Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 459895, 14 pages
http://dx.doi.org/10.1155/2015/459895

*Research Article*

# A Subspace Iteration Algorithm for Fredholm Valued Functions

**Christian Engström[1] and Luka Grubišić[2]**

[1]*Department of Mathematics and Mathematical Statistics, Umeå University, MIT-Huset, 90187 Umeå, Sweden*
[2]*Department of Mathematics, University of Zagreb, Bijenička 30, 10000 Zagreb, Croatia*

Correspondence should be addressed to Luka Grubišić; luka.grubisic@math.hr

Received 2 July 2015; Revised 2 October 2015; Accepted 5 October 2015

Academic Editor: Ruben Specogna



We present an algorithm for approximating an eigensubspace of a spectral component of an analytic Fredholm valued function. Our approach is based on numerical contour integration and the analytic Fredholm theorem. The presented method can be seen as a variant of the FEAST algorithm for infinite dimensional nonlinear eigenvalue problems. Numerical experiments illustrate the performance of the algorithm for polynomial and rational eigenvalue problems.

## 1. Introduction

In this paper we study analytic operator eigenvalue problems defined on an open connected subset $\Omega \subseteq \mathbb{C}$ in a separable Hilbert space $\mathscr{H}$. Throughout this paper we assume that $S : \Omega \to \mathscr{L}(\mathscr{H})$ is an analytic function with values in the space $\mathscr{L}(\mathscr{H})$ of bounded linear operators. A scalar $\lambda \in \Omega$ is called an eigenvalue of $S$ if $S(\lambda)$ is not injective. Hence, the eigenvalue problem is to find $\lambda \in \Omega$ and $u \in \mathscr{H} \setminus \{0\}$ such that

$$S(\lambda)\, u = 0. \tag{1}$$

Such problems are, for example, used to study the dispersion and damping properties of waves [1–3]. Given a closed contour $\Gamma$, we would like to approximate all the eigenvalues of $S$ inside $\Gamma$ to the sufficient degree of accuracy. In this paper the numerical method is based on contour integrals of the generalized resolvent $S^{-1}$. The state-of-the-art results in the contour integration based methods for solving nonlinear matrix eigenvalue problems are presented in [4–6] including the references therein. Results for contour integration based solution methods for Fredholm valued eigenvalue problems can be found in, for example, [7, 8].

The spectrum $\sigma(S)$ of the operator function $S$ is the set of all $\lambda \in \Omega$ such that $S(\lambda)$ is not invertible in $\mathscr{L}(\mathscr{H})$ and the resolvent set is defined as the complement $\rho(S) = \Omega \setminus \sigma(S)$. For $F \in \mathscr{L}(\mathscr{H})$ we define the operator norm by the expression $\|F\| = \sqrt{\operatorname{spr}(F^* F)}$, where $\operatorname{spr}(\cdot)$ denotes the spectral radius. We call an operator $F \in \mathscr{L}(\mathscr{H})$ a Fredholm operator if the dimensions of its null space $\operatorname{Ker}(F)$ and of the orthogonal complement of its range $\operatorname{CoKer}(F) = \operatorname{Ran}(F)^{\perp}$ are finite. By $\Phi(\mathscr{H})$ we denote the set of all Fredholm operators on $\mathscr{H}$ and the number $\operatorname{ind}(F) = \dim \operatorname{Ker}(F) - \dim \operatorname{CoKer}(F)$ is called the index of $F \in \Phi(\mathscr{H})$. In what follows we will assume that $S(\lambda) \in \Phi(\mathscr{H})$ for all $\lambda \in \Omega$. If in addition the resolvent set of such $S$ is nonempty, the analytic Fredholm theorem, for example, [9, Theorem 1.3.1], implies that the generalized resolvent $z \mapsto S(z)^{-1}$ is finitely meromorphic. This in turn implies that the spectrum $\sigma(S)$ is countable and the geometric multiplicity of $\lambda$, that is, $\dim(\operatorname{Ker} S(\lambda))$, is finite. Moreover, the associated Jordan chains of generalized eigenvectors have finite length bounded by the algebraic multiplicity; see [9].

The results of this paper are a combination of matrix techniques from [6] with the specialization of the results from [7] to Hilbert spaces. In particular, we leverage the technique of block operator matrix representation of Fredholm valued operator functions and prove that the convergence rate for the inexact subspace iteration algorithm depends primarily on the spectral properties of the operator function. Also, we make the case for the problem dependent determination of the number of integration nodes depending on the clustering of eigenvalues towards the contour of integration (cf. [6]), where the use of 16 nodes of Gauss-Legendre integration formula is recommended. To assess the quality of a computed eigenpair, rank one perturbations of the operator function are studied. In particular, we construct a perturbation based on the residual functional and estimate the approximation

errors by estimating the norm of the residual by an auxiliary subspace technique. Our algorithm consists of the inexact subspace iteration for the zeroth moment of the resolvent to construct the approximate eigenspace for the eigenvalues contained inside a contour $\Gamma$. We then use the moment method of Beyn et al. [4, 7] to extract information on eigenvalues from the computed approximate eigenspace. As the convergence criterion we use a hierarchical residual estimate. In the case in which the convergence criterion is not satisfied the procedure is repeated. This structure of the algorithm will be replicated directly in the structure of the paper and is presented as Algorithm 2.

The paper is organized as follows. In Section 2 we establish a criterion based on the residual norm for assessing approximations of simple eigenvalues. In Section 3 we present inexact subspace iteration algorithm based on contour integration and prove that its convergence rate essentially depends on the properties of the operator function. It is shown that the influence of the integration formula diminishes exponentially with the number of integration nodes. Finally, in Section 4 we present numerical experiments.

## 2. Notation and Basic Analytic Results

In this section we present the machinery of quasi-matrices from [10–12]. In particular we present basic results on the angles between finite dimensional subspaces of a Hilbert space in terms of quasi-matrix notation. Finally, we will prove an error estimation result for simple eigenvalues of a Fredholm valued function.

A quasi-matrix is a bounded linear operator $V$ from the finite dimensional space $\mathbb{C}^r$ to an (in general) infinite dimensional Hilbert space $\mathscr{H}$. Then, the product $V^*V$ denotes the Gram matrix:

$$\left(V^*V\right)_{ij} = \left(Ve_i, Ve_j\right), \quad i, j = 1, 2, \ldots, r, \tag{2}$$

which depends on the inner product $(\cdot,\cdot)$ of $\mathscr{H}$. Let $P_1$, $P_2$, $Q_1$, and $Q_2$ be orthogonal projections such that $P_1 \oplus P_2 = I$ and $Q_1 \oplus Q_2 = I$, where $I$ is the identity operator on $\mathscr{H}$. Furthermore, let $\mathscr{H}_i := P_i\mathscr{H}$, $\widetilde{\mathscr{H}}_i := Q_i\mathscr{H}$, $i = 1, 2$, and identify $\mathscr{H}$ and the two spaces $\mathscr{H}_1 \oplus \mathscr{H}_2$ and $\widetilde{\mathscr{H}}_1 \oplus \widetilde{\mathscr{H}}_2$ by isomorphism. Let $B \in \mathscr{L}(\mathscr{H})$ and take

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathscr{H}_1 \oplus \mathscr{H}_2 := \mathscr{H}. \tag{3}$$

Then, $Q_iBu = B_{i1}u_1 + B_{i2}u_2$, $i = 1, 2$, where $B_{ij} : \mathscr{H}_j \to \widetilde{\mathscr{H}}_i$ is defined by restricting $B_{ij} := Q_iBP_j$ onto appropriate spaces. Hence, the bounded operator $B$ has a block operator matrix representation $B : \mathscr{H}_1 \oplus \mathscr{H}_2 \to \widetilde{\mathscr{H}}_1 \oplus \widetilde{\mathscr{H}}_2$ in terms of $B_{ij}$ and $Bu$ is computed following the rules of matrix algebra:

$$Bu = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} B_{11}u_1 + B_{12}u_2 \\ B_{21}u_1 + B_{22}u_2 \end{bmatrix}. \tag{4}$$

The multiplication of two block operator matrices also follows the rules of matrix algebra. Represent, for example, $A : \mathscr{H} \to \mathscr{H}$ by $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$, where $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2$ and $A_1 : \mathscr{H}_1 \to \mathscr{H}$ and $A_2 : \mathscr{H}_2 \to \widetilde{\mathscr{H}}$. Then, the operator $AB : \mathscr{H} \to \mathscr{H}$ has the block representation:

$$AB = \begin{bmatrix} A_1B_{11} + A_2B_{21} & A_1B_{12} + A_2B_{22} \end{bmatrix}. \tag{5}$$

Here we have exemplarily taken the trivial partition of unity $Q_1 = I$ and $Q_0 = 0$ and we assume that both $P_1 \neq 0$ and $P_2 \neq 0$. This illustrates the flexibility we have in choosing the block operator matrix representations for bounded operators. For more details on this construction see a recent book by C. Tretter [13]. To make the paper more readable we will use $I_r$ to denote the identity operator on the finite dimensional space $\mathbb{C}^r$ and $I$ will denote the identity operator on $\mathscr{H}$.

Let $P_1 \oplus P_2 = I$ be given such that $\dim \operatorname{Ran}(P_1) = r$. Let $V$ be a unitary operator such that $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ and $\operatorname{Ran}(P_1) = \operatorname{Ran}(V_1)$ and $\operatorname{Ran}(P_2) = \operatorname{Ran}(V_2)$. Note that in this setting we have $P_1 = V_1V_1^*$. For the quasi-matrix $X : \mathbb{C}^r \to \mathscr{H}$ such that $X^*X = I_r$ we can write

$$X = V_1C + V_2S, \tag{6}$$

where $C = V_1^*X$ and $S = V_2^*X$. With this notation we compute

$$I_r = X^*X = C^*C + S^*S, \tag{7}$$

and so $\|S\| \leq 1$ and $\|C\| \leq 1$. Furthermore, note that $Q = XX^*$ is the orthogonal projection onto $\operatorname{Ran}(X)$ and so from [14, Theorem 2] it follows that

$$\begin{aligned} \left\|P_1 - Q\right\| &= \left\|V_1V_1^* - XX^*\right\| = \left\| \begin{bmatrix} I_r - CC^* & -CS^* \\ -SC^* & -SS^* \end{bmatrix} \right\| \\ &= \|S\|. \end{aligned} \tag{8}$$

The last identity has been established by spectral calculus using the fact that $\dim\operatorname{Ran}(P_1) = r$; see [15]. Let $\|S\| < 1$; then we define the unique number $\theta \in [0, \pi/2\rangle$ such that

$$\sin\theta = \|S\|. \tag{9}$$

We call $\theta$ the maximal canonical angle between the spaces $\operatorname{Ran}(X)$ and $\operatorname{Ran}(P_1) = \operatorname{Ran}(V_1)$. Since $\|S\| < 1$ (7) implies that $C^*C$ is a positive definite matrix, therefore $C$ must be invertible. By direct computation using spectral calculus, as in [15], we establish that

$$\begin{aligned} \cos\theta &= \|C\|, \\ \tan\theta &= \left\|SC^{-1}\right\|. \end{aligned} \tag{10}$$

When we are considering several pairs of quasi-matrices $V_1$ and $X$ we will write $\theta(V_1, X)$ to denote the maximal canonical angle between subspaces $\operatorname{Ran}(V_1)$ and $\operatorname{Ran}(X)$.

This definition can be extended to subspaces of different dimensions using pairs of orthogonal projections and their singular value decomposition [14, 15]. In this case we call the norm of the difference of the two projections the gap distance between subspaces; see [14, 15].

*2.1. Application of Abstract Results to Operators Defined by Sesquilinear Forms.* The abstract Fredholm analytic theorem is stated for bounded operators between Hilbert spaces. Since we are interested in finite element computations, our problems will always be stated in the variational form which assumes working with unbounded sectorial operators. See [16] for definitions and the terminology relating to unbounded sectorial operators and forms. Below we will formulate the Fredholm analytic theorem in this setting.

Let $\mathfrak{t}_0$ be a densely defined closed symmetric form in $\mathscr{H}$ which is semibounded from below with a strictly positive lower bound; see [16, Section VI.2]. We will call such quadratic forms positive definite forms and use $\mathrm{Dom}(\mathfrak{t}_0) \subset \mathscr{H}$ to denote its domain of definition. To simplify the notation we will write $\mathscr{V} := \mathrm{Dom}(\mathfrak{t}_0)$ in the rest of the paper. Additionally, let $\mathfrak{c}_i$, $i = 1, \ldots, d$, be a sequence of sesquilinear (not necessarily symmetric) forms which are relatively compact [16] with regard to $\mathfrak{t}_0$. Moreover, assume that $f_i(\cdot)$, $i = 1, \ldots, d$, is a sequence of scalar analytic functions in $\Omega$. Define the family of sesquilinear forms:

$$\mathfrak{t}(z)[\phi, \psi] := \mathfrak{t}_0[\phi, \psi] + \sum_{i=1}^{d} f_i(z)\, \mathfrak{c}_i[\phi, \psi], \qquad \phi, \psi \in \mathscr{V},\ z \in \Omega. \tag{11}$$

In a variationally posed eigenvalue problem we are seeking a scalar $\lambda \in \Omega$ and a vector $\phi \in \mathscr{V} \setminus \{0\}$ such that

$$\mathfrak{t}(\lambda)[\phi, \psi] = 0, \quad \psi \in \mathscr{V}. \tag{12}$$

To this variational formulation we construct a representation by an operator-valued function $T(\cdot)$ and then apply the analytic Fredholm theorem to establish the structure of the spectrum.

Let $T_0$, $\mathrm{Dom}(T_0^{1/2}) = \mathscr{V}$ be a self-adjoint positive definite operator which represents the form $\mathfrak{t}_0$ in the sense of Kato's second representation theorem; see [16, Theorem VI.2.23]. Let $K_i$ be defined by

$$(K_i u, v) = \mathfrak{c}_i\left[T_0^{-1/2} u, T_0^{-1/2} v\right], \quad u, v \in \mathscr{H}. \tag{13}$$

The operators $K_i$, $i = 1, \ldots, d$, are obviously compact and for $z \in \Omega \setminus \sigma(T)$ the value of the generalized resolvent is the operator:

$$T^{-1}(z) = T_0^{-1/2}\left(I + \sum_{j=1}^{d} f_j(z) K_j\right)^{-1} T_0^{-1/2}. \tag{14}$$

Here $T(z)$ is the unbounded sectorial operator with domain $\mathrm{Dom}(T(z))$ such that

$$(T(z)\phi, \psi) = \mathfrak{t}_0[\phi, \psi] + \sum_{i=1}^{d} f_i(z)\, \mathfrak{c}_i[\phi, \psi], \qquad \phi \in \mathrm{Dom}(T(z)),\ \psi \in \mathscr{V}. \tag{15}$$

Obviously the operator-valued function

$$S(z) = I + \sum_{j=1}^{d} f_j(z) K_j, \quad z \in \Omega \tag{16}$$

satisfies the requirements of the analytic Fredholm theorem and $\sigma(S) = \sigma(T)$. Let us note that we will use

$$\left(T'(z)\phi, \psi\right) = \sum_{i=1}^{d} f_i'(z)\, \mathfrak{c}_i[\phi, \psi], \qquad \phi \in \mathrm{Dom}\left(T'(z)\right),\ \psi \in \mathscr{V}, \tag{17}$$

to define a derivative of an operator-valued analytic function. We can now define the notion of a semisimple eigenvalue.

*Definition 1.* Let $T$ be as in (15) and let $\mu \in \sigma(T)$ be an eigenvalue. The eigenvalue $\mu$ is semisimple if for each $\psi \in \mathrm{Ker}(T(\mu)) \setminus \{0\}$ there is a $\psi \in \mathrm{Ker}(T^*(\mu))$ such that

$$\left(T'(\mu)\phi, \psi\right) \neq 0. \tag{18}$$

If $\dim \mathrm{Ker}(T(\mu)) = 1$, then $\mu$ is called a simple eigenvalue.

Note that in the quasi-matrix notation we will freely write

$$\left(T'(\mu)\phi, \psi\right) = \psi^* T'(\mu)\phi. \tag{19}$$

To this end we identify the vectors with a mapping $\psi : \mathbb{C} \to \mathscr{H}$.

With these conventions we state, informally, the generalized argument principle proved by Gohberg and Sigal in [17–19]. It states that for the closed contour $\Gamma \subset \rho(T)$ the number

$$M(T, \Gamma) = \mathrm{tr}\left(\int_\Gamma T(z)^{-1} T'(z)\, dz\right) = \mathrm{tr}\left(\int_\Gamma T'(z) T(z)^{-1}\, dz\right) \tag{20}$$

satisfies $M(T, \Gamma) \in \mathbb{N}$ and it equals the total multiplicity of the eigenvalues enclosed by $\Gamma$. We also have the following consequence of [9, Theorem 1.3.1].

**Proposition 2.** *Let us assume that we have a variational eigenvalue problem (12) with the operator representation $T : \Omega \to \mathscr{H}$ given by (15). Then the spectrum $\sigma(T)$ consists of a countable collection of eigenvalues with finite multiplicity. Further, let the component of $\sigma(T)$ inside a contour $\Gamma$ consist only of semisimple eigenvalues $\lambda_i$, $i = 1, \ldots, r$, such that $n_{\lambda_i} = \dim \mathrm{Ker}(T(\lambda_i))$. Then there are quasi-matrices $X_i, Y_i : \mathbb{C}^{n_{\lambda_i}} \to \mathscr{H}$, $i = 1, \ldots, r$, such that $\mathrm{Ran}(X_i) \subset \mathrm{Ker}(T(\lambda_i))$ and $\mathrm{Ran}(Y_i) \subset \mathrm{Ker}(T^*(\lambda_i))$, $i = 1, \ldots, r$, and an open once connected neighborhood $\mathscr{U}$ containing $\lambda_i$, $i = 1, \ldots, r$ and $\Gamma$ and an operator-valued function $H$ which is analytic on $\mathscr{U}$ and taking values in $\mathscr{L}(\mathscr{H})$ such that*

$$T^{-1}(z) = \sum_{i=1}^{r} \frac{1}{z - \lambda_i} X_i Y_i^* + H(z), \quad z \in \mathscr{U}, \tag{21}$$

*and $Y_i^*(T'(\lambda_i)X_i) = I_{n_{\lambda_i}}$, $i = 1, \ldots, r$.*

*Proof.* For $z \in \Omega \setminus \sigma(T)$ write

$$T^{-1}(z) = T_0^{-1/2}\left(I + \sum_{j=1}^{d} f_j(z) K_j\right)^{-1} T_0^{-1/2} \tag{22}$$

and define the Fredholm valued function:

$$S(z) = I + \sum_{j=1}^{d} f_j(z) K_j, \quad z \in \Omega. \tag{23}$$

Recall that $\mathrm{Dom}(T_0^{1/2}) = \mathscr{V}$ and that $T_0^{1/2}$ maps $\mathscr{V}$ one to one on $\mathscr{H}$. Now apply [9, Theorem 1.3.1] on $S$. □

*2.2. Estimating Eigenvalues inside a Contour.* To count the semisimple eigenvalues inside a contour we will use the approach of [20]. We will limit our consideration on the case of semisimple eigenvalues and rank one perturbations. First we will present results for Fredholm valued operator functions and then formulate the result for operators defined by sesquilinear forms.

**Lemma 3.** *Let $S : \Omega \to \Phi(\mathscr{H})$ be an analytic function and let $E$ be a bounded operator such that* $\dim \mathrm{Ran}(E) = 1$. *Assume that $\Gamma \subset \rho(S)$ is a simple closed contour such that the component of $\sigma(S)$ inside $\Gamma$ consists solely of a simple eigenvalue $\lambda$. If $S(z) + \tau E$ is invertible for all $z \in \Gamma$ and all $\tau \in [0,1]$, then $T + E$ has a simple eigenvalue $\widetilde{\lambda}$ inside $\Gamma$ and*

$$\left|\lambda - \widetilde{\lambda}\right| \le C \|E\|, \tag{24}$$

*where $C$ essentially depends on* $\max_{z\in\Gamma}\|S(z)^{-1}|_{\mathrm{Ran}(E)}\|$ *and* $\max_{z\in\Gamma}\|S(z)^{-*}|_{\mathrm{Ran}(E^*)}\|$ *and the length of the the integration curve $\Gamma$.*

*Proof.* Recall that $(S(z) + E)' = S'(z)$ and define the function:

$$f(\tau) = \frac{1}{2\pi i}\mathrm{tr}\left(\int_\Gamma S'(z)(S(z) + \tau E)^{-1} dz\right). \tag{25}$$

By the generalized argument principle—see [17–19]—the value of $f(\tau)$ equals the total multiplicity of the eigenvalues inside $\Gamma$. In particular, by Proposition 2 we have that there are vectors $x$ and $y$ such that $(S'(\lambda)x, y) = y^* S'(\lambda)x = 1$ and

$$f(0) = \mathrm{tr}\left(S'(\lambda) x y^*\right) = \mathrm{tr}\left(y^* S'(\lambda) x\right) = 1, \tag{26}$$

where we used the circularity of the trace. By the assumptions of the theorem $S(z)$ is invertible for all $z \in \Gamma$ and $E$ is a rank one bounded operator. Therefore, there are vectors $v$ and $u$ so that $E = uv^*$ and using Sherman-Morrison formula (see [21–23] and the references therein), we write

$$\begin{aligned}(S(z) + \tau E)^{-1} &= S(z)^{-1} - \frac{\tau}{1 + v^* S^{-1}(z) u} S(z)^{-1}(uv^*) S(z)^{-1},\end{aligned} \tag{27}$$

and so it follows that

$$\begin{aligned}&\left(S(z) + \tau_1 E\right)^{-1} - \left(S(z) + \tau_2 E\right)^{-1} \\ &\quad = \frac{\tau_2 - \tau_1}{1 + v^* S^{-1}(z) u} S(z)^{-1}(uv^*) S(z)^{-1} \\ &\quad = \frac{\tau_2 - \tau_1}{1 + v^* S^{-1}(z) u}\left(S(z)^{-1} u\right)\left(S(z)^{-*} v\right)^*.\end{aligned} \tag{28}$$

And in particular $f$ is a smooth function. Since $f$, due to Rouche's theorem, conclude that $f$ takes values only in natural numbers, it must be constant for all $\tau \in [0,1]$. Let us denote this eigenvalue of $\widetilde{S} = S + E$ by $\widetilde{\lambda}$. Define the operators

$$\begin{aligned}A^{(q)} &= \int_\Gamma z^q S^{-1}(z) dz, \\ \widetilde{A}^{(q)} &= \int_\Gamma z^q \widetilde{S}^{-1}(z) dz, \\ &\qquad q = 0, 1.\end{aligned} \tag{29}$$

Proposition 2 implies $A^{(1)} = \lambda x y^*$, $\widetilde{A}^{(1)} = \widetilde{\lambda}\widetilde{x}\widetilde{y}^*$, where $S(\lambda)x = 0$, $S(\lambda)^* y = 0$, $\widetilde{S}(\widetilde{\lambda})\widetilde{x} = 0$, and $\widetilde{S}(\lambda)^*\widetilde{y} = 0$ and so there exists a vector $\phi$ such that $\phi^* A^{(0)}\phi \neq 0$ and $\phi^*\widetilde{A}^{(0)}\phi \neq 0$. We now compute

$$\begin{aligned}\lambda &= \frac{\phi^*\left(\int_\Gamma z S^{-1}(z) dz\right)\phi}{\phi^*\left(\int_\Gamma S^{-1}(z) dz\right)\phi}, \\ \widetilde{\lambda} &= \frac{\phi^*\left(\int_\Gamma z \widetilde{S}^{-1}(z) dz\right)\phi}{\phi^*\left(\int_\Gamma \widetilde{S}^{-1}(z) dz\right)\phi}\end{aligned} \tag{30}$$

and so the second assertion follows by the following computation:

$$\begin{aligned}\left|\lambda - \widetilde{\lambda}\right| &= \left|\frac{\phi^*\left(\int_\Gamma z S^{-1}(z) dz\right)\phi}{\phi^*\left(\int_\Gamma S^{-1}(z) dz\right)\phi} - \frac{\phi^*\left(\int_\Gamma z \widetilde{S}^{-1}(z) dz\right)\phi}{\phi^*\left(\int_\Gamma \widetilde{S}^{-1}(z) dz\right)\phi}\right| \\ &\le \frac{1}{\left|\phi^*\left(\int_\Gamma S^{-1}(z) dz\right)\phi\right|}\left[\left|\phi^*\left(\int_\Gamma z S^{-1}(z) dz\right)\phi - \phi^*\left(\int_\Gamma z \widetilde{S}^{-1}(z) dz\right)\phi\right|\right. \\ &\quad \left. + \left|\frac{\phi^*\left(\int_\Gamma z \widetilde{S}^{-1}(z) dz\right)\phi}{\phi^*\left(\int_\Gamma \widetilde{S}^{-1}(z) dz\right)\phi}\right|\left|\phi^*\left(\int_\Gamma \widetilde{S}^{-1}(z) dz\right)\phi - \phi^*\left(\int_\Gamma S^{-1}(z) dz\right)\phi\right|\right] \\ &\le \frac{1}{\left|\phi^*\left(\int_\Gamma S^{-1}(z) dz\right)\phi\right|}\left[\left|\phi^*\left(\int_\Gamma z S^{-1}(z) - z\widetilde{S}^{-1}(z) dz\right)\phi\right| + \left|\widetilde{\lambda}\right|\left|\phi^*\left(\int_\Gamma \widetilde{S}^{-1}(z) - S^{-1}(z) dz\right)\phi\right|\right]\end{aligned}$$

$$= \frac{1}{\left|\phi^* \left(\int_\Gamma S^{-1}(z)\,dz\right)\phi\right|} \left[\left|\phi^* \left(\int_\Gamma \frac{z}{1+v^* S^{-1}(z)\,u} S(z)^{-1} ES(z)^{-1}\,dz\right)\phi\right|\right.$$

$$\left. + \left|\widetilde{\lambda}\right| \left|\phi^* \left(\int_\Gamma \frac{1}{1+v^* S^{-1}(z)\,u} S(z)^{-1} ES(z)^{-1}\,dz\right)\phi\right|\right] \le C\,\|E\|. \tag{31}$$

The claim on the constant $C$ follows from the observation that $\mathrm{Ran}(E) = \mathrm{span}\{u\}$ and $\mathrm{Ran}(E^*) = \mathrm{span}\{v\}$. □

With this result we can now formulate the main result of this section. It will be used to assess the quality of a given approximation, regardless of its origin.

**Proposition 4.** *Let $\mathfrak{t}(\cdot)$ denote the family of sesquilinear forms (11) with operator representation $T(\cdot) : \Omega \to \mathscr{H}$ given by (15). Assume that a contour $\Gamma \subset \rho(T)$ encloses only a simple eigenvalue $\lambda$ and $\mu \in \mathbb{C}$ but no other points of $\sigma(T)$. Let $u \in \mathscr{V}$, with $\|u\| = 1$. Then there is $\lambda \in \sigma(T)$ such that*

$$|\lambda - \mu| \le C \sup_{\phi \in \mathscr{V}\setminus\{0\}} \frac{|\mathfrak{t}(\mu)[u,\phi]|}{\sqrt{\mathfrak{t}_0[\phi,\phi]}}. \tag{32}$$

*The constant $C$ does not depend on $\mu$ and $u$ but on contour $\Gamma$ and the restriction of $\|T^{-1}(\cdot)\|$ and $\|T^{-*}(\cdot)\|$ to $\Gamma$.*

*Proof.* We will now construct a particular operator $E_{\mu,u}$ which will be used to assess the quality of the pair $\mu$, $u$. Define the sesquilinear form $\mathfrak{e}_{\mu,u} : \mathscr{V} \times \mathscr{V} \to \mathbb{C}$ by the formula

$$\mathfrak{e}_{\mu,u}[\psi,\phi] = -\mathfrak{t}(\mu)[u,\phi](\psi,u), \quad \psi,\phi \in \mathscr{V}. \tag{33}$$

It is obviously relatively compact with respect to $\mathfrak{t}_0$ and so we can define the Fredholm valued operator function $\Omega \ni z \to \widetilde{T}(z)$, where $\widetilde{T}(z)$ is the operator defined by the form

$$\widetilde{\mathfrak{t}}(\lambda)[\psi,\phi] = \mathfrak{t}(\lambda)[\psi,\phi] + \mathfrak{e}_{\mu,u}[\psi,\phi]. \tag{34}$$

Further,

$$\begin{aligned}\left(\widetilde{T}(\mu)u,\phi\right) &= \mathfrak{t}(\mu)[u,\phi] + \mathfrak{e}_{\mu,u}[u,\phi] \\ &= \mathfrak{t}(\mu)[u,\phi] - \mathfrak{t}(\mu)[u,\phi](u,u) \\ &= \mathfrak{t}(\mu)[u,\phi] - \mathfrak{t}(\mu)[u,\phi] = 0, \\ &\qquad \phi \in \mathscr{V},\end{aligned} \tag{35}$$

and so $\mu \in \sigma(\widetilde{T})$. We construct the operator $E_{\mu,u}$ as the operator defined by

$$\left(E_{\mu,u}\phi,\psi\right) = \mathfrak{e}_{\mu,u}\left[T_0^{-1/2}\phi, T_0^{-1/2}\psi\right], \quad \phi,\psi \in \mathscr{H}. \tag{36}$$

Now recall that (22) and (35) imply

$$\widetilde{T}^{-1}(z) = T_0^{-1/2}\left(I + \sum_{j=1}^{d} f_j(z) K_j + E_{\mu,u}\right)^{-1} T_0^{-1/2}. \tag{37}$$

By construction the operator function $\widetilde{S} : \Omega \to \mathscr{L}(\mathscr{H})$

$$\widetilde{S}(z) = I + \sum_{j=1}^{d} f_j(z) K_j + E_{\mu,u} = S(z) + E_{\mu,u} \tag{38}$$

satisfies the assumption of Lemma 3. Finally, note that

$$\begin{aligned}\sup_{\phi \in \mathscr{V}\setminus\{0\}} \frac{|\mathfrak{e}_{\mu,u}(\mu)[u,\phi]|}{\sqrt{\mathfrak{t}_0[\phi,\phi]}} &= \sup_{\phi \in \mathscr{V}\setminus\{0\}} \frac{|\mathfrak{t}(\mu)[u,\phi]|}{\sqrt{\mathfrak{t}_0[\phi,\phi]}} \\ &=: \left\|\mathfrak{t}_{\mu,u}(\mu)[u,\cdot]\right\|_{\mathfrak{t}_0,-1},\end{aligned} \tag{39}$$

and so from (36) it follows that

$$\left\|E_{\mu,u}\right\| = \left\|\mathfrak{t}_{\mu,u}(\mu)[u,\cdot]\right\|_{\mathfrak{t}_0,-1}. \tag{40}$$

Now recall the definition of $C$ from Lemma 3 to conclude the proof. □

*2.3. A Sketch for a Practical Algorithm for Error Estimation.* Proposition 4 will be the basis for practical error estimation. Let $\mathscr{V} \supset \mathscr{Q}_h$, $h > 0$ be a family of finite dimensional spaces such that the orthogonal projections $Q_h$ onto $\mathscr{Q}_h$ converge strongly to $I$ as $h \to \infty$. Let further $\mathscr{Q}_{h_1} \subset \mathscr{Q}_{h_2}$ for $h_1 \le h_2$. Assume that $\phi \in \mathscr{Q}_{h_1}$ and $\mu \in \mathbb{C}$ are given. For $h_2$, $h_2 > h_1$ define

$$\left\|\mathfrak{t}(\mu)[u,\cdot]\right\|_{\mathscr{Q}_{h_2},\mathfrak{t}_0,-1} = \sup_{\phi \in \mathscr{Q}_{h_2}\setminus\{0\}} \frac{|\mathfrak{t}(\mu)[u,\phi]|}{\sqrt{\mathfrak{t}_0[\phi,\phi]}}. \tag{41}$$

Obviously it holds, recalling the definition from (39), that

$$\left\|\mathfrak{t}(\mu)[u,\cdot]\right\|_{\mathscr{Q}_{h_2},\mathfrak{t}_0,-1} \le \left\|\mathfrak{t}_{\mu,u}(\mu)[u,\cdot]\right\|_{\mathfrak{t}_0,-1}. \tag{42}$$

However, in the case in which $\mathscr{Q}_h$, $h > 0$ satisfies the standard saturation property, for example, $\|(I-Q_{h_2})\phi\| \le q\|(I-Q_{h_1})\phi\|$ for fixed $q$, $0 < q < 1$ and some $\phi$, we can also prove

$$\begin{aligned}\left\|\mathfrak{t}(\mu)[u,\cdot]\right\|_{\mathscr{Q}_{h_2},\mathfrak{t}_0,-1} &\le \left\|\mathfrak{t}(\mu)[u,\cdot]\right\|_{\mathfrak{t}_0,-1} \\ &\le C\left\|\mathfrak{t}(\mu)[u,\cdot]\right\|_{\mathscr{Q}_{h_2},\mathfrak{t}_0,-1}.\end{aligned} \tag{43}$$

The constant $C$ essentially depends on $q$ which in turn depends on $h_2 - h_1$ but not on the magnitude of $h_1$; for more details see [24, 25].

## 3. Contour Integration Based Subspace Iteration

In the following section we will present the inexact subspace iteration algorithm based on contour integration and prove basic convergence results using quasi-matrix notation. We consider the spectral transform functions from [5, 6, 26, 27] in the context of eigenvalues of operator-valued analytic functions. Let $T(\cdot)$ be given and let $\Gamma \subset \Omega$ be a closed curve which encloses either a set of $r$ simple eigenvalues or a single semisimple eigenvalue whose multiplicity is $r$.

Let us consider the operators

$$A^{(q)} = \frac{1}{2\pi i}\int_{\Gamma} z^q T(z)^{-1}\,dz,$$
$$B^{(q)} = \frac{1}{2\pi i}\int_{\Gamma} z^q T(z)^{-1} T'(z)\,dz, \qquad q = 0, 1, \tag{44}$$

and their approximations:

$$A_N^{(q)} = \sum_{k=1}^{N} \omega_k z_k^q T(z_k)^{-1},$$
$$B_N^{(q)} = \sum_{k=1}^{N} \omega_k z_k^q T(z_k)^{-1} T'(z_k), \qquad q = 0, 1. \tag{45}$$

Here $\omega_i$ and $z_i \in \Gamma$, $i = 1, \ldots, N$, are integration weights and integration nodes. Based on Theorem 4.7 in [4] we establish, for $\omega_i$ and $z_i \in \Gamma$ defined by the $N$ node trapeze integration formula for the contour integral (44), the following estimates:

$$\left\|A^{(q)} - A_N^{(q)}\right\| \le C_1 d(T)^{-\kappa} e^{-C_2 N d(T)},$$
$$\left\|B^{(q)} - B_N^{(q)}\right\| \le C_1 d(T)^{-\kappa} e^{-C_2 N d(T)}. \tag{46}$$

Here $\Gamma$ is simple closed contour in $\Omega$ such that $\sigma(T) \cap \Gamma = \emptyset$, $d(T) = \min_{\lambda\in\sigma(T)} \operatorname{dist}(\lambda, \Gamma)$, and $\kappa$ is the maximum order of poles for the inverse of $T$. The constants in (46) are in general different and depend on the maximum of the integrand on the contour $\Gamma$. For more details see [4, 7]. Subsequently we conclude that $A_N^{(q)} \to A^{(q)}$ and $B_N^{(q)} \to B^{(q)}$, $q = 0, 1$ in the norm resolvent sense.

*3.1. Extracting Information on Eigenvalues Enclosed by a Contour.* Based on Proposition 2 we see that operators $A^{(0)}$ and $B^{(0)}$ are finite rank operators such that $\operatorname{Ran}(A^{(0)}) = \operatorname{Ran}(B^{(0)})$ is the space spanned by linearly independent eigenvectors associated with semisimple eigenvalues which are enclosed by $\Gamma$. Rather than providing a technical proof of these claims, which will be a subject of subsequent reports, we will present numerical evidence on two judiciously chosen examples. Further, we see that based on [7] we can establish the following technical result for the operators $A^{(1)}$ and $B^{(1)}$.

**Proposition 5.** *Let $T$ be the operator-valued function from (15) and let $\Gamma$ be the contour which encloses solely the $r$, counting according to algebraic multiplicity, semisimple eigenvalues of $T$. Define the matrices $A^{(q)}$ and $B^{(q)}$, $q = 0, 1$, as in (44) and let $Q : \mathbb{C}^r \to \widetilde{\mathcal{H}}$, $Q^*Q = I$, be a quasi-matrix such that $Q^*A^{(0)}Q$ and $Q^*B^{(0)}Q$ are invertible. Then the eigenvalues of the matrix pairs $(Q^*A^{(1)}Q, Q^*A^{(0)}Q)$ and $(Q^*B^{(1)}Q, Q^*B^{(0)}Q)$ are precisely the eigenvalues $\lambda_i$, $i = 1, \ldots, r$, where we count according to multiplicity. Furthermore, if $a_i, b_i \in \mathbb{C}^r$ satisfy $Q^*A^{(1)}Qa_i = \lambda_i Q^*A^{(0)}Qb_i$ and $Q^*B^{(1)}Qa_i = \lambda_i Q^*B^{(0)}Qb_i$, then $Qa_i$ and $Qb_i$ are eigenvectors of $T$ associated with $\lambda_i$, $i = 1, \ldots, r$.*

*Proof.* We will prove the statement for the matrix pair $(Q^*B^{(1)}Q, Q^*B^{(0)}Q)$ and note that the proof for the matrix pair $(Q^*A^{(1)}Q, Q^*A^{(0)}Q)$ is equivalent. Based on Proposition 2 we see that there are quasi-matrices $X : \mathbb{C}^r \to \widetilde{\mathcal{H}}$ and $Y : \mathbb{C}^r \to \widetilde{\mathcal{H}}$ and a neighborhood $\mathcal{U}$ of $\Gamma$ such that

$$T(z)^{-1} = X \begin{bmatrix} (z-\lambda_1)^{-1} & 0\cdots & 0 & \\ 0 & (z-\lambda_2)^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & (z-\lambda_r)^{-1} \end{bmatrix} Y^* + H(z), \quad z \in \mathcal{U}. \tag{47}$$

Here $H$ is an analytic operator-valued function. Now we compute

$$Q^*B^{(1)}Q = (Q^*X) \begin{bmatrix} \lambda_1 & 0\cdots & 0 & \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \lambda_r \end{bmatrix} (Y^*Q)$$
$$Q^*B^{(0)}Q = (Q^*X)(Y^*Q). \tag{48}$$

Based on the assumptions of the theorem we conclude that $(Q^*X)$ and $(Y^*Q)$ have to be invertible and so the conclusion readily follows. □

Before we proceed, note that norm resolvent convergence of $A_N^{(q)}$ and $B_N^{(q)}$ implies the convergence of spectra and associated spaces to those of $A^{(q)}$ and $B^{(q)}$, $q = 0, 1$. In particular, this means that $A_N^{(0)}$ and $B_N^{(0)}$, for $N$ large enough, will have two well separated components in the spectrum and so we are motivated to use subspace iteration on the operator level to improve the quality of the approximate eigenvalues. Note that for a vector $v$ the vectors $A_N^{(0)}v$ and $B_N^{(0)}v$ can be computed using formula (45) without ever forming a representation for the underlying operator.

*3.2. Inexact Subspace Iteration Based on Quadrature Formula.* We have the following algorithm for a generic bounded

Pick a random quasi-matrix $Q_1 = [q_1 \quad q_2 \quad \cdots \quad q_r]$, $Q_1 : \mathbb{C}^r \to L^2(\Omega)$
Set $n = 1$
**repeat**
  Compute $Y_{n+1} = B_N Q_n$
  Compute an orthogonal quasi-matrix $Q_{n+1}$, $Q_{n+1}^* Q_{n+1} = I$ using Gram-Schmidt so that $\mathrm{Ran}(Q_{n+1}) = \mathrm{Ran}(Y_{n+1})$.
  Compute `criterion` $= \|Q_{n+1} - Q_n(Q_n^* Q_{n+1})\|$
  $n := n + 1$
**until** `criterion` < `tol`
**return** $Q_{n+1}$

Algorithm 1: Contour based inexact subspace iteration for $B$, `tol`, and $N$.

**Require** $Q_1 : \mathbb{C}^r \to \widetilde{\mathscr{H}}$, $Q_1^* Q_1 = I_r$, $\mathtt{tol}_1$ and $N$.
  Compute $Q$ as output from Algorithm 1 with tolerance $\mathtt{tol}_1$ and $Q_1$ as starting value.
  Compute $Z_1 = B_N^{(1)} Q$ and $Z_0 = B_N^{(0)} Q$.
  Form $W^{(q)} = Q^* Z_q$, $q = 0, 1$
  Compute $v_i$ and $\lambda_i$ so that $W^{(1)} u_i = \lambda_i W^{(0)} u_i$, $i = 1, \ldots, r$
  Compute eigenvectors $v_i = Q_{n+1} u_i$, $i = 1, \ldots, r$.
  Compute $\|\mathfrak{t}(\lambda_i)[v_i, \cdot]\|_{\mathscr{Q}_{h_2}, \mathfrak{t}_0, -1}$
  Form the index set $\{i_1, \ldots, i_{r_2}\} = \{i : \|\mathfrak{t}(\lambda_i)[v_i, \cdot]\|_{\mathscr{Q}_{h_2}, \mathfrak{t}_0, -1} \le \mathtt{tol}_2\}$
  Set $V = \begin{bmatrix} v_{i_1} & \cdots & v_{i_{r_2}} \end{bmatrix}$
**return** $V : \mathbb{C}^{r_2} \to \widetilde{\mathscr{H}}$ and $\lambda_{i_j}$, $j = 1, \ldots, r_2$

Algorithm 2: Approximating eigenvectors and eigenvalues with tolerance $\mathtt{tol}_2$.

operator $B$ of type (44) which has a component of $r$, $r \in \mathbb{N}$ dominant eigenvalues and whose action on a vector can be represented by a formula (45).

For Algorithm 1 we present standard convergence results in Theorems 6 and 8. First we consider a special case in which $B$ is a Hermitian (bounded self-adjoint) operator. We will see that a possible presence of singularities in the operator function will not pollute the convergence rate.

**Theorem 6.** *Let $B$ be a bounded Hermitian operator such that $\sigma(B) = \{\lambda_1, \ldots, \lambda_r\} \cup \Sigma$, where $\lambda_i$, $i = 1, \ldots, r$, are eigenvalues of finite multiplicity and $|\lambda_1| \ge \cdots \ge |\lambda_r| > \tau > 0$, and $\tau > |\lambda|$, $\lambda \in \Sigma$. Then for the sequence $V^{(n+1)} := BV^{(n)}$, where $V^{(0)} : \mathbb{C}^r \to \mathscr{H}$, $(V^{(0)})^* V^{(0)} = I_r$, we have*

$$\tan\theta\left(V^{(n+1)}, U\right) \le \frac{\tau^n}{|\lambda_r|^n} \tan\theta\left(V^{(0)}, U\right), \tag{49}$$

*where $U : \mathbb{C}^r \to \mathscr{H}$, $U^* U = I$, is such that $U = [u_1 \quad u_2 \quad \cdots \quad u_r]$ and $Bu_i = \lambda_i u_i$, $i = 1, \ldots, r$.*

*Proof.* Let $P_1 = UU^*$ and represent $U : \mathbb{C}^r \to \mathscr{H}$ with respect to the partition of unity $P_1 \oplus P_2 = I$ by $U = [U_1 \quad U_2]$, with $\mathrm{Ran}(P_1) = \mathrm{Ran}(U_1)$. Since $B$ is self-adjoint it can be represented by the block operator matrix:

$$B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}. \tag{50}$$

Let us represent, with respect to the same partition of unity, the quasi-matrix $V^{(0)}$ as

$$V^{(0)} = \begin{bmatrix} V_1^{(0)} \\ V_2^{(0)} \end{bmatrix}. \tag{51}$$

Without reducing the level of generality we may assume that $V_1^{(0)}$ is invertible, since otherwise $\mathrm{Ran}(V^{(0)})$ would contain eigenvectors of $B$. This corresponds to the trivial situation and will not be further discussed.

Since $(V^{(0)})^* V^{(0)} = I_r$ it follows from (6) that

$$\tan\theta\left(U_1, V^{(0)}\right) = \left\| V_2^{(0)} \left(V_1^{(0)}\right)^{-1} \right\|. \tag{52}$$

The assumption of the separation of the spectra of $B$ implies that $B_{11}$ must also be invertible, and so

$$V^{(n)} := B^n V^{(0)} = \begin{bmatrix} B_{11}^n V_1^{(0)} \\ B_{22}^k V_2^{(0)} \end{bmatrix} = \begin{bmatrix} I \\ F^{(n)} \end{bmatrix} B_{11}^n V_1^{(0)} \tag{53}$$

for $F^{(n)} = B_{22}^n V_2^{(0)} (V_1^{(0)})^{-1} B_{11}^{-n}$ which implies that

$$V^{(n)} \left(V_1^{(0)}\right)^{-1} B_{11}^{-n} = \begin{bmatrix} I \\ F^{(n)} \end{bmatrix} = U_1 + U_2 F^{(n)}. \tag{54}$$

And so from the optimality of the canonical angles (see [15]), it follows that

$$\tan\theta\left(V^{(n)}, U\right) \le \left\| F^{(n)} \right\|. \tag{55}$$

Consequently, it follows that

$$\begin{aligned}\tan\theta\left(V^{(n)},U\right) &\le \left\|F^{(n)}\right\| \le \left\|B_{22}^{n}\right\|\left\|F^{(0)}\right\|\left\|B_{11}^{-n}\right\| \\ &\le \frac{\tau^{n}}{\left|\lambda_{r}\right|^{n}}\tan\theta\left(V^{(0)},U_{1}\right).\end{aligned} \tag{56}$$

Now the conclusion of the theorem readily follows. □

*Remark 7.* Note that from (8) and (9) it is clear that $\theta(V^{(n)},U)$ does not depend on the quasi-matrices $V^{(n)}$ and $U$, but rather on the orthogonal projections onto $\mathrm{Ran}(V^{(n)})$ and $\mathrm{Ran}(U) = \mathrm{Ran}(P_1)$. In Theorem 6 we have stated the convergence result for the angle $\theta(V^{(n)},U)$. Here we have ignored the choice of a basis of $\mathrm{Ran}(V^{(n)})$ which is a very important part of Algorithm 1. Note that in practical computations the choice of the basis is crucial to achieve an efficient implementation; see [28].

In the case in which the operator $B$ is not Hermitian but has clearly separated invariant subspace associated with the finite collection, counting according to the algebraic multiplicity of dominant eigenvalues we have the theorem below.

**Theorem 8.** *Let $B$ be a bounded operator such that its spectrum has two disjoint components $\Sigma_1$ and $\Sigma_2$ so that $\sigma(B) = \Sigma_1 \cup \Sigma_2$. Let further $\Sigma_1 = \{\lambda_1,\lambda_2,\ldots,\lambda_r\}$, where we have counted $\lambda_i \in \sigma(B)$ according to the algebraic multiplicity of an eigenvalue. Further let a partition of unity $P_1 \oplus P_2 = I$ exist such that $r = \dim(\mathrm{Ran}(P_1))$ and*

$$B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22}\end{bmatrix}, \tag{57}$$

*and let further $sep(B_{11},B_{22}) > 1$ and $\|B_{22}\|\|B_{11}^{-1}\| < 1$. Then for a quasi-matrix $V:\mathbb{C}^r \to \mathscr{H}$ we have*

$$\begin{aligned}&\tan\theta\left(B^{n}V,P_{1}\right) \\ &\quad\le \left(\left\|B_{22}\right\|\left\|B_{11}^{-1}\right\|\right)^{n}\left(1+\|G\|\right)\tan\theta\left(V,P_{1}\right).\end{aligned} \tag{58}$$

*Here $G$ is the solution of the Sylvester equation $GB_{22} - B_{11}G = B_{12}$ and $sep(B_{11},B_{22}) = \inf_{\|G\|=1}\|GB_{22} - B_{11}G\|$.*

*Proof.* For $B$ such that $\sigma(B) = \Sigma_1 \cup \Sigma_2$ with $\Sigma_1 = \{\lambda_1,\lambda_2,\ldots,\lambda_r\}$ satisfying the assumptions of the theorem choose $P_1$ as the orthogonal projection onto the maximal invariant subspace associated with $\{\lambda_1,\lambda_2,\ldots,\lambda_r\}$ and $P_2 = I - P_1$; then $B$ has the block operator representation as claimed in the theorem. We note that

$$B = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22}\end{bmatrix} = \begin{bmatrix} I & -G \\ 0 & I\end{bmatrix}\begin{bmatrix} B_{11} & 0 \\ 0 & B_{22}\end{bmatrix}\begin{bmatrix} I & G \\ 0 & I\end{bmatrix}, \tag{59}$$

where $GB_{22} - B_{11}G = B_{12}$. Operator $G$, such that $GB_{22} - B_{11}G = B_{12}$ exists and is unique, for example, [29, Lemma A.1], providing the spectra of $B_{11}$ and $B_{22}$, are disjoint. The rest of the proof follows analogously as in Theorem 6. □

*Remark 9.* In [29, Lemma A.1] there is also an estimate of $\mathrm{sep}(B_{11},B_{22})$ in terms of the pseudospectra distances.

This result implies that the convergence rate for the subspace iteration essentially depends on the properties of the operator and not on the subspace which is iterated. Further, the dependence on the number of integration nodes $N$ diminishes exponentially. To extract eigenvalues and eigenvectors we use the method from [7] and apply Proposition 5 on the quasi-matrix $Q_{n+1}$ which is returned by Algorithm 1. More precisely we have the following. For simplicity we write $B = B^{(0)}$.

**Theorem 10.** *Let the conditions of Theorem 8 be satisfied and also assume the same notation conventions. For a quasi-matrix $V:\mathbb{C}^r \to \mathscr{H}$, we set $Q_1 := V$ and apply Algorithm 1 with fixed $N \in \mathbb{N}$. Then*

$$\begin{aligned}&\sin\theta\left(\mathrm{Ran}\left(Q_{n}\right),P_{1}\right) \\ &\quad\le \left(\left\|B_{22}\right\|\left\|B_{11}^{-1}\right\| + C_{3}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}\right)^{n} \\ &\quad\cdot\left(1+\|G\|+C_{4}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}\right)\tan\theta\left(Q_{1},P_{1}\right) \\ &\quad+ C_{5}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}.\end{aligned} \tag{60}$$

*Proof.* Recall that

$$\left\|B - B_{N}\right\| \le C_{1}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}. \tag{61}$$

From standard results on norm convergence of bounded operators (see [16]) we conclude that

$$\begin{aligned}\sin\theta\left(P_{1},\left(P_{N}\right)_{1}\right) &= \left\|P_{1} - \left(P_{N}\right)_{1}\right\| \\ &\le C_{5}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}.\end{aligned} \tag{62}$$

Here $(P_N)_1 + (P_N)_2 = I$ is the partition of unity in which $B_N$ has a block triangular form:

$$B_{N} = \begin{bmatrix} \left(B_{N}\right)_{11} & \left(B_{N}\right)_{12} \\ 0 & \left(B_{N}\right)_{22}\end{bmatrix}. \tag{63}$$

Assumption on the spectral separation from Theorem 8 implies that $B_{11}$ is invertible and obviously

$$\left\|B_{11} - \left(B_{N}\right)_{11}\right\| \le C_{1}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)} \tag{64}$$

$$\left\|B_{22} - \left(B_{N}\right)_{22}\right\| \le C_{1}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}. \tag{65}$$

Norm convergence (64) implies that for $N$ large enough $(B_N)_{11}$ must also be invertible; for details see [30]. For such $N$ we have

$$\begin{aligned}\left\|B_{11}^{-1} - \left(B_{N}\right)_{11}^{-1}\right\| &\le \left\|\left(B_{11}\right)^{-1}\left(B_{11} - \left(B_{N}\right)_{11}\right)\left(B_{N}\right)_{11}^{-1}\right\| \\ &\le \left\|\left(B_{11}\right)^{-1}\right\|\left\|\left(B_{N}\right)_{11}^{-1}\right\|\left\|B_{11} - \left(B_{N}\right)_{11}\right\| \\ &\le C_{3}d\left(T\right)^{-\kappa}e^{-C_{2}Nd(T)}.\end{aligned} \tag{66}$$

Triangle inequality implies that

$$\begin{aligned} \left\|(B_N)_{22}\right\| &\le \left\|B_{22}\right\| + \left\|(B_N)_{22} - B_{22}\right\|, \\ \left\|(B_N)_{11}^{-1}\right\| &\le \left\|B_{11}^{-1}\right\| + \left\|(B_N)_{11}^{-1} - B_{11}^{-1}\right\| \end{aligned} \tag{67}$$

and so

$$\begin{aligned} \left\|(B_N)_{22}\right\| \left\|(B_N)_{11}^{-1}\right\| &\le \left\|B_{22}\right\| \left\|B_{11}^{-1}\right\| \\ &\quad + C_4 d(T)^{-\kappa} e^{-C_2 N d(T)}. \end{aligned} \tag{68}$$

Let us now assume that $N$ is large enough so that

$$\left\|B_{22}\right\| \left\|B_{11}^{-1}\right\| + C_4 d(T)^{-\kappa} e^{-C_2 N d(T)} < 1. \tag{69}$$

Then we can apply Theorem 8 to $B_N$ and conclude that

$$\begin{aligned} \sin\theta\left(B_N^n V, (P_N)_1\right) &\le \tan\theta\left(B_N^n V, (P_N)_1\right) \\ &\le \left(\left\|(B_N)_{22}\right\| \left\|(B_N)_{11}^{-1}\right\|\right)^n \left(1 + \left\|G_N\right\|\right) \\ &\quad \cdot \tan\theta\left(V, (P_N)_1\right). \end{aligned} \tag{70}$$

Equivalently since $\operatorname{sep}((B_N)_{22}, (B_N)_{11}) \to \operatorname{sep}(B_{22}, B_{11})$ and $(B_N)_{12} \to B_{12}$ in norm we conclude that

$$\left\|G_N\right\| \le \|G\| + C_5 d(T)^{-\kappa} e^{-C_2 N d(T)}. \tag{71}$$

We now apply the triangle inequality for the sine of the maximal canonical angle to conclude the proof. □

Using Theorem 10 we will define the *effective convergence rate* of the inexact contour based subspace iteration. Recall that

$$\begin{aligned} B &= \frac{1}{2\pi i} \int_\Gamma T(z)^{-1} T'(z)\, dz, \\ B_N &= \sum_{k=1}^{N} \omega_k z_k^q T(z_k)^{-1} T'(z_k). \end{aligned} \tag{72}$$

Starting from (63) we define the *effective convergence rate* for the inexact subspace iteration for $B$ as

$$\eta(\Gamma, N) := \left\|(B_N)_{22}\right\| \left\|(B_N)_{11}^{-1}\right\|. \tag{73}$$

*Remark 11.* We will project $B_N$ on carefully constructed finite dimensional spaces, for example, finite element spaces, to experimentally estimate $\eta(\Gamma, N)$ in the following section. We will not further elaborate on this procedure but solely present the results of experiments for illustration purposes.

*3.3. Auxiliary Subspace Error Estimates and the Convergence Criterion.* In this paper we will use the inequality $\|\mathfrak{t}(\mu)[u,\cdot]\|_{\mathcal{Q}_{h_2},\mathfrak{t}_0,-1} \le \|\mathfrak{t}(\mu)[u,\cdot]\|_{\mathfrak{t}_0,-1}$ to filter the eigenpairs which have been extracted using Proposition 5 from the subspace returned by Algorithm 1. More precisely, eigenpairs for which $\|\mathfrak{t}(\mu)[u,\cdot]\|_{\mathcal{Q}_{h_2},\mathfrak{t}_0,-1}$ is too large will be discarded. We will indicate the importance of this step of the algorithm in the experiments section.

In the case in which the number of eigenvalues returned by Algorithm 2 is not satisfactory, the procedure can be extended as an innerouter iteration scheme by setting $Q_1 = V$ and repeating the procedure. One further possibility to improve performance is to modify Algorithm 1 so that at the beginning of the repeat loop we remove all those directions from $Q_1$ which are identified by Algorithm 2 as converged and then proceed as in the original versions of Algorithm 1.

## 4. Numerical Experiments

For the finite element discretizations, we use the space of piecewise linear and continuous finite elements on a given subdivision $\delta_h$ of an interval $[a,b]$. We denote this space by $\mathcal{V}_h(a,b)$. For our computations we set the tolerance `tol` so as to balance the error when solving the source problem by finite element approximation with the error in the integration. For the contour $\Gamma$ we chose a circle and as integration nodes and weights we use the trapeze formula from [7] and Gauss-Legendre nodes and weights from [5].

To estimate the approximation errors in our experiments we will use an auxiliary subspace error estimation technique. To this end we use $\mathcal{Q}_h(a,b)$ to denote the space of piecewise quadratic and continuous function on the same subdivision of the interval $[a,b]$ which was used to define the space $\mathcal{V}_h(a,b)$. To estimate the error $E_{\mu,u}$ from (39) for the function $u \in \mathcal{V}_h(a,b)$ and $\mu \in \mathbb{C}$ we use the formula; see [24] for further information:

$$\left\|\mathfrak{t}(\mu)[u,\cdot]\right\|_{\mathcal{Q}_h(a,b),\mathfrak{t}_0,-1} = \sup_{\phi \in \mathcal{Q}_h(a,b)\setminus\{0\}} \frac{\left|\mathfrak{t}(\mu)[u,\phi]\right|}{\sqrt{\mathfrak{t}_0[\phi,\phi]}}. \tag{74}$$

*Example 12.* We study the quadratic eigenvalue problem:

$$\begin{aligned} \partial_{xx} u &= \lambda\left(\gamma \partial_{xx} u + \delta u\right) + \lambda^2 u \\ u(-2) &= u(2) = 0 \end{aligned} \tag{75}$$

for $\gamma = 0.05$ and $\delta = 0.3$. As a benchmark we use the eigenvalue computed with a spectral discretization using the `chebfun` system; see [10]. The timings are 2.5 seconds for `chebfun` and 1.6 seconds for subspace iteration. We used the trapeze formula to define the operator $B_N$ with $N = 25$ and achieved the residuals as small as `1e-12`. The results are presented in Figure 1. From Figure 1 we see that the spectrum of (75) is well separated, and so the sufficient separation of the spectra of $(B_N)_{11}$ and $(B_N)_{22}$ can be achieved with few integration nodes $N$: for example, $N = 25$. Figure 2(a) depicts the dependence of the effective convergence rate for the inexact subspace iteration $\eta(\Gamma, N) = \|(B_N)_{22}\|\|(B_N)_{11}^{-1}\|$ on the number of integration nodes $N$ for the trapeze formula. The contour was a circle which enclosed the five eigenvalues marked by crosses.

*Remark 13.* Note that in [5, 6] it was shown that Gauss-Legendre and trapeze integration require similar number of integration nodes to be applicable in inexact subspace iteration algorithm. A slight experimental advantage was noticed in favor of Gauss-Legendre, but both approaches

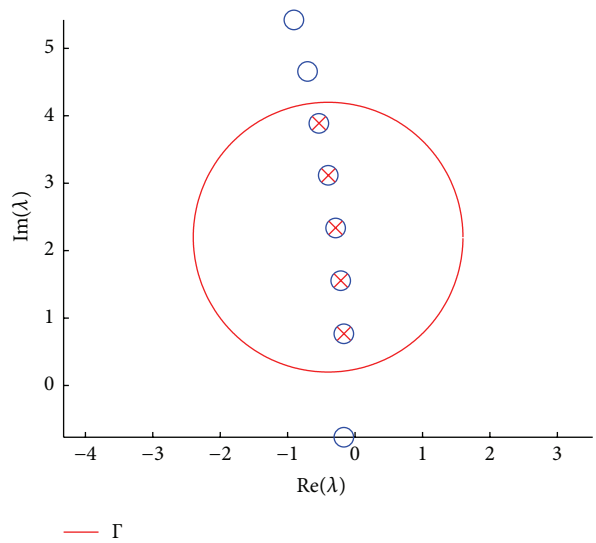
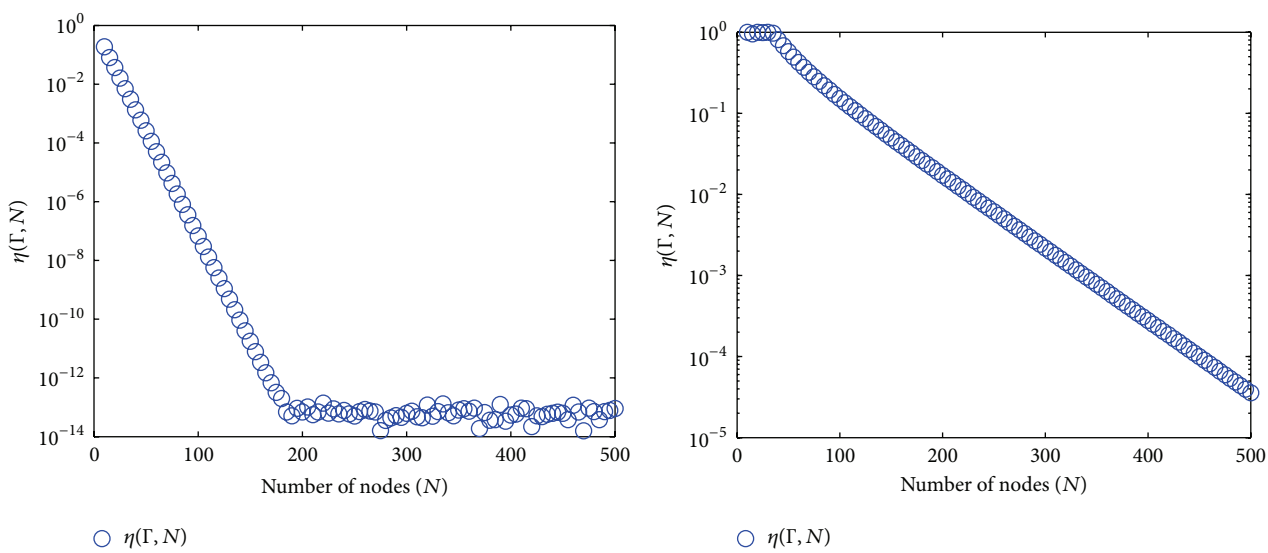



FIGURE 1: Eigenvalues for the quadratic eigenvalue problem.




(a) Quadratic eigenvalue problem with well-separated spectrum from Example 12 and Γ as in Figure 1

(b) Rational eigenvalue problem with tightly clustered eigenvalues from Example 14 and Γ as in Figure 4

FIGURE 2: The dependence of the effective convergence rate $\eta(\Gamma, N)$ as defined in (73) on the number of integration nodes $N$. This illustrates the negative effect of the distance to Γ on the convergence rate.

can be used in competitive algorithms depending on the particular context of the application.

*Example 14.* We will now consider the following rational eigenvalue problem:

$$\frac{d^2}{dx^2}u + \frac{\lambda}{1-\lambda}\chi_{|x|\le 9/10}\cdot u + \frac{\lambda}{3-\lambda}\chi_{|x|\le 9/10}\cdot u + \lambda\left(\chi_{|x|\le 9/10}u + 10\chi_{|x|>9/10}u\right) = 0 \qquad (76)$$

$$u \in H^1(-1,1) \text{ with periodic b.c.,}$$

where

$$\chi_{|x|\le 9/10} : x \longmapsto \begin{cases} 0: & x \notin \left[-\frac{9}{10}, \frac{9}{10}\right] \\ 1: & x \in \left[-\frac{9}{10}, \frac{9}{10}\right] \end{cases} \qquad (77)$$

is the characteristic function of the interval $[-9/10, 9/10]$.

This eigenvalue problem has a singularity at $\lambda = 1$ and $\lambda = 3$. The eigenvalues below the singularity $\lambda = 1$ and those in $\langle 1, 3\rangle$ are discrete (finite multiplicity) and they accumulate at $\lambda = 1$ and at $\lambda = 3$ from below [31].

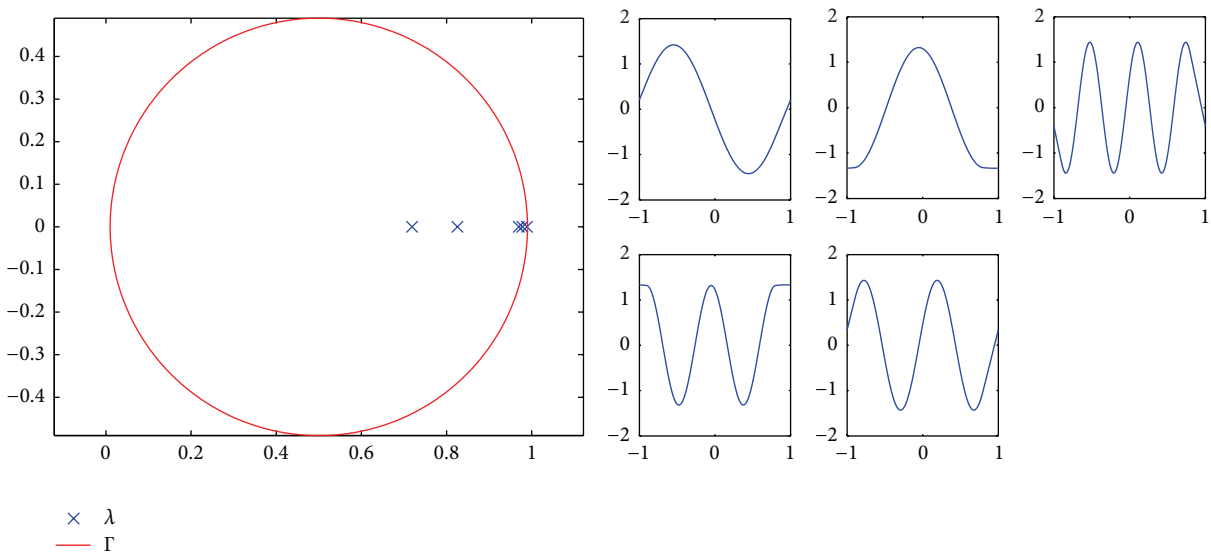



Figure 3: Eigenvectors for eigenvalues in $[0, 1\rangle$.

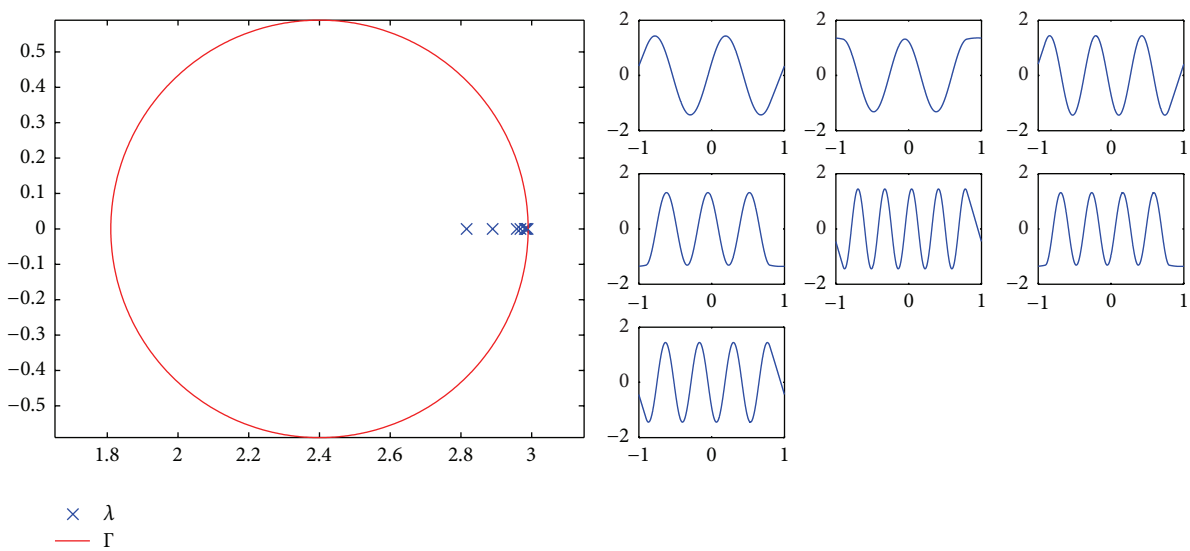



Figure 4: Eigenvectors for eigenvalues in $\langle 1, 3\rangle$.

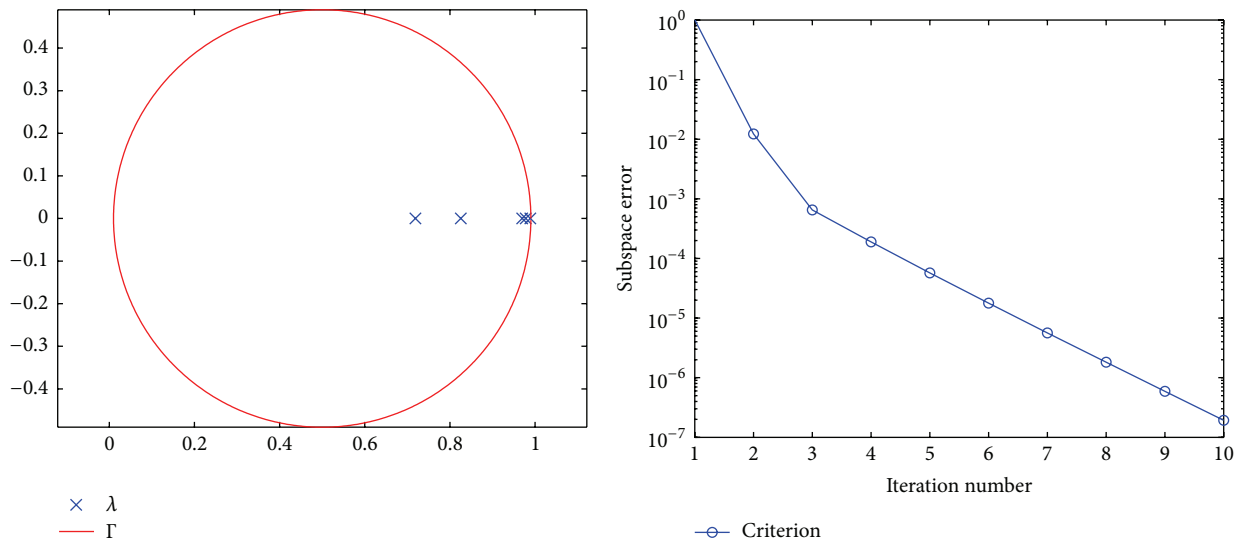
The results in Figures 3 and 4 are presented solely for benchmark purposes. They were computed using trapeze formula with 4500 nodes and validated using the approach from [8]. Looking at Figures 3 and 4 we see that in general there are two classes of eigenfunctions: those that are confined to the interval $[-9/10, 9/10]$ and those that are not. Further discussion of the qualitative properties of eigenfunctions is outside the scope of this paper. We will address the modeling aspects elsewhere.

As a first experiment in Table 1 we present the empirical study of the convergence of the residual measures and the relative eigenvalue error computed from the sequence $V^{(n+1)} = B_N V^{(n)}$, $V_0 : \mathbb{C}^2 \to \mathscr{V}$, where $N = 16$ and $\omega_i$ and $z_i$ are Gauss-Legendre integration nodes and weights from [5]. We have used the contour which encloses only two well-separated eigenvalues, two left most crosses, in Figure 3.

Table 1: Convergence history for the rational eigenvalue problem Example 14 and the two leftmost eigenvalues from Figure 3.

| Iteration number $n$ | Maximal eigenvalue error | Maximal residual estimate |
|---|---|---|
| 1 | $2.0880e-008$ | $1.2607e-009$ |
| 2 | $4.9292e-009$ | $1.3579e-010$ |
| 3 | $4.9282e-009$ | $1.3580e-010$ |

The convergence rate for the well-separated eigenvalues and the standard choice of the integration nodes from [6] seem satisfactory. However, the eigenvalues of (76) cluster towards $z = 1$. In Figure 2(b) we see that many more integration nodes are necessary to achieve a similar estimate of the effective convergence rate $\eta(\Gamma, N)$ compared to the




(a) Eigenvalues clustering towards Γ

(b) Decay of $\|Q_{n+1} - Q_n(Q_n^* Q_{n+1})\|$ for the rational eigenvalue problem and $N = 350$

FIGURE 5: We display $\|Q_{n+1} - Q_n(Q_n^* Q_{n+1})\|$ for $n = 1, \ldots, 10$.

quadratic eigenvalue problem in Figure 2(a). To this end we will compare below the measured convergence rates for the criterion in Algorithm 1 on Examples (76) and (75).

*Remark 15.* Results presented in Figure 5 show that, unlike what was suggested in [6], the number of integration nodes might have to be adaptively adjusted for some contours Γ. Also we emphasize that the residual criterion

$$\left\| \mathtt{t}(\mu)[u, \cdot] \right\|_{\mathcal{C}_h(a,b), \mathtt{t}_0, -1} \leq \mathtt{tol}_2 \tag{78}$$

in Algorithm 2 is particularly important in the case of the clustering of eigenvalues towards Γ like in Figures 3 and 4. In comparison with benchmark results from Figures 3 and 4 three spurious eigenvalues, which would otherwise have been declared as converged, were discarded because their residuals were larger than a threshold. With this in mind we justify the application of the modified algorithm from [7]. The algorithm in [7] had residual error control but controlled the accuracy solely by increasing the number of integration nodes (e.g., we used 4500 nodes for benchmark results in Figures 3 and 4). In our modification we combine subspace iteration with contour integration with problem adapted number of nodes (e.g., with $N = 16$ and 3 steps of inexact subspace iteration acceleration we reach the same accuracy (see Table 1) as a priori predicted for $N = 4500$ based on [8]). Finally, we note that performance issues are outside the scope of this paper. As is indicated in [6] the algorithm offers large potential to leverage parallel processing. This will be the topic of further research.

*4.1. Stability of Convergence Rate of Inexact Subspace Iteration.* From Figure 2 we see that the convergence rate for the subspace iteration essentially depends on $\|B_{22}\|\|B_{11}^{-1}\| \approx \eta(\Gamma, N)$. We will now see that this convergence rate is relatively robust

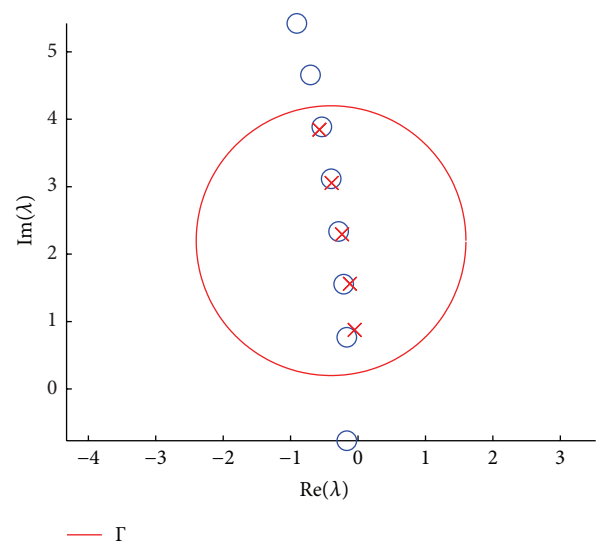
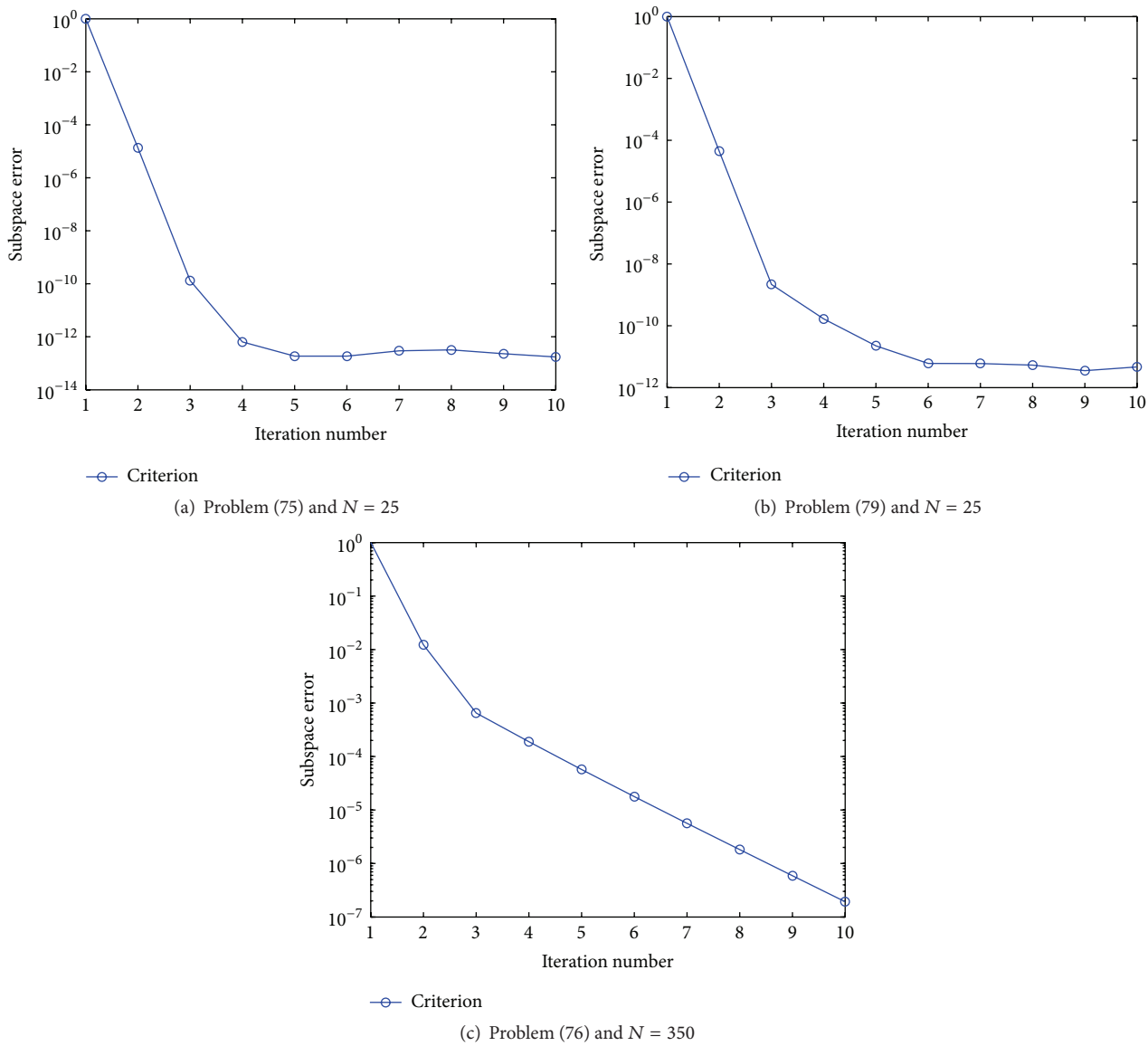



FIGURE 6: Eigenvalues for the eigenvalue problem (79).

to perturbations as long as the distance from the curve Γ and $\sigma(T)$ is not too small. To this end we consider the problem

$$\begin{aligned} \partial_{xx} u &= \lambda(\gamma \partial_{xx} u + \delta u) + \lambda^2 u + \frac{1}{4}\exp(-\lambda) \\ u(-2) &= u(2) = 0 \end{aligned} \tag{79}$$

as a perturbation of (75). Its eigenvalues are presented on Figure 6. We see that both the eigenvalues of (75) and (79) are equally well separated from Γ and so the convergence rates, as measured by the decay rate of $\|Q_{n+1} - Q_n(Q_n^* Q_{n+1})\|$ in Algorithm 1, are similar. However, when comparing the results from Figure 7(c) with those from Figure 7(a) we see that we need more iterations to achieve a comparable reduction in the convergence criterion for Algorithm 1 in the




(a) Problem (75) and $N = 25$

(b) Problem (79) and $N = 25$

(c) Problem (76) and $N = 350$

FIGURE 7: We display $\|Q_{n+1} - Q_n(Q_n^* Q_{n+1})\|$ for $n = 1, \ldots, 10$. For $\Gamma$ we have used the contours as in Figures 1, 6, and 5(a), respectively.

case of eigenvalue clustering towards $\Gamma$. On the other hand, Figure 7(b) indicates that the nonlinear perturbation did not change the convergence rate significantly. This justifies the consideration of the adaptivity both in the choice of integration nodes and in representing the operators.

## 5. Conclusion

We have shown that the subspace iteration nonlinear eigensolver based on spectral transformation of the analytic Fredholm valued function converges at rates that depend primarily on the problem and not on the discretization. Further, we have seen that the distance from the contour and the spectrum does limit the accuracy of the approximation based on numerical integration of the resolvent. Also, residual norm is a reliable estimator of the approximation error even in the presence of poles. In comparison, the convergence in the case of the quadratic eigenvalue problem which had a more pronounced spectral gap than the rational problem was much faster and more robust. This suggests that a subspace iteration algorithm that increases the number of integration nodes based on a posteriori computable criterion is promising for nonlinear eigenvalue problems.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

Christian Engström gratefully acknowledges the support of the Swedish Research Council under Grant no. 621-2012-3863. The work of Luka Grubišić has in part been supported

by the Croatian Research Foundation Grant HRZZ-9345 and the MZOS-NSF grant "Estimates for Finite Element Approximation Error by Auxiliary Subspace Method." Luka Grubišić is thankful to J. Ovall, Portland State University, for the hospitality during the initial phase of the work on this paper as well as for many helpful discussions.

## References

[1] C. Effenberger, D. Kressner, and C. Engström, "Linearization techniques for band structure calculations in absorbing photonic crystals," *International Journal for Numerical Methods in Engineering*, vol. 89, no. 2, pp. 180–191, 2012.

[2] K. F. Gurski, R. Kollár, and R. L. Pego, "Slow damping of internal waves in a stably stratified fluid," *Proceedings A: Mathematical, Physical and Engineering Sciences*, vol. 460, no. 2044, pp. 977–994, 2004.

[3] P. J. Schmid and D. S. Henningson, *Stability and Transition in Shear Flows*, vol. 142 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 2001.

[4] W.-J. Beyn, "An integral method for solving nonlinear eigenvalue problems," *Linear Algebra and Its Applications*, vol. 436, no. 10, pp. 3839–3863, 2012.

[5] E. Polizzi, "Density-matrix-based algorithm for solving eigenvalue problems," *Physical Review B: Condensed Matter and Materials Physics*, vol. 79, no. 11, Article ID 115112, 2009.

[6] P. T. P. Tang and E. Polizzi, "Feast as a subspace iteration eigensolver accelerated by approximate spectral projection," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 2, pp. 354–390, 2014.

[7] W.-J. Beyn, Y. Latushkin, and J. Rottmann-Matthes, "Finding eigenvalues of holomorphic Fredholm operator pencils using boundary value problems and contour integrals," *Integral Equations and Operator Theory*, vol. 78, no. 2, pp. 155–211, 2014.

[8] L. Grubišić and A. Grbić, "Discrete perturbation estimates for eigenpairs of fredholm operator-valued functions," *Applied Mathematics and Computation*, vol. 267, pp. 632–647, 2015.

[9] R. Mennicken and M. Möller, *Non-Self-Adjoint Boundary Eigenvalue Problems* , vol. 192 of *North-Holland Mathematics Studies*, North-Holland Publishing Company, Amsterdam, The Netherlands, 2003.

[10] R. B. Platte and L. N. Trefethen, "Chebfun: a new kind of numerical computing," in *Progress in Industrial Mathematics at ECMI 2008*, vol. 15 of *Mathematics in Industry*, pp. 69–87, Springer, Berlin, Germany, 2010.

[11] A. Townsend and L. N. Trefethen, "Continuous analogues of matrix factorizations," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2173, 2015.

[12] L. N. Trefethen, "Householder triangularization of a quasimatrix," *IMA Journal of Numerical Analysis*, vol. 30, no. 4, pp. 887–897, 2010.

[13] C. Tretter, *Spectral Theory of Block Operator Matrices and Applications*, Imperial College Press, London, UK, 2008.

[14] P. R. Halmos, "Two subspaces," *Transactions of the American Mathematical Society*, vol. 144, pp. 381–389, 1969.

[15] C. Davis and W. M. Kahan, "The rotation of eigenvectors by a perturbation. III," *SIAM Journal on Numerical Analysis*, vol. 7, no. 1, pp. 1–46, 1970.

[16] T. Kato, *Perturbation Theory for Linear Operators*, Classics in Mathematics, Springer, Berlin, Germany, 1995, Reprint of the 1980 edition.

[17] I. C. Gohberg and E. I. Sigal, "Global factorization of a meromorphic operator-function and some of its applications," *Matematicheskie Issledovaniya*, vol. 6, no. 1, pp. 63–82, 1971.

[18] I. C. Gohberg and E. I. Sigal, "An operator generalization of the logarithmic residue theorem and the theorem of Rouché," *Matematicheskii Sbornik*, vol. 84, no. 126, pp. 607–629, 1971.

[19] I. Gohberg and E. I. Sigal, "The root multiplicity of the product of meromorphic operator functions," *Istoriko-Matematicheskie Issledovaniya*, vol. 6, pp. 33–50, 1971.

[20] D. Bindel and H. Amanda, "Localization theorems for nonlinear eigenvalue problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 4, pp. 1728–1749, 2013.

[21] W. W. Hager, "Updating the inverse of a matrix," *SIAM Review*, vol. 31, no. 2, pp. 221–239, 1989.

[22] R. Harte, *Invertibility and Singularity for Bounded Linear Operators*, vol. 109 of *Monographs and Textbooks in Pure and Applied Mathematics*, Marcel Dekker, New York, NY, USA, 1988.

[23] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *The Annals of Mathematical Statistics*, vol. 21, no. 1, pp. 124–127, 1950.

[24] R. E. Bank, "Hierarchical bases and the finite element method," *Acta Numerica*, vol. 5, pp. 1–43, 1996.

[25] J. S. Ovall, "Function, gradient, and Hessian recovery using quadratic edge-bump functions," *SIAM Journal on Numerical Analysis*, vol. 45, no. 3, pp. 1064–1080, 2007.

[26] L. Krämer, E. Di Napoli, M. Galgon, B. Lang, and P. Bientinesi, "Dissecting the FEAST algorithm for generalized eigenproblems," *Journal of Computational and Applied Mathematics*, vol. 244, no. 1, pp. 1–9, 2013.

[27] L. Krämer, *Integration based solvers for standard and generalized Hermitian eigenvalue problems [Ph.D. thesis]*, Bergische Universität, Wuppertal, Germany, 2014.

[28] G. W. Stewart, "Simultaneous iteration for computing invariant subspaces of non-hermitian matrices," *Numerische Mathematik*, vol. 25, no. 2, pp. 123–136, 1976.

[29] C. Beattie, "Galerkin eigenvector approximations," *Mathematics of Computation*, vol. 69, no. 232, pp. 1409–1434, 2000.

[30] T. Kato, *Perturbation Theory for Linear Operators*, vol. 132 of *Grundlehren der Mathematischen Wissenschaften, Band*, Springer, Berlin, Germany, 2nd edition, 1976.

[31] C. Engström, H. Langer, and C. Tretter, "Non-linear eigenvalue problemsand applications to photonic crystals," http://arxiv.org/abs/1507.06381.