 10.8131 |
| 2 | 103 | 1 | 71.6865 | 62.0467 | 8.9268 | 2.1554 | 11.0823 |
| 3 | 230 | 1 | 124.0862 | 114.4465 | 8.9370 | 2.8001 | 11.7371 |
| 4 | 533 | 1 | 253.2587 | 243.6190 | 9.1481 | 4.0423 | 13.1904 |
| 5 | 1193 | 3 | 9.7568 | 0.1171 | 0.1237 | 0.0508 | 0.1745 |
| 6 | 1533 | 4 | 9.6856 | 0.0459 | 0.0769 | 0.0095 | 0.0865 |
| 7 | 2310 | 4 | 9.6642 | 0.0245 | 0.0450 | 0.0089 | 0.0539 |
| 8 | 4102 | 4 | 9.6536 | 0.0138 | 0.0257 | 0.0088 | 0.0346 |
| 9 | 7313 | 5 | 9.6466 | 0.0069 | 0.0145 | 0.0013 | 0.0157 |

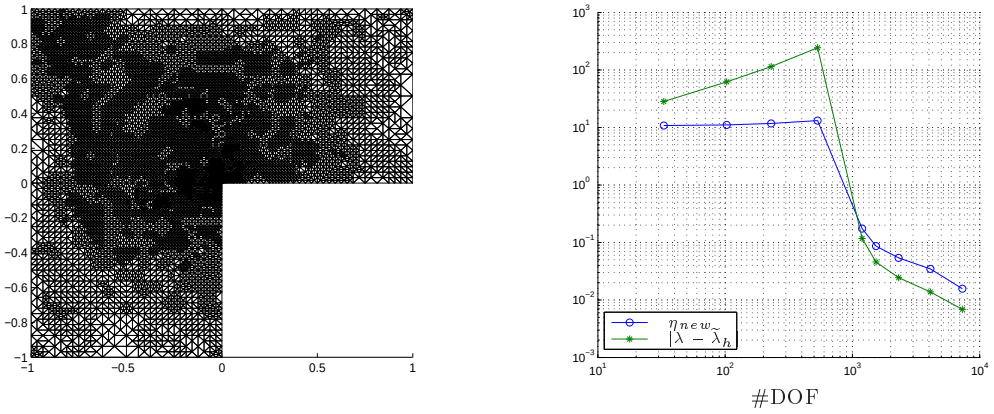


FIG. 4. *The final mesh with 7313 degrees of freedom and the convergence history for the smallest eigenvalue for (3.1) on the L-shape domain with $\omega = 0.5$ and no restriction on the minimal number of iterations.*

the behavior of our new balanced AFEM algorithm where the continuous estimator was chosen as $\eta_{DPR}(\widetilde{\lambda}_h, \widetilde{u}_h)$ [12]. Let us now first compare the performance of our new balanced AFEM algorithm with the standard AFEM algorithm based on the $\eta_{DPR}(\widetilde{\lambda}_h, \widetilde{u}_h)$ estimator, where the underlying algebraic eigenvalue problem will be solved with MATLAB function *eigs* up to final accuracy. The resulting meshes are presented in Figure 7, while Figure 8 shows the convergence history. Both algorithms result in the optimal convergence with respect to the number of degrees of freedom. The new combined error estimator, due to the right choice of the discrete norm, estimates the real error much better than the standard residual type estimator $\eta_{DPR}(\widetilde{\lambda}_h, \widetilde{u}_h)$. Furthermore, we compare our balancing algorithm with the standard AFEM algorithm where $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ is chosen as the edge residual error estimator $\eta_{CG}(\widetilde{\lambda}_h, \widetilde{u}_h)$ [8]. We observe that the original $\eta_{CG}(\widetilde{\lambda}_h, \widetilde{u}_h)$ guarantees the optimal convergence rate with respect to the number of degrees of freedom. Although, our balanced AFEM algorithm deviate slightly from the optimal convergence, it almost perfectly captures

TABLE 5
*Approximations of the smallest eigenvalue for (3.1) on the L-shape domain with $\omega = 0.5$ and the minimal number of required iterations equal 2.*

| ref. level | #DOF | # iteration | $\widetilde{\lambda}_1$ | $\lvert\lambda_1 - \widetilde{\lambda}_1\rvert$ | $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ | $\frac{\lVert \mathbf{r} \rVert_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ | $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ |
|---|---|---|---|---|---|---|---|
| 1 | 33 | 2 | 10.8075 | 1.1678 | 2.3114 | 0.1185 | 2.4299 |
| 2 | 86 | 2 | 10.5196 | 0.8799 | 1.1524 | 0.1641 | 1.3165 |
| 3 | 157 | 2 | 10.2919 | 0.6522 | 0.6536 | 0.1726 | 0.8262 |
| 4 | 339 | 3 | 9.7909 | 0.1512 | 0.2915 | 0.0415 | 0.3330 |
| 5 | 610 | 3 | 9.7367 | 0.0970 | 0.1552 | 0.0498 | 0.2049 |
| 6 | 1159 | 4 | 9.6802 | 0.0404 | 0.0852 | 0.0089 | 0.0941 |
| 7 | 2103 | 4 | 9.6592 | 0.0195 | 0.0466 | 0.0085 | 0.0550 |
| 8 | 3856 | 4 | 9.6517 | 0.0119 | 0.0265 | 0.0086 | 0.0351 |
| 9 | 6736 | 5 | 9.6459 | 0.0062 | 0.0150 | 0.0013 | 0.0163 |

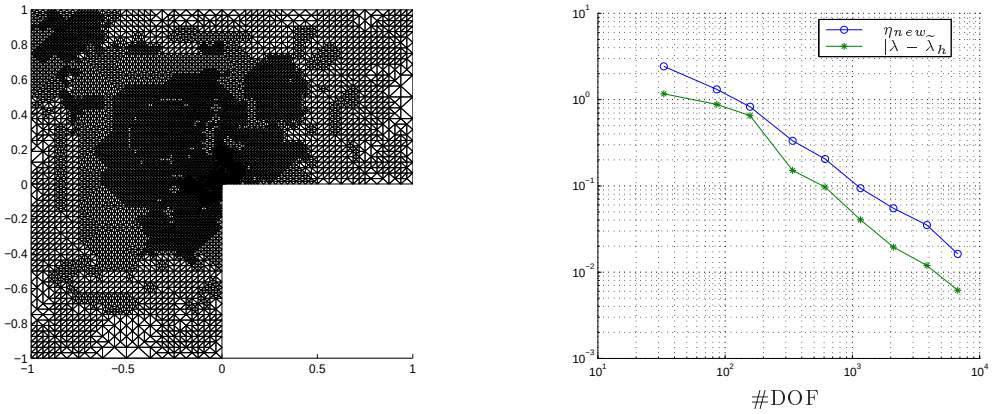


FIG. 5. *The final mesh with 6736 degrees of freedom and the convergence history for the smallest eigenvalue for (3.1) on the L-shape domain with $\omega = 0.5$ and the minimal number of required iterations equal 2.*

the behavior of the real error, see Figures 9 – 10.

**6. Conclusions.** We have formulated functional perturbation results for PDE eigenvalue problems including the functional backward error and the functional condition number. These results are used to relate the eigenvalue and the eigenvector error to the residual in the $H^{-1}(\Omega)$-norm and furthermore to establish a combined a posteriori error estimator embodying the discretization and the approximation error.

We have analyzed the behavior of the balanced AFEM algorithm which significantly reduces the number of eigensolver iteration by incorporating the discrete residual into the adaptation process. The eigensolver stopping criterion is based on an equilibration strategy. Several numerical examples confirm the reliability of our estimator. A formal proof for the efficiency of the combined error estimator and the convergence of the balancing algorithm is still an open question. Also there are several other choices of the eigensolver stopping criterion which are an interesting subject for further research. The convergence of the inexact adaptive algorithm based on a

TABLE 6
*Approximations of the smallest eigenvalue for (3.1) on the L-shape domain refined uniformly with $\omega = 0.5$ and the minimal number of required iterations equal 2.*

| ref. level | #DOF | # iteration | $\widetilde{\lambda}_1$ | $\lvert\lambda_1 - \widetilde{\lambda}_1\rvert$ | $\eta(\widetilde{\lambda}_h, \widetilde{u}_h)$ | $\frac{\lVert \mathbf{r} \rVert_{A_\sigma^{-1}\mathbf{H}A_\sigma^{-1}}}{\lVert \widetilde{\mathbf{u}}_h \rVert_{\mathbf{H}}}$ | $\eta_{new}(\widetilde{\lambda}_h, \widetilde{u}_h)$ |
|---|---|---|---|---|---|---|---|
| 1 | 33 | 2 | 11.2225 | 1.5828 | 2.4189 | 0.1851 | 2.6040 |
| 2 | 161 | 2 | 10.4686 | 0.8289 | 0.6589 | 0.1867 | 0.8455 |
| 3 | 705 | 3 | 9.7616 | 0.1219 | 0.1711 | 0.0476 | 0.2186 |
| 4 | 2945 | 4 | 9.6710 | 0.0312 | 0.0520 | 0.0092 | 0.0612 |
| 5 | 12033 | 5 | 9.6504 | 0.0107 | 0.0162 | 0.0012 | 0.0174 |

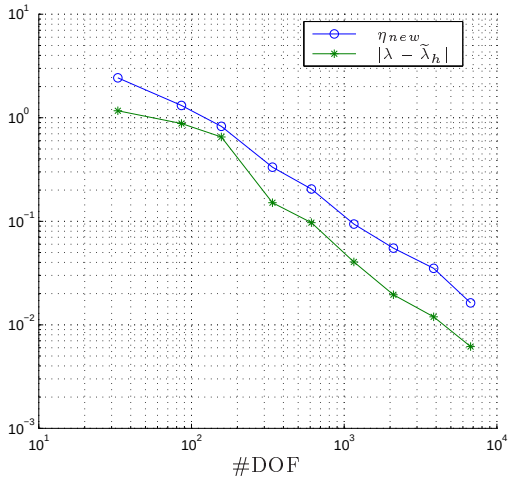


FIG. 6. *The convergence history for the smallest eigenvalue for (3.1) on the uniformly refined L-shape domain with 12033 degrees of freedom, $\omega = 0.5$ and the minimal number of required iterations equal 2.*

slightly different combined a posteriori error estimator is a subject of ongoing work.

**7. Acknowledgment.** The author would like to thank Mario Arioli and Ulrich Hetmaniuk for valuable comments and suggestions. We also thank the DFG Research Center MATHEON and the Berlin Mathematical School for the financial support.

REFERENCES

[1] M. Arioli, E. H. Georgoulis, and D. Loghin. Convergence of inexact adaptive finite element solvers for elliptic problems. Technical Report RAL-TR-2009-21, Rutherford Appleton Laboratory, 2009.
[2] M. Arioli, E. Noulard, and A. Russo. Stopping criteria for iterative methods: applications to PDE's. *CALCOLO*, 38:97–112, 2001.
[3] M. Bennani and T. Braconnier. Stopping criteria for eigensolvers. Technical Report TR/PA/94/22, CERFACS, 1994.
[4] J. Bramble, J. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55:1–22, 1990.
[5] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springer Ser. Comput. Math. 15. Springer-Verlag, New York, 1991.

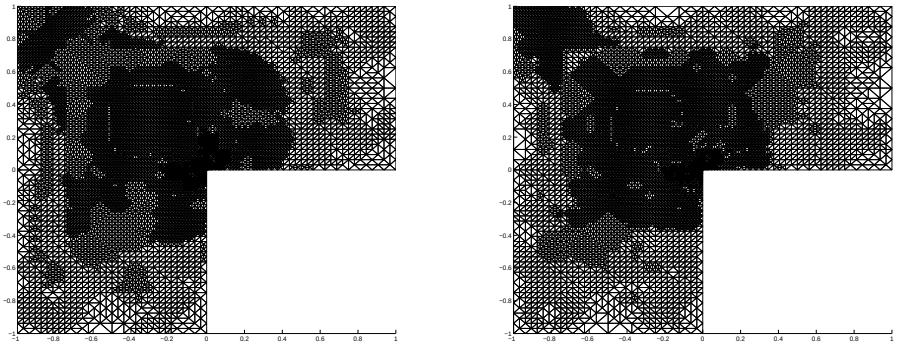

FIG. 7. *The final mesh for the AFEM with $\eta_{DPR}$ with 6990 degrees of freedom (left) and with $\eta_{new}$ with 6610 degrees of freedom (right).*

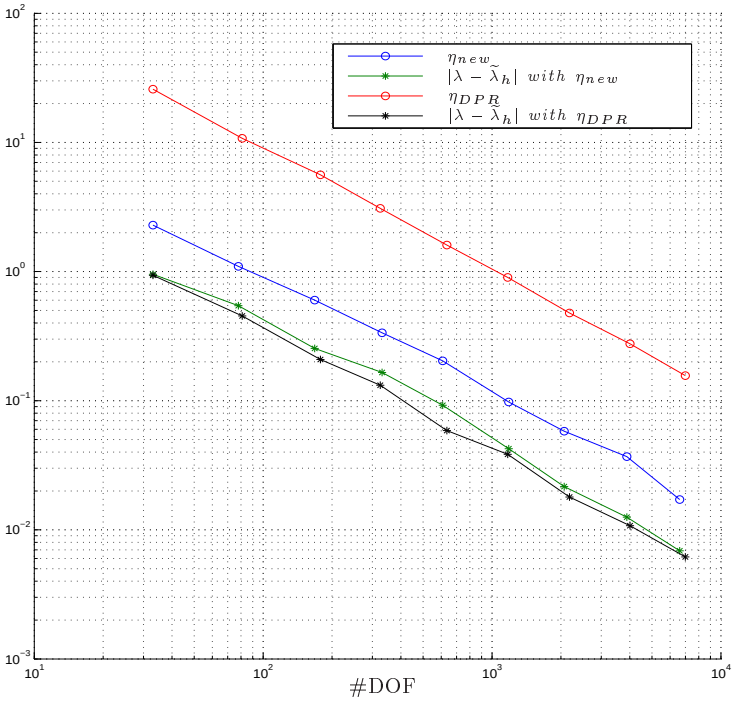


FIG. 8. *The convergence history for the smallest eigenvalue for (3.1) on the L-shape domain.*

[6] A. Byfut, J. Gedicke, D. Günther, J. Reininghaus, and S. Wiedemann. OPENFFW, The Finite Element Framework. GNU General Public License v.3.
[7] C. Carstensen and J. Gedicke. An adaptive finite element eigenvalue solver of quasi-optimal computational complexity. Preprint 662, DFG Research Center MATHEON, Strasse des 17. Juni 136, D-10623 Berlin, 2009. http://www.matheon.de/.
[8] C. Carstensen and J. Gedicke. An oscillation-free adaptive FEM for symmetric eigenvalue

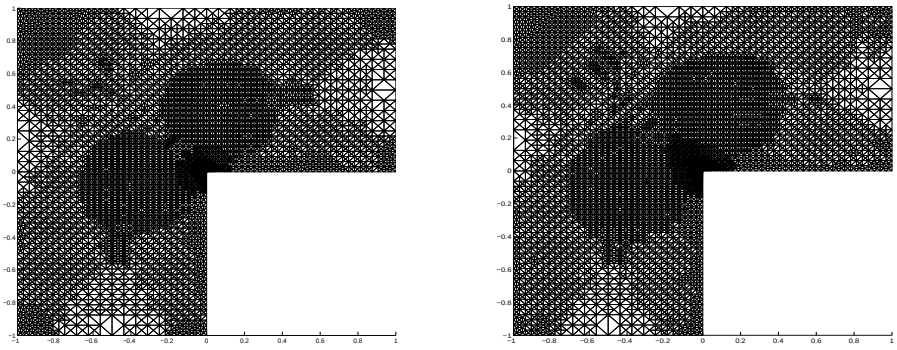

FIG. 9. *The final mesh for the AFEM with $\eta_{CG}$ with 5577 degrees of freedom (left) and the final mesh for AFEM with $\eta_{new}$ with 5995 degrees of freedom (right).*

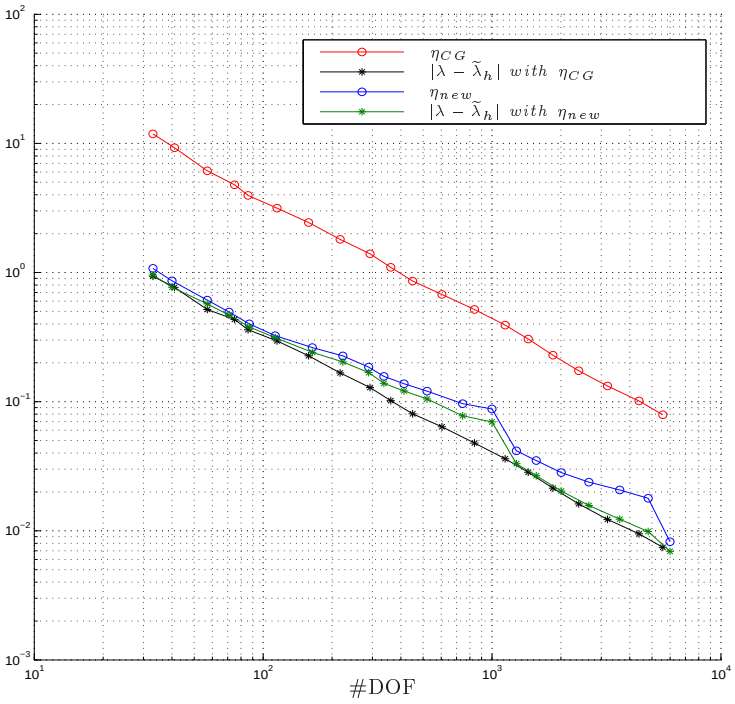


FIG. 10. *The convergence history for the smallest eigenvalue for (3.1) on the L-shape domain.*

problems. *Numer. Math.*, 118(3):401 – 427, 2011.

[9] J. M. 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.

[10] F. Chaitin-Chatelin. The influence of nonnormality on matrix computations. In R. J. Plemmons and C. D. Meyer, editors, *Linear Algebra, Markov Chains and Queing Models*, pages 13–19. Springer-Verlag, New York, 1992.

[11] F. Chaitin-Chatelin and V. Frayssé. *Lectures on Finite Precision Computations - Software, Environments, and Tools.* SIAM Publications, 1996.
[12] R. G. Durán, C. Padra, and R. Rodríguez. A posteriori error estimates for the finite element approximation of eigenvalue problems. *Math. Mod. Meth. Appl. Sci.*, 13:1219–1229, 2003.
[13] H. Harbrecht and R. Schneider. On error estimation in finite element methods without having Galerkin orthogonality. Preprint 457, Berichtsreihe des SFB 611, Universität Bonn, 2009.
[14] U. L. Hetmaniuk and R. B. Lehoucq. Uniform accuracy of eigenpairs from a shift-invert Lanczos method. *SIAM J. Matrix Anal. Appl.*, 28:927–948, 2006.
[15] P. Jiranek, Z. Strakoš, and M. Vohralik. A posteriori error estimates including algebraic error: computable upper bounds and stopping criteria for iterative solvers. *SIAM J. Sci. Comput.*, 32:1567–1590, 2010.
[16] R. B. Lehoucq, D. C. Sorensen, and C. Yang. *ARPACK User's Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods.* SIAM Publications, Philadelphia, 1998.
[17] A. Miedlar. *Inexact Adaptive Finite Element Methods for Elliptic PDE Eigenvalue Problems.* PhD thesis, Technische Universität Berlin, 2011.
[18] L. N. Trefethen and T. Betcke. Computed eigenmodes of planar regions. *Contemp. Math.*, 412:297–314, 2006.
[19] J. H. Wilkinson. *The Algebraic Eigenvalue Problem.* Oxford University Press, Oxford, 1965.
[20] Ch.-T. Wu and H. C. Elman. *Stopping criteria for iterative methods in solving convection-diffusion equations on adaptive meshes*, 2004. Presented in the third international congress of Chinese mathematicians, 2004.
[21] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, 2010.