Numer. Math. (2015) 129:353–381
DOI 10.1007/s00211-014-0639-8

Numerische Mathematik

Local convergence of Newton-like methods for degenerate eigenvalues of nonlinear eigenproblems. I. Classical algorithms

Daniel B. Szyld · Fei Xue

Received: 5 October 2012 / Revised: 16 May 2014 / Published online: 19 June 2014

Abstract We study the local convergence rates of several most widely used single-vector Newton-like methods for the solution of a degenerate eigenvalue of nonlinear algebraic eigenvalue problems of the form $T(\lambda)v = 0$. This problem has not been completely understood, since the Jacobian associated with Newton's method is singular at the desired eigenpair, and the standard convergence theory is not applicable. In fact, Newton's method generally converges only linearly towards singular roots. In this paper, we show that the local convergence of inverse iteration, Rayleigh functional iteration and the Jacobi–Davidson method are at least quadratic for semi-simple eigenvalues. For defective eigenvalues, Newton-like methods converge only linearly in general. The results are illustrated by numerical experiments.

Mathematics Subject Classification (2000) 65F15 · 65F10 · 65F50 · 15A18 · 15A22

This work was supported by the U. S. National Science Foundation under grant DMS-1115520.

D. B. Szyld
Department of Mathematics, Temple University, 1805 N. Broad Street,
Philadelphia, PA 19122-6094, USA
e-mail: szyld@temple.edu

F. Xue (✉)
Department of Mathematics, University of Louisiana at Lafayette,
P.O. Box 41010, Lafayette, LA 70504-1010, USA
e-mail: fxue@louisiana.edu

1 Introduction

In this paper, we study the local convergence rates of several classical single-vector Newton-like methods for computing a degenerate eigenvalue and a corresponding eigenvector of nonlinear algebraic eigenvalue problems of the form

$$T(\lambda)v = 0, \tag{1}$$

where $T(\cdot) : U \to \mathbb{C}^{n\times n}$ is holomorphic on a domain $U \subset \mathbb{C}$, $\lambda \in U$ and $v \in \mathbb{C}^n\backslash\{0\}$. λ is an eigenvalue of $T(\cdot)$ if and only if $T(\lambda)$ is singular. Assume that $\det T(\cdot) \not\equiv 0$ in any neighborhood of λ; that is, eigenvalues of $T(\cdot)$ are isolated. The algebraic multiplicity of an eigenvalue λ of (1) is defined as the multiplicity of λ as a root of the characteristic function $\det T(\mu)$. The eigenvalue λ is called *degenerate* if its algebraic multiplicity is greater than one. A degenerate eigenvalue λ is called *semi-simple* if its geometric multiplicity, defined as $\dim(\text{null } T(\lambda))$, equals its algebraic multiplicity; it is called *defective* if its geometric multiplicity is smaller than its algebraic multiplicity.

Our study of numerical methods for degenerate eigenvalues of (1) is motivated by recent rapid development of theory and algorithms for nonlinear eigenvalue problems. These problems arise in a variety of applications; see, e.g., [15,16,30] and references therein. Polynomial and rational eigenvalue problems, which can be transformed by *linearization* to standard linear eigenvalue problems multiple times larger in dimension, have been studied intensively [12,13,26]. For these problems of small or medium size, linearization is usually the most effective approach. For very large or irrational (truly nonlinear) problems where linearization is not effective or applicable, or if there are only a small number of eigenpairs of interest, other methods might be more appropriate. For example, one can use the contour integral method [1,4,8] to find initial approximations to the eigenvalues in a specified domain, and apply Newton-like methods to refine the eigenpair approximations.

Variants of Newton-like methods have been proposed and studied for the solution of a single eigenpair of the nonlinear eigenvalue problem (1); see the early works [19–21], a few recent developments [2,9,22,25], and a study of inexact variants of these algorithms [27]. These analyses are focused on *simple* eigenpairs, for which the local convergence of Newton-like methods are generally quadratic (possibly cubic for problems with local symmetry). On the other hand, convergence of Newton-like methods for degenerate eigenpairs, which arise in important scenarios such as certain gyroscopic systems, hyperbolic Hermitian polynomial problems, and delay differential equations, are not completely understood. The major difficulty is that the Jacobian (Fréchet derivative) of the augmented system of problem (1) at the desired degenerate eigenpair is singular, and the standard convergence theory of Newton's method is not applicable. To work around this difficulty, a block variant of Newton's method was proposed to compute a simple invariant pair including the whole eigenspace of the desired degenerate eigenvalues [11], where in each iteration the number of linear systems to be solved equals the sum of algebraic multiplicities of the desired eigenvalues. This method has a nonsingular Fréchet derivative at the desired invariant pair and exhibits quadratic local convergence, but it is computationally prohibitive for degenerate eigenvalues of high algebraic multiplicities.

For many applications, fortunately, we are mostly interested in eigenvalues alone. In this case, single-vector Newton-like methods are probably the most suitable algorithms. The main purpose of this paper is to gain a better understanding of several widely used single-vector Newton-like methods for the solution of a degenerate eigenpair of problem (1). We investigate the local convergence rates of inverse iteration, Rayleigh functional iteration (RFI), and the single-vector Jacobi-Davidson (JD) method. For semi-simple eigenvalues, we show that the convergence of these algorithms is quadratic or cubic, depending on the symmetry of $T(\lambda)$; in other words, the convergence rates of Newton-like methods for semi-simple eigenpairs are identical to those for simple ones. For defective eigenvalues, however, these methods exhibit only linear convergence in general. In a companion paper [29], we analyze accelerated algorithms that exhibit quadratic convergence in the defective case.

The rest of the paper is organized as follows. In Sect. 2 we review some definitions and preliminary results for degenerate eigenvalues. In Sect. 3, we show that the singularity of the Jacobian of the augmented system at the desired semi-simple eigenpair has no impact on the local quadratic convergence of inverse iteration; in addition, RFI and single-vector JD converge quadratically or cubically towards a semi-simple eigenpair, respectively, for nonsymmetric or symmetric problems. We then show in Sect. 4 that the convergence of Newton-like methods towards a defective eigenpair is generally linear. Numerical experiments are provided throughout the paper to illustrate the convergence results. Section 5 is the conclusion of the paper.

2 Preliminaries

In this section, we review some preliminary results on degenerate eigenvalues for the study of local convergence of Newton-like methods. Theories for semi-simple and defective eigenvalues can be presented in a uniform manner, but we review them separately for the purpose of clarity. For the two types of degenerate eigenvalues, as we shall see, there exist important differences in the structure of the resolvent $T(\mu)^{-1}$ near the eigenvalue λ, and the sets of right and left eigenvectors satisfy different normalization conditions. These properties are fundamental in the understanding of the quadratic (possibly cubic) and the linear convergence of Newton-like methods, respectively, for a semi-simple and a defective eigenpair.

2.1 Semi-simple eigenvalues

Let λ be an eigenvalue of (1), $alg_T(\lambda)$ the algebraic multiplicity of λ, i.e., the multiplicity of λ as a root of the characteristic function $\det T(\mu)$, and $geo_T(\lambda) = \dim(\text{null } T(\lambda))$ the geometric multiplicity. It is shown in [28] that $alg_T(\lambda) \geq geo_T(\lambda)$. The eigenvalue λ is semi-simple if $alg_T(\lambda) = geo_T(\lambda) \geq 2$. Intuitively, a semi-simple eigenpair can be considered as a set of multiple simple eigenpairs sharing an identical eigenvalue. This perspective is helpful for our understanding of the quadratic convergence of Newton-like methods.

The major theorem on the structure of the resolvent $T(\mu)^{-1}$ near a semi-simple eigenvalue and the normalization condition of the sets of left and right eigenvectors

is described as follows, and it can be obtained directly from Theorem A.10.2 of [10], where all Jordan chains are of length 1.

Theorem 1 *Let $T : U \to \mathbb{C}^{n\times n}$ be a Fredholm holomorphic operator in a neighborhood of a semi-simple eigenvalue $\lambda \in U$. Let $alg_T(\lambda) = J$, and $\{\varphi_k\}$ $(k = 1, \ldots, J)$ be the corresponding right eigenvectors. Then there exists a unique set of corresponding left eigenvectors $\{\psi_k\}$ $(k = 1, \ldots, J)$ such that*

$$T(\mu)^{-1} = \sum_{k=1}^{J} \frac{\langle \cdot, \psi_k \rangle \varphi_k}{\mu - \lambda} + Q(\mu)$$

in a neighborhood of λ, where Q is holomorphic in this neighborhood. The two sets of eigenvectors satisfy the following normalization condition

$$\langle T'(\lambda)\varphi_k, \psi_j \rangle = \delta_{jk}, \quad (j, k = 1, \ldots, J). \tag{2}$$

In addition, the right eigenvectors satisfy

$$T(\lambda)Q(\lambda)T'(\lambda)\varphi_k = 0, \quad (k = 1, \ldots, J). \tag{3}$$

Note that λ is a simple eigenvalue if $J = 1$. Therefore, we assume throughout the paper that $J \geq 2$ for the semi-simple case.

Let $x \neq 0$ be an eigenvector approximation which has a significant component in the eigenspace $\text{span}\{\varphi_1, \ldots, \varphi_J\}$ corresponding to a semi-simple eigenvalue λ. We propose a decomposition of x which we use later for the study of the convergence of RFI. Let $G \in \mathbb{C}^{J\times J}$ be a nonsingular matrix, $\Phi_J = [\varphi_1 \ \ldots \ \varphi_J]G$ and $\Psi_J = [\psi_1 \ \ldots \ \psi_J]G^{-*}$ such that

$$\Psi_J^* T'(\lambda)\Phi_J = I_J. \tag{4}$$

From (2), $T'(\lambda)\Phi_J$ and $\Psi_J^* T'(\lambda)$ are of full rank. Let $W_{n-J} \in \mathbb{C}^{n\times(n-J)}$ and $U_{n-J} \in \mathbb{C}^{n\times(n-J)}$, respectively, have orthonormal columns such that

$$W_{n-J}^* T'(\lambda)\Phi_J = 0_{(n-J)\times J} \quad \text{and} \quad \Psi_J^* T'(\lambda)U_{n-J} = 0_{J\times(n-J)}. \tag{5}$$

Assume that $x \notin \text{range}(U_{n-J})$, such that $\|\Psi_J^* T'(\lambda)x\| \neq 0$. A decomposition of x can be formed as follows

$$x = \gamma(cv + sg), \tag{6}$$

where

$$\gamma = \left\| \begin{bmatrix} \Psi_J^* \\ W_{n-J}^* \end{bmatrix} T'(\lambda)x \right\|, \quad c = \frac{\|\Psi_J^* T'(\lambda)x\|}{\gamma}, \quad s = \frac{\|W_{n-J}^* T'(\lambda)x\|}{\gamma},$$
$$v = \Phi_J \frac{\Psi_J^* T'(\lambda)x}{\|\Psi_J^* T'(\lambda)x\|} \in null\ T(\lambda), \quad \text{and} \quad g = \frac{1}{s}\left(\frac{x}{\gamma} - cv\right). \tag{7}$$

Here, γ is a generalized norm of x, and c and s with $c^2 + s^2 = 1$ are generalized sine and cosine, respectively, of the angle between x and $v \in \text{span}\{\varphi_1, \ldots, \varphi_J\}$. It can be shown without difficulty that

$$\Psi_J^* T'(\lambda) g = 0 \quad \text{and} \quad \|W_{n-J}^* T'(\lambda) g\| = 1, \tag{8}$$

and $g \in \text{range}(U_{n-J})$ is an error vector normalized as above. The eigenvector approximation error can be measured by s or the generalized tangent $t = s/c$.

2.2 Defective eigenvalues

For a defective λ, the structure of the resolvent $T(\mu)^{-1}$ near λ and the normalization conditions of the left and the right eigenvectors are more complicated than they are in the semi-simple case.

Definition 2 Let λ be a defective eigenvalue of the holomorphic operator $T(\cdot)$, and $\varphi_{\cdot,0}$ a corresponding right eigenvector. Then the nonzero vectors $\varphi_{\cdot,1}, \ldots, \varphi_{\cdot,m-1}$ are called generalized eigenvectors (principal vectors) if

$$\sum_{j=0}^{\ell} \frac{1}{j!} T^{(j)}(\lambda) \varphi_{\cdot,\ell-j} = 0, \quad \ell = 1, \ldots, m-1, \tag{9}$$

where $T^{(j)}(\lambda) = \frac{d^j}{d\mu^j} T(\mu)|_{\mu=\lambda}$. The ordered collection $\{\varphi_{\cdot,0}, \varphi_{\cdot,1}, \ldots, \varphi_{\cdot,m-1}\}$ is called a right Jordan chain corresponding to λ. If (9) is satisfied for some $m = m_*$ and no more vectors can be introduced such that (9) is satisfied for $m = m_* + 1$, then m_* is called the length of the Jordan chain and a *partial multiplicity* of λ. Similarly, a left Jordan chain $\{\psi_{\cdot,0}, \psi_{\cdot,1}, \ldots, \psi_{\cdot,m-1}\}$ can be defined by replacing $T^{(j)}\varphi_{\cdot,\ell-j}$ in (9) with $\psi_{\cdot,\ell-j}^* T^{(j)}$.

Let $\{\varphi_{k,0}, \ldots, \varphi_{k,m_k-1}\}$ be a right Jordan chain with $m_k \geq 2$, for which $T(\lambda)\varphi_{k,1} + T'(\lambda)\varphi_{k,0} = 0$ holds from Definition 2. Premultiplying by $\psi_{j,0}^*$, we have

$$\psi_{j,0}^* T(\lambda)\varphi_{k,1} + \psi_{j,0}^* T'(\lambda)\varphi_{k,0} = \psi_{j,0}^* T'(\lambda)\varphi_{k,0} = 0 \tag{10}$$

for $j = 1, \ldots, J$. Similarly, $\psi_{j,0}^* T'(\lambda)\varphi_{k,0} = 0$ also holds for all $\psi_{j,0}$ with $m_j \geq 2$ and $k = 1, \ldots, J$. This relation is important in the analysis of the linear convergence of Newton-like methods towards defective eigenvalues.

We end this subsection by reviewing the structure of the resolvent $T(\mu)^{-1}$ near a defective λ. This structure plays a critical role in understanding the local convergence of RFI and single-vector JD. We reiterate the result as follows (cf. Theorem A.10.2 of [10]).

Theorem 3 *Let $T : U \to \mathbb{C}^{n\times n}$ be holomorphic in a neighborhood of a defective eigenvalue λ of T, and J and $m_1, \ldots, m_J$ be the geometric and partial multiplicities of λ. Suppose that $\{\varphi_{k,s}\}$, $k = 1, \ldots, J$, $s = 0, \ldots, m_k - 1$ is a canonical system of right Jordan chains of T corresponding to λ. Then*

(i) There is a unique canonical system of left Jordan chains $\{\psi_{k,s}\}$, $k = 1, \ldots, J$, $s = 0, \ldots, m_k - 1$ *such that*

$$T(\mu)^{-1} = \sum_{k=1}^{J} \sum_{h=0}^{m_k-1} \frac{\sum_{s=0}^{h} \langle \cdot, \psi_{k,s} \rangle \varphi_{k,h-s}}{(\mu - \lambda)^{m_k - h}} + Q(\mu) \tag{11}$$

in a neighborhood of λ, *where* Q *is holomorphic in this neighborhood.*

(ii) The left and the right Jordan chains satisfy the following normalization conditions

$$\sum_{s=0}^{\ell} \sum_{\sigma=s+1}^{m_k+s} \frac{1}{\sigma!} \langle T^{(\sigma)}(\lambda) \varphi_{k,m_k+s-\sigma}, \psi_{j,\ell-s} \rangle = \delta_k^j \delta_\ell^0, \tag{12}$$

$$\sum_{s=0}^{\ell} \sum_{\sigma=s+1}^{m_k+s} \frac{1}{\sigma!} \langle T^{(\sigma)}(\lambda) \varphi_{k,\ell-s}, \psi_{j,m_k+s-\sigma} \rangle = \delta_k^j \delta_\ell^0,$$

where $\psi_{j,p} = 0$, $\varphi_{k,q} = 0$ *for* $p \geq m_j$, $q \geq m_k$ *by convention.*

2.3 Review of Newton-like algorithms

Consider the computation of a single eigenpair (λ, v) of (1) under the normalization condition $u^* v = 1$, where $u \in \mathbb{C}^n \backslash \{0\}$ is a fixed normalization vector such that $u \not\perp v$; see, e.g., [16,21]. This is equivalent to seeking a solution of the augmented system

$$F(\lambda, v) = \begin{bmatrix} T(\lambda) v \\ u^* v - 1 \end{bmatrix} = 0. \tag{13}$$

The augmented system can be solved by Newton's method:

$$\begin{bmatrix} x_{k+1} \\ \mu_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \mu_k \end{bmatrix} - \begin{bmatrix} T(\mu_k) & T'(\mu_k) x_k \\ u^* & 0 \end{bmatrix}^{-1} \begin{bmatrix} x_k \\ \mu_k \end{bmatrix}, \tag{14}$$

which is equivalent to inverse iteration:

$$\begin{cases} p_k = T(\mu_k)^{-1} T'(\mu_k) x_k \\ x_{k+1} = \dfrac{1}{u^* p_k} p_k \\ \mu_{k+1} = \mu_k - \dfrac{1}{u^* p_k} \end{cases}. \tag{15}$$

The convergence analysis of inverse iteration is usually carried out using the theory of Newton's method.

For a degenerate λ with $geo_T(\lambda) = J > 1$, $T(\lambda) = 0$ together with a single normalization vector may not uniquely determine the eigenvector v, since there could be infinitely many ways to form $v \in \text{span}\{\varphi_1, \ldots, \varphi_J\}$ such that $u^* v = 1$. In many

cases, fortunately, one is mostly interested in the eigenvalue alone, and any single eigenvector $v \in \text{span}\{\varphi_1, \ldots, \varphi_J\}$ would satisfy the needs of the applications. We assume this is the case throughout the paper. If the whole eigenspace of a degenerate eigenvalue is needed, one may have to use a block variant of Newton's method based on invariant pairs [11,28].

Inverse iteration can be modified to incorporate alternative eigenvalue approximations. RFI is the most important variant of this type. Given a right eigenvector approximation $x_k \approx v$, one can choose a vector y_k, such that $y_k^* T(\rho_k) x_k = 0$ for some scalar ρ_k, and $\rho_k = \rho_F(x_k; T, y_k)$ is called the value of the nonlinear Rayleigh functional $\rho_F(\cdot; T, y)$; see [23] and references therein. RFI is described as follows

$$\begin{cases} \text{choose } y_k \text{ and compute } \rho_k = \rho_F(x_k; T, y_k) \text{ s.t. } y_k^* T(\rho_k) x_k = 0; \\ p_k = T(\rho_k)^{-1} T'(\rho_k) x_k; \\ \text{normalize } p_k \text{ and assign it to } x_{k+1}. \end{cases} \tag{16}$$

RFI exhibits cubic local convergence for problems with symmetry, where $T(\lambda)$ is real (skew) symmetric, complex (skew) Hermitian, or complex (skew) symmetric, and quadratic convergence in the nonsymmetric case.

RFI bears a close connection to the single-vector JD method, another widely used Newton-like eigenvalue algorithm. For instance a variant of single-vector JD is proposed in [27] to compute a simple eigenpair of (1):

$$\begin{cases} \text{choose a vector } y_k \text{ and compute } \rho_k = \rho_F(x_k; T, y_k) \\ \quad \text{s.t. } y_k^* T(\rho_k) x_k = 0 \text{ and } y_k^* T'(\rho_k) x_k \neq 0; \\ \text{define } \Pi_k^{(1)} = I - \dfrac{T'(\rho_k) x_k y_k^*}{y_k^* T'(\rho_k) x_k}, \Pi_k^{(2)} = I - \dfrac{x_k u^*}{u^* x_k}, \text{ and solve the} \\ \quad \text{correction equation } \Pi_k^{(1)} T(\rho_k) \Pi_k^{(2)} \Delta x_k = -T(\rho_k) x_k \text{ for } \Delta x_k \perp u; \\ \text{set } x_{k+1} = x_k + \Delta x_k, \text{ and normalize when necessary.} \end{cases} \tag{17}$$

It can be shown that the exact solution of the correction equation is

$$\Delta x_k = \frac{T(\rho_k)^{-1} T'(\rho_k) x_k}{u^* T(\rho_k)^{-1} T'(\rho_k) x_k} - x_k,$$

so that $x_k + \Delta x_k = T(\rho_k)^{-1} T'(\rho_x) x_k$ up to a scaling factor. That is, single-vector JD (17) is equivalent to RFI (16), provided that $y_k T'(\rho_k) x_k \neq 0$ and $u^* T(\rho_k)^{-1} T'(\rho_x) x_k \neq 0$. In this paper, we restrict our discussion to the local convergence of single-vector JD, which can be used as a worst-case estimate of the convergence rates of full JD working with a search subspace.

2.4 Convergence of Newton's method near singular roots

To prepare for the convergence analysis of inverse iteration towards degenerate eigenvalues, we need to review the local convergence of Newton's method for a nonlinear

system of algebraic equations $F(z) = 0$ towards a root $z^* \in \mathbb{C}^m$ for which the Jacobian $F'(z^*)$ is singular. A sequence of related results on this topic can be found in [5–7]. Our review is primarily based on [7].

Let $\mathcal{N}_1 = \text{null } F'(z^*)$, $\mathcal{M}_2 = \text{range } F'(z^*)$, so that $\text{codim}(\mathcal{M}_2) = \dim(\mathcal{N}_1)$. We choose complementary subspaces $\mathcal{M}_1, \mathcal{N}_2$ such that $\mathbb{C}^m = \mathcal{N}_1 \oplus \mathcal{M}_1 = \mathcal{N}_2 \oplus \mathcal{M}_2$. Define $P_{\mathcal{N}_i}$ as the projections onto $\mathcal{N}_i$ along $\mathcal{M}_i$, and $P_{\mathcal{M}_i} = I - P_{\mathcal{N}_i}$. Given these subspaces, the Jacobian $F'(z)$ can be decomposed as

$$F'(z) = A_F(z) + B_F(z) + C_F(z) + D_F(z),$$

where

$$\begin{aligned} A_F(z) &= P_{\mathcal{M}_2} F'(z) P_{\mathcal{M}_1}, \quad B_F(z) = P_{\mathcal{M}_2} F'(z) P_{\mathcal{N}_1}, \\ C_F(z) &= P_{\mathcal{N}_2} F'(z) P_{\mathcal{M}_1}, \quad D_F(z) = P_{\mathcal{N}_2} F'(z) P_{\mathcal{N}_1}. \end{aligned} \tag{18}$$

Let $A_{F^*} = P_{\mathcal{M}_2} A_F(z^*) P_{\mathcal{M}_1}$, which is a bijection when regarded as a mapping from $\mathcal{M}_1$ into $\mathcal{M}_2$. Consider the Taylor expansions of the terms in (18) at z^*, i.e., the restrictions of $F'(z)$ onto $\mathcal{M}_i$ and $\mathcal{N}_i$ $(i = 1, 2)$:

$$\begin{aligned} A_F(z) &= A_{F^*} + \sum_{j=a}^{\ell} A_{(j)}(z) + \mathcal{O}_{\ell+1}(e_z), \quad B_F(z) = \sum_{j=b}^{\ell} B_{(j)}(z) + \mathcal{O}_{\ell+1}(e_z), \\ C_F(z) &= \sum_{j=c}^{\ell} C_{(j)}(z) + \mathcal{O}_{\ell+1}(e_z), \quad D_F(z) = \sum_{j=d}^{\ell} D_{(j)}(z) + \mathcal{O}_{\ell+1}(e_z), \end{aligned} \tag{19}$$

where

$$\begin{aligned} A_{(j)}(z) &= \frac{1}{j!} P_{\mathcal{M}_2} F^{(j+1)}(z^*)(e_z^j, P_{\mathcal{M}_1}\cdot), \ B_{(j)}(z) = \frac{1}{j!} P_{\mathcal{M}_2} F^{(j+1)}(z^*)(e_z^j, P_{\mathcal{N}_1}\cdot), \\ C_{(j)}(z) &= \frac{1}{j!} P_{\mathcal{N}_2} F^{(j+1)}(z^*)(e_z^j, P_{\mathcal{M}_1}\cdot), \ D_{(j)}(z) = \frac{1}{j!} P_{\mathcal{N}_2} F^{(j+1)}(z^*)(e_z^j, P_{\mathcal{N}_1}\cdot) \end{aligned}$$

are square matrices. Here, $F^{(j+1)}(z^*)(\cdot, \cdots, \cdot)$ stands for the $(j+1)$st derivative of F at z^* (a multilinear form, or tensor, with $j+1$ arguments), and e_z^j means that all the first j arguments of $F^{(j+1)}$ are e_z. The values $j = a, b, c, d > 0$ in (19) are the smallest integers, independent of z, for which the *vectors* $A_{(j)}(z)e_z$, $B_{(j)}(z)e_z$, $C_{(j)}(z)e_z$, and $D_{(j)}(z)e_z$, respectively, are nonzero. In addition, let $j = \bar{a}, \bar{b}, \bar{c}, \bar{d}$ be the smallest integers independent of z for which the *matrices* $A_{(j)}(z)$, $B_{(j)}(z)$, $C_{(j)}(z)$ and $D_{(j)}(z)$ are nonzero. By definition, $\bar{a} \le a$, $\bar{b} \le b$, $\bar{c} \le c$ and $\bar{d} \le d$.

With the above notation and definitions, a major convergence result of Newton's method near singular points [7] is summarized as follows.

Theorem 4 (Theorem 5.9 in [7]) *Let z^* be a singular root of $F(z) = 0$, and $e_z = z - z^*$. Define the operator $\widetilde{D}_{(j)} = \frac{1}{j!} P_{\mathcal{N}_2} F^{(j+1)}(z^*)((P_{\mathcal{N}_1} e_z)^j, P_{\mathcal{N}_1}\cdot)$, and let $j = \widetilde{d}$ be the smallest integer, independent of z, for which $\widetilde{D}_{(j)} e_z \neq 0$. Assume that $\widetilde{d} = d \le c$,*

$\bar{d} \le \bar{c}$ and $\widetilde{D}_{(\widetilde{d})}$ is nonsingular for all $P_{\mathcal{N}_1} e_z \neq 0$. Let z_0 be the initial Newton iterate, and define $\eta = \min(a, b, c, d)$. Then, for sufficiently small $\delta > 0$ and $\theta > 0$, $F'(z_0)$ is nonsingular for all

$$z_0 \in W(\delta, \theta) \equiv \left\{ z : 0 < \|e_z\| \le \delta,\ \|P_{\mathcal{M}_1} e_z\| \le \theta \|P_{\mathcal{N}_1} e_z\| \right\},$$

all subsequent Newton iterates remain in $W(\delta, \theta)$; in addition, $z_k \to z^*$ with

$$\|P_{\mathcal{M}_1}(z_k - z^*)\| \le C \|P_{\mathcal{M}_1}(z_{k-1} - z^*)\|^{\eta+1}$$

for some constant $C > 0$ independent of k, and

$$\lim_{k\to\infty} \frac{\|P_{\mathcal{N}_1}(z_k - z^*)\|}{\|P_{\mathcal{N}_1}(z_{k-1} - z^*)\|} = \frac{d}{d+1}.$$

Theorem 4 states that under certain assumptions, as Newton's method converges towards z^* for which $F'(z^*)$ is singular, the error component lying in $\mathcal{M}_1$ converges at least quadratically, whereas the error component lying in $\mathcal{N}_1$ converges only linearly. This observation will be used in Sects. 3 and 4 to show the quadratic and the linear convergence of inverse iteration, respectively, towards a semi-simple and a defective eigenpair.

3 Convergence for semi-simple eigenvalues

In this section, we study the local convergence of Newton-like methods for computing a semi-simple eigenvalue of the nonlinear eigenproblem (1). As we shall see, the singularity of the Jacobian of the augmented system (13) at a semi-simple eigenpair does not hamper the quadratic convergence of Newton's method; in addition, RFI also exhibits the same order of convergence for semi-simple eigenvalues as it does for simple ones.

3.1 Inverse iteration

Assume that (λ, v) is a semi-simple eigenpair of the holomorphic operator $T(\cdot)$, and $\varphi_1, \ldots, \varphi_J$ and $\psi_1, \ldots, \psi_J$ are the corresponding left and right eigenvectors. Since $\dim(\operatorname{null} T(\lambda)) = J$, there exists a singular value decomposition of $T(\lambda)$ of the form $Y^* T(\lambda) X = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma_{n-J} \end{bmatrix}$, where $X = [X_J \ X_{n-J}]$ and $Y = [Y_J \ Y_{n-J}]$ are unitary matrices, $X_J \in \mathbb{C}^{n\times J}$ and $Y_J \in \mathbb{C}^{n\times J}$ have orthonormal columns forming a basis of $\operatorname{span}\{\varphi_1, \ldots, \varphi_J\}$ and $\operatorname{span}\{\psi_1, \ldots, \psi_J\}$, respectively, and Σ_{n-J} is a diagonal matrix of positive singular values of $T(\lambda)$. Therefore there exist nonsingular matrices $K_J, M_J \in \mathbb{C}^{J\times J}$ such that $X_J = [\varphi_1 \ \ldots \ \varphi_J] K_J$ and $Y_J = [\psi_1 \ \ldots \ \psi_J] M_J$. Since $v \in \operatorname{span}\{\varphi_1, \ldots, \varphi_J\}$, we have $v = [\varphi_1 \ \ldots \ \varphi_J] d_v$ for some $d_v \in \mathbb{C}^J \setminus \{0\}$. It follows from (2) that $Y_J^* T'(\lambda) v = M_J^* [\psi_1 \ \ldots \ \psi_J]^* T'(\lambda) [\varphi_1 \ \ldots \ \varphi_J] d_v = M_J^* d_v$, and

$$\begin{bmatrix} Y^* & \\ & 1 \end{bmatrix} \begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix} \begin{bmatrix} X & \\ & 1 \end{bmatrix} = \begin{bmatrix} Y^*T(\lambda)X & Y^*T'(\lambda)v \\ u^*X & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & M_J^* d_v \\ 0 & \Sigma_{n-J} & Y_{n-J}^* T'(\lambda)v \\ u^*X_J & u^*X_{n-J} & 0 \end{bmatrix}. \quad (20)$$

Let $h = [h_a^T \ h_b^T \ h_c]^T$ be a vector in the null space of the above square matrix, where $h_a \in \mathbb{C}^J$, $h_b \in \mathbb{C}^{n-J}$ and $h_c \in \mathbb{C}$. Then

$$\begin{bmatrix} 0 & 0 & M_J^* d_v \\ 0 & \Sigma_{n-J} & Y_{n-J}^* T'(\lambda)v \\ u^*X_J & u^*X_{n-J} & 0 \end{bmatrix} \begin{bmatrix} h_a \\ h_b \\ h_c \end{bmatrix} = \begin{bmatrix} M_J^* d_v h_c \\ \Sigma_{n-J} h_b + Y_{n-J}^* T'(\lambda)v h_c \\ u^*X_J h_a + u^*X_{n-J} h_b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

In the first J rows, since M_J is nonsingular and $d_v \neq 0$, we have $M_J^* d_v \neq 0$, and therefore $h_c = 0$. The next $n - J$ rows is thus simplified as $\Sigma_{n-J} h_b$, and since Σ_{n-J} is diagonal with all nonzero entries, we have $h_b = 0$, and the last row is equivalent to $u^*X_J h_a = 0$. Now, since u with $u^*v = 1$ specifies the scaling of $v = [\varphi_1 \ \dots \ \varphi_J]d_v$ in the desired eigenspace range(X_J), we have $u^*X_J \neq 0$. Without loss of generality, assume that $u^*X_J = [\eta_1 \ \dots \ \eta_J]$ where $\eta_J \neq 0$. Then we can determine h_a and h.

$$h = \begin{bmatrix} h_a \\ h_b \\ h_c \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{\eta_1}{\eta_J} \\ 0_{n-J+1} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -\frac{\eta_2}{\eta_J} \\ 0_{n-J+1} \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{\eta_{J-1}}{\eta_J} \\ 0_{n-J+1} \end{bmatrix} \right\}.$$

From (20), we can premultiply h by $\begin{bmatrix} X & \\ & 1 \end{bmatrix}$ to obtain null $\left(\begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix} \right)$.

$$\mathcal{N}_1 \equiv \text{null} \left(\begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix} \right) = \begin{bmatrix} X & \\ & 1 \end{bmatrix} \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{\eta_1}{\eta_J} \\ 0_{n-J+1} \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{\eta_{J-1}}{\eta_J} \\ 0_{n-J+1} \end{bmatrix} \right\}$$
$$= \text{span} \left\{ \begin{bmatrix} X\mathbf{e_1} - \frac{\eta_1}{\eta_J} X\mathbf{e_J} \\ 0 \end{bmatrix}, \begin{bmatrix} X\mathbf{e_2} - \frac{\eta_2}{\eta_J} X\mathbf{e_J} \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} X\mathbf{e_{J-1}} - \frac{\eta_{J-1}}{\eta_J} X\mathbf{e_J} \\ 0 \end{bmatrix} \right\}$$
$$\subset \text{span} \left\{ \begin{bmatrix} \varphi_1 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} \varphi_J \\ 0 \end{bmatrix} \right\}, \quad (21)$$

where $\dim(\mathcal{N}_1) = J - 1$ ($J \geq 2$), and $\mathbf{e_i}$ is the standard basis vector with a one in the i-th entry and zeros elsewhere.

The special structure of the null space $\mathcal{N}_1$ indicates that the convergence of Newton's method towards the semi-simple eigenvalue and the corresponding *eigenspace* (instead of the eigenvector v) is quadratic. To see this, define

$$\mathcal{M}_1 = \mathcal{M}_{1\alpha} \oplus \mathcal{M}_{1\beta} = \operatorname{span}\left\{\begin{bmatrix} X\mathbf{e_J} \\ 0 \end{bmatrix}\right\} \oplus \operatorname{range}\left(\begin{bmatrix} X_{n-J} & \\ & 1 \end{bmatrix}\right),$$

so that $\dim(\mathcal{M}_1) = n - J + 2$. Therefore

$$\begin{aligned}\mathbb{C}^{n+1} &= \operatorname{span}\left\{\begin{bmatrix} X\mathbf{e_1} - \frac{\eta_1}{\eta_J} X\mathbf{e_J} \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} X\mathbf{e_{J-1}} - \frac{\eta_{J-1}}{\eta_J} X\mathbf{e_J} \\ 0 \end{bmatrix}\right\} \\ &\quad \oplus \left(\operatorname{span}\left\{\begin{bmatrix} X\mathbf{e_J} \\ 0 \end{bmatrix}\right\} \oplus \operatorname{range}\left(\begin{bmatrix} X_{n-J} & \\ & 1 \end{bmatrix}\right)\right) \\ &= \mathcal{N}_1 \oplus \left(\mathcal{M}_{1\alpha} \oplus \mathcal{M}_{1\beta}\right) = \mathcal{N}_1 \oplus \mathcal{M}_1.\end{aligned}$$

Let $\mathcal{P}_{\mathcal{N}_1}$ be the projector onto $\mathcal{N}_1$ along $\mathcal{M}_1$, $\mathcal{P}_{\mathcal{M}_1} = I - \mathcal{P}_{\mathcal{N}_1}$, and $e_k = [x_k; \mu_k] - [v; \lambda]$ the error between the Newton iterate and the *particular eigenpair* (λ, v) in the kth iteration. We know from Theorem 4 that $\|\mathcal{P}_{\mathcal{N}_1}(e_k)\|$ and $\|\mathcal{P}_{\mathcal{M}_1}(e_k)\|$ converge to zero linearly and quadratically, respectively. Thus Newton's method converges towards the *particular eigenpair* (λ, v) linearly.

Fortunately, the linear convergence of Newton's method towards (λ, v) does not affect its quadratic convergence towards λ and the desired *eigenspace*. The key observation is that the error between the Newton iterate and the eigenvalue together with its eigenspace lies in $\mathcal{M}_1$, and $\|\mathcal{P}_{\mathcal{M}_1}(e_k)\|$ converges quadratically. In fact, the eigenspace approximation error in $\operatorname{range}\left(\begin{bmatrix} X_{n-J} \\ 0 \end{bmatrix}\right)$ together with the eigenvalue approximation error in $\operatorname{span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ can be represented by $\mathcal{P}_{\mathcal{M}_{1\beta}}(e_k)$, the projection of e_k onto $\mathcal{M}_{1\beta} = \operatorname{range}\left(\begin{bmatrix} X_{n-J} & \\ & 1 \end{bmatrix}\right)$ along $\mathcal{N}_1 \oplus \mathcal{M}_{1\alpha} = \operatorname{span}\left\{\begin{bmatrix} \varphi_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} \varphi_J \\ 0 \end{bmatrix}\right\}$. Therefore, $\mathcal{P}_{\mathcal{M}_{1\beta}}(e_k)$, instead of $\mathcal{P}_{\mathcal{N}_1}(e_k)$, represents the error between the Newton iterate and the space $\mathcal{E} = \operatorname{span}\left\{\begin{bmatrix} \varphi_1 \\ \lambda \end{bmatrix}, \ldots, \begin{bmatrix} \varphi_J \\ \lambda \end{bmatrix}\right\}$ spanned by all valid candidate eigenpairs. In addition, (21) shows that any vector lying in $\mathcal{N}_1$ can be represented as the difference between two candidate eigenpairs, and thus $\mathcal{P}_{\mathcal{N}_1}(e_k)$ bears no connection to the error between the Newton iterate $z_k = [x_k; \mu_k]$ and $\mathcal{E}$ spanned by all candidate eigenpairs. Since $\mathcal{M}_{1\beta} \subset \mathcal{M}_1$, the quadratic convergence of Newton's method towards the semi-simple λ and its eigenspace (represented by $\mathcal{E}$) is established from the quadratic convergence of $\|\mathcal{P}_{\mathcal{M}_1}(e_k)\|$.

Theorem 5 *Let λ be a semi-simple eigenvalue of the holomorphic operator $T(\cdot) : U \to \mathbb{C}^{n\times n}$, and* $\operatorname{null} T(\lambda) = \operatorname{span}\{\varphi_1, \ldots, \varphi_J\}$. *Let (μ_0, x_0) be a right eigenpair approximation such that $|\mu_0 - \lambda|$ and $\angle(x_0, \operatorname{null} T(\lambda))$ are sufficiently small. Assume that the conditions in* Theorem 4 *hold. Then inverse iteration* (15) *converges towards λ together with its eigenspace quadratically.*

Table 1 Description of the test problems for semi-simple eigenvalues

Problem	Source	Type	Size	Eigenvalue	Multiplicity	Symmetry
*tols*1090	MM	lep	1,090	−12.098	200	real unsymm.
plasma_drift	NLEVP	pep	128	10.004129 − 0.19324032*i*	2	cplx unsymm.
schrödinger	NLEVP	qep	1998	−0.33111181 + 0.24495497*i*	2	cplx symm.
ss_art_symm	Artificial	nep	256	0	5	real symm.
ss_art_unsymm	Artificial	nep	256	0	5	real unsymm.

3.2 Numerical experiments for inverse iteration

In this section, we provide numerical results illustrating the local quadratic convergence of inverse iteration for semi-simple eigenvalues. The experiments are performed on five benchmark problems, one from the MatrixMarket [14], two from the NLEVP toolbox [3], and two constructed artificially. A description of these problems is given in Table 1.

For example, the problem *tols*1090 taken from the MatrixMarket (MM) is a linear eigenvalue problem (lep), defined by a matrix $A \in \mathbb{R}^{1090\times 1090}$; it has a semi-simple eigenvalue $\lambda = -12.098$ of multiplicity 200, and the matrix $T(\lambda) = \lambda I - A$ is real unsymmetric (see the column of "symmetry"). Similarly, *plasma_drift* and *schrödinger* taken from NLEVP, respectively, are a polynomial eigenvalue problem (pep) of degree 3 and a quadratic eigenvalue problem (qep); they both have a semi-simple eigenvalue λ of multiplicity 2, and $T(\lambda)$ is complex unsymmetric and complex symmetric, respectively. The two artificial problems are both truly nonlinear eigenvalue problems (nep), where $T(\mu)$ cannot be represented as a polynomial of μ; specifically,

$$\begin{aligned} T_{symm}(\mu) &= G_A^* D(\mu) G_A \quad \text{for} \quad \text{“}ss\_art\_symm\text{”}, \quad \text{and} \\ T_{unsymm}(\mu) &= G_A^* D(\mu) G_B \quad \text{for} \quad \text{“}ss\_art\_unsymm\text{”}, \end{aligned} \tag{22}$$

where $D(\mu) = \operatorname{diag}([e^{\mu} - 1;\ \ 2\sin\mu;\ \ -5\ln(1+\mu);\ \ 8\mu;\ \ \tan^{-1}(\mu);\ \ c_0 + c_1\mu + c_2\mu^2])$, $c_i \in \mathbb{R}^{251}\,(i = 0, 1, 2)$ and $G_A, G_B \in \mathbb{R}^{256\times 256}$ generated by MATLAB's function `randn` are random matrices whose entries follow the standard normal distribution. Both artificial problems have a semi-simple eigenvalue $\lambda = 0$ of multiplicity 5.

To test the local convergence of inverse iteration for these problems, we first construct an initial eigenvector approximation using the MATLAB commands

$$\texttt{x\_0} = \texttt{v} * \texttt{cos(err\_theta)} + \texttt{g} * \texttt{sin(err\_theta)}; \tag{23}$$

where v is a normalized eigenvector corresponding to the desired eigenvalue of the given problem, g is a normalized random vector that has been orthogonalized against v, and $\texttt{err\_theta} = \angle(x_0, v)$. An appropriate value for *err_theta* is obtained by

trial and error, such that quadratic convergence of inverse iteration can be clearly observed. The initial eigenvalue approximation μ_0 is defined as the value of the one-sided Rayleigh functional $\rho_F(x_0; T, y)$ where the auxiliary vector $y = T'(\lambda)v$, and therefore $|\mu_0 - \lambda| = \mathcal{O}(\angle(x_0, v))$.

To estimate the convergence rate of inverse iteration, we monitor how quickly the eigenresidual norm $\|r_k\| = \|T(\mu_k)x_k\|$ decreases as the algorithm proceeds. The residual norm is a natural measure of the eigenpair approximation error $\|e_k\| = \|[x_k; \mu_k] - [v; \lambda]\|$, and it can be evaluated at little cost. In fact, assuming that $x_k = v\cos\angle(x_k, v) + g\sin\angle(x_k, v)$ as in (23), we have

$$\begin{aligned} T(\mu_k)x_k &= T(\lambda)x + (\mu_k - \lambda)T'(\lambda)x + \mathcal{O}\left((\mu_k - \lambda)^2\right) \\ &= s_k T(\lambda)g + (\mu_k - \lambda)T'(\lambda)v + \mathcal{O}(s_k^2), \end{aligned}$$

and thus $\|T(\mu_k)x_k\| = \mathcal{O}(s_k)$ unless in very special circumstances where the dominant components of $s_k T(\lambda)g$ and $(\mu_k - \lambda)T'(\lambda)v$ happen to cancel out each other such that $\|T(\mu_k)x_k\| = \mathcal{O}(s_k^2)$; see [29, Lemma 1] for such a counterexample. In our experiment, this special case does not happen, and thus we use $\|r_k\| = \mathcal{O}(s_k) = \mathcal{O}(\|e_k\|)$ to explore the convergence rates experimentally.

For all the test problems, we find that inverse iteration converges at least superlinearly. Nevertheless, our goal in this section is to show that the order of convergence η of this algorithm is exactly 2, and the standard criterion $\|r_{k+1}\|/\|r_k\|^\eta \le C$ may not be very descriptive for this purpose. For example, if $\|r_0\| = 10^{-2}$, $\|r_1\| = 10^{-5}$ and $\|r_2\| = 10^{-12}$, it is then difficult to tell if the convergence is quadratic, superquadratic or cubic, due to the very small number of iterations for which the standard criterion holds.

We now discuss an alternative approach to estimate the order of convergence η in a more descriptive manner. First, we generate a *sequence* of initial approximations $(\mu_0^{(j)}, x_0^{(j)})$, such that $\angle(v, x_0^{(j+1)}) = \frac{1}{2}\angle(v, x_0^{(j)})$, and thus $|\mu_0^{(j+1)} - \lambda| \approx \frac{1}{2}|\mu_0^{(j)} - \lambda|$ since the value of the Rayleigh functional $\mu_0^{(j)} = \rho_F(x_0^{(j)}; T, y)$ satisfies $|\mu_0^{(j)} - \lambda| = \mathcal{O}\big(\angle(x_0^{(j)}, v)\big)$. It can be shown that $\|r_0^{(j+1)}\| = \|T(\mu_0^{(j+1)})x_0^{(j+1)}\| \approx \frac{1}{2}\|r_0^{(j)}\| = \frac{1}{2}\|T(\mu_0^{(j)})x_0^{(j)}\|$. We then apply *one step* of inverse iteration to generate a sequence of new iterates $(\mu_1^{(j)}, x_1^{(j)})$, for which $\|r_1^{(j+1)}\| = \left(\frac{1}{2}\right)^\eta \|r_1^{(j)}\|$ holds, and an estimate of η can be obtained. This approach is more descriptive because a relatively large number of initial approximations can be generated, for which the algorithm exhibits the ηth order of convergence for at least one iteration.

The estimated order of convergence of inverse iteration is summarized in Table 2. To explain the results, take for instance the problem *tols*1090. We generated 20 initial approximations $(\mu_0^{(j)}, x_0^{(j)})$ $(j = 1, 2, \ldots, 20)$, where $\angle(x_0^{(1)}, v) = 10^{-2}$ and $\angle(x_0^{(j+1)}, v) = \frac{1}{2}\angle(x_0^{(j)}, v)$. The estimates of the order of convergence ℓ are obtained by applying the least-squares line formula to $\left(\log\|r_0^{(j)}\|, \log\|r_1^{(j)}\|\right)$. As we see, the estimated values of η are very close to 2, indicating that inverse iteration converges quadratically, independent of the symmetry of $T(\lambda)$; see Theorem 5.

Table 2 Estimated order of convergence η for inverse iteration

Problem	$\angle(x_0^{(1)}, v)$	# Init. approx.	Estimated η
*tols*1090	10^{-2}	20	2.022
plasma_drift	2×10^{-1}	20	2.002
schrödinger	10^{-4}	13	1.979
ss_art_symm	2×10^{-2}	16	2.008
ss_art_unsymm	2×10^{-2}	16	1.983

3.3 Rayleigh functional iteration and single-vector JD

We see that a semi-simple eigenvalue is a class of degenerate eigenvalue for which Newton's method exhibit the same order of convergence as for simple eigenvalues. In fact, we can take a step further and show that RFI converges cubically towards a semi-simple λ, if $T(\lambda)$ is (skew) real or complex symmetric, or (skew) Hermitian. In this case, note that the right and the left eigenvectors satisfy $\varphi_j = \psi_j$ or $\varphi_j = \overline{\psi}_j$ ($j = 1, \ldots, J$), so that a left eigenvector approximation y can be obtained directly from the right eigenvector approximation x. To realize the cubic convergence, one should take advantage of the symmetry of $T(\lambda)$ when generating the Rayleigh functional value $\rho = \rho_F(x; T, y)$.

We now present the local convergence analysis of RFI. First, note that for any $v = \sum_{j=1}^{J} \alpha_j \varphi_j$ such that $T(\lambda)v = 0$, we have

$$\sum_{k=1}^{J} \left(\psi_k^* T'(\lambda) v\right) \varphi_k = \sum_{k=1}^{J} \left(\sum_{j=1}^{J} \alpha_j \left(\psi_k^* T'(\lambda) \varphi_j\right) \right) \varphi_k = \sum_{k=1}^{J} \alpha_k \varphi_k = v. \tag{24}$$

Note from (3) that for such v, we have $Q(\lambda)T'(\lambda)v \in \operatorname{span}\{\varphi_1, \ldots, \varphi_J\}$ where both $Q(\cdot)$ and $T'(\cdot)$ are holomorphic in a neighborhood of λ; that is, there exists some constant α such that

$$Q(\lambda)T'(\lambda)v = \alpha \widetilde{v} \text{ with } \widetilde{v} \in \operatorname{span}\{\varphi_1, \ldots, \varphi_J\}. \tag{25}$$

In the following derivation, let $\rho = \rho_F(x; T, y)$ be the Rayleigh functional value corresponding to x and an auxiliary vector y. To prepare for the analysis, we also need decompositions of the following vectors

$$Q(\lambda)T'(\lambda)g = \gamma_1(c_1 v_1 + s_1 g_1), \tag{26}$$
$$Q'(\lambda)T'(\rho)x = \gamma_2(c_2 v_2 + s_2 g_2), \quad \text{and} \tag{27}$$
$$Q(\rho)T''(\lambda)x = \gamma_3(c_3 v_3 + s_3 g_3), \tag{28}$$

which can be obtained by substituting the left-hand sides in (26)–(28), respectively, for x in (6). Therefore $v_j \in \operatorname{span}\{\varphi_1, \ldots, \varphi_J\}$ and

$$T'(\lambda)g_j \perp \operatorname{span}\{\psi_1, \ldots, \psi_J\} \tag{29}$$

for $j = 1, 2, 3$; see (6)–(8). Without loss of generality, assume that the right eigenvector approximation x has a unit generalized norm, i.e., $\gamma = 1$; see (6) and (7). Assuming that ρ is in a neighborhood of λ where $Q(\cdot)$ is analytic, the generalized norms γ_j in (26)–(28) are bounded above by $\mathcal{O}(1)$. It follows that the new unnormalized eigenvector approximation p computed by RFI is

$$
\begin{aligned}
p = T(\rho)^{-1}T'(\rho)x &= \frac{1}{\rho-\lambda}\sum_{k=1}^{J}\left(\psi_k^* T'(\rho)x\right)\varphi_k + Q(\rho)T'(\rho)x \\
&= \frac{1}{\rho-\lambda}\sum_{k=1}^{J}\psi_k^*\left(T'(\lambda)(cv+sg)+(\rho-\lambda)T''(\lambda)x+\frac{(\rho-\lambda)^2}{2}T'''(\lambda)x\right)\varphi_k \\
&\quad + Q(\lambda)T'(\lambda)(cv+sg)+(\rho-\lambda)\left(Q'(\lambda)T'(\rho)+Q(\rho)T''(\lambda)\right)x+\mathcal{O}\left(|\rho-\lambda|^2\right) \\
&= \frac{cv}{\rho-\lambda}+\sum_{k=1}^{J}\left(\psi_k^* T''(\lambda)x+\frac{\rho-\lambda}{2}\psi_k^* T'''(\lambda)x\right)\varphi_k + c\alpha\widetilde{v}+sQ(\lambda)T'(\lambda)g \\
&\quad + (\rho-\lambda)\left(Q'(\lambda)T'(\rho)+Q(\rho)T''(\lambda)\right)x+\mathcal{O}\left(|\rho-\lambda|^2\right) \text{ (see (8), (24) and (25))} \\
&= \frac{cv}{\rho-\lambda}+\sum_{k=1}^{J}\eta_k\varphi_k + c\alpha\widetilde{v}+s\gamma_1 c_1 v_1+(\rho-\lambda)(\gamma_2 c_2 v_2+\gamma_3 c_3 v_3) \\
&\quad + s\gamma_1 s_1 g_1+(\rho-\lambda)(\gamma_2 s_2 g_2+\gamma_3 s_2 g_3)+\mathcal{O}\left(|\rho-\lambda|^2\right) \text{ (see (26)–(28))} \\
&= \widehat{v}_p+\widehat{g}_p+\mathcal{O}\left(|\rho-\lambda|^2\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\widehat{v}_p &= \frac{cv}{\rho-\lambda}+\sum_{k=1}^{J}\eta_k\varphi_k + c\alpha\widetilde{v}+s\gamma_1 c_1 v_1+(\rho-\lambda)(\gamma_2 c_2 v_2+\gamma_3 c_3 v_3), \\
\eta_k &= \psi_k^* T''(\lambda)x+\frac{\rho-\lambda}{2}\psi_k^* T'''(\lambda)x, \quad \text{and} \\
\widehat{g}_p &= s\gamma_1 s_1 g_1+(\rho-\lambda)(\gamma_2 s_2 g_2+\gamma_3 s_2 g_3),
\end{aligned}
$$

so that $\widehat{v}_p \in \operatorname{span}\{\varphi_1, \ldots, \varphi_J\}$ and $T'(\lambda)\widehat{g}_p \perp \operatorname{span}\{\psi_1, \ldots, \psi_J\}$.

The convergence rates of RFI can be obtained by studying the error angle $\angle(p, \widehat{v}_p)$. Recall that $\rho = \rho_F(x; T, y)$ is the Rayleigh functional value such that $y^*T(\rho)x = 0$. Suppose that x is a good right eigenvector approximation for which the generalized sine s is sufficiently small, then $|\rho-\lambda| = \mathcal{O}(s^2)$ for $T(\lambda)$ with symmetry, and $|\rho-\lambda| = \mathcal{O}(s)$ otherwise; see Theorem 5 in [23]. In both cases, it is easy to see that $(\rho-\lambda)^{-1}cv$ is the unique dominant term in $\widehat{v}_p$, and $\|\widehat{v}_p\| = \mathcal{O}(|\rho-\lambda|^{-1})$; in addition, $\|\widehat{g}_p\| = \mathcal{O}(s)+\mathcal{O}(|\rho-\lambda|) = \mathcal{O}(s)$. Following the discussion of the error angle $\angle(x, v)$ (see (6) and (7)), we see that the generalized tangent of the error angle $\angle(p, \widehat{v}_p)$ is bounded above by a quantity proportional to the ratio between the magnitude of the error component $\widehat{g}_p$ and that of the eigenvector component $\widehat{v}_p$. That is,

$$\operatorname{gtan}\angle(p,\widehat{v}_p) \le \mathcal{O}\left(\frac{\|\widehat{g}_p\| + \mathcal{O}(|\rho-\lambda|^2)}{\|\widehat{v}_p\| - \mathcal{O}(|\rho-\lambda|^2)}\right) = \mathcal{O}(|\rho-\lambda|s)$$
$$= \begin{cases} \mathcal{O}(s^3) & \text{for } T(\lambda) \text{ with symmetry} \\ \mathcal{O}(s^2) & \text{otherwise} \end{cases}.$$

In conclusion, the convergence rates of RFI for semi-simple and simple eigenvalues are identical. This observation also holds for the single-vector JD method (17), because it is mathematically equivalent to RFI. The result is summarized in the following theorem.

Theorem 6 *Let λ be a semi-simple eigenvalue of the holomorphic operator $T(\cdot) : U \to \mathbb{C}^{n\times n}$, and $x_0 = \gamma(c_0 v + s_0 g)$ be a corresponding initial right eigenvector approximation with a sufficiently small error angle $\angle(x_0, v)$. Then RFI (16) and single-vector JD (17) converge towards λ and its eigenspace at least quadratically. The local convergence is at least cubic for $T(\lambda)$ with symmetry, assuming that the two-sided Rayleigh functional is properly used; that is, $\rho = \rho_F(x; T, x)$ if $T(\lambda)$ is real (skew) symmetric or complex (skew) Hermitian, and $\rho = \rho_F(x; T, \bar{x})$ if $T(\lambda)$ is complex (skew) symmetric, such that $|\rho - \lambda| = \mathcal{O}(s^2)$.*

3.4 Numerical experiments for RFI and JD

We test the convergence of RFI and single-vector JD on the five problems with semi-simple eigenvalues introduced in Sect. 3.2. We follow the same approach used for inverse iteration to estimate the order of convergence for RFI, with the only difference in the generation of the initial eigenvalue approximations $\mu_0^{(j)}$. To achieve the maximum order of convergence for RFI, as discussed in Sect. 3.3, we use the two-sided Rayleigh functional whenever the symmetry of $T(\lambda)$ exists; that is, for *schrödinger* and *ss_art_symm* (see Table 1), we choose $y = \operatorname{conj}(x_0^{(j)})$ and $y = x_0^{(j)}$, respectively, and let $\mu_0^{(j)} = \rho_F(x_0^{(j)}; T, y)$, such that $|\mu_0^{(j)} - \lambda| = \mathcal{O}\big(\angle(x_0^{(j)}, v)^2\big)$. As a result, RFI converges at least cubically for these problems.

Table 3 presents the estimated order of convergence η for RFI. The results also hold for single-vector JD because the two algorithms are mathematically equivalent. We see that RFI converges quadratically for the three problems without symmetry of $T(\lambda)$, and it converges cubically for the two problems with symmetry (see the numbers in bold in Table 3). All these results are consistent with Theorem 6.

3.5 Semi-simple eigenvalues with perturbation: a practical example

In practical applications, semi-simple eigenvalues typically arise due to symmetries of the continuous operator and of the domain. A classical example is the Laplacian in a three-dimensional cube. A finite element discretization by a standard mesh generator usually neglects the symmetry of the domain and thus yields a discrete Laplacian that does not fully reflect the symmetry of the problem. Consequently, the discrete Laplacian has groups of tightly clustered simple eigenvalues that result from small

Table 3 Estimated order of convergence η for RFI/JD

Problem	$\angle(x_0^{(1)}, v)$	# Init. approx.	Estimated η
*tols*1090	10^{-6}	16	2.013
plasma_drift	5×10^{-4}	11	2.012
schrödinger	5×10^{-4}	10	**2.972**
ss_art_symm	2×10^{-2}	12	**3.006**
ss_art_unsymm	2×10^{-2}	17	1.997

perturbations to semi-simple eigenvalues. Here, we provide some intuitive insight into the behavior of single-vector Newton-like methods for solving one simple eigenvalue that belongs to such a cluster, and we illustrate our analysis with a numerical example.

For the purpose of clarity, we assume that the *relative* distance among the cluster of eigenvalues is moderately small, say, about 10^{-6}, and we want to solve one eigenvalue in this cluster to a high accuracy close to machine precision. We also assume that the initial eigenvector approximation x_0 is such that $\angle(x_0, \mathcal{V})$ is small, e.g., $\angle(x_0, \mathcal{V}) = 10^{-2}$, where $\mathcal{V}$ stands for the whole eigenspace associated with the cluster of eigenvalues. Note that this assumption does *not* mean that x_0 is necessarily close to any particular eigenvector in this eigenspace. Such x_0 could be obtained, for example, from the inverse power method with a fixed shift near this cluster of eigenvalues.

Consider using Rayleigh quotient iteration (RQI) with such an initial iterate x_0 to compute an eigenvalue of a Hermitian matrix A that is tightly clustered with a few other eigenvalues. The relative distance between the initial Rayleigh quotient $\rho_0 = (x_0^T A x_0)/(x_0^T x_0)$ and the eigenvalue cluster, by assumption, is mildly small, say, about 10^{-2}. Since this quantity is much larger than the relative distance among the eigenvalues in the cluster, the whole cluster can be viewed as a single semi-simple eigenvalue at the current step. As a result, RQI converges cubically in the first iteration, and it yields the new eigenvector approximation x_1 such that $\angle(x_1, \mathcal{V}) \approx 10^{-6}$, and x_1 is not dominated by any particular eigenvector. Therefore, certain eigenvalues in this cluster could be as close to the new Rayleigh quotient $\rho_1 = (x_1^T A x_1)/(x_1^T x_1)$ as to other ones in the cluster. At this stage, the clustered eigenvalues can no longer be viewed as one semi-simple eigenvalue; instead, since (ρ_1, x_1) is *not* an accurate approximation to any of the eigenpairs in the cluster, it may take a few iterations for RQI to gradually move towards one particular eigenpair. Thereafter, RQI begins to exhibit asymptotic cubic convergence again.

To illustrate the above analysis, we use the MATLAB code developed by Smith and Knyazev [24] to generate an ideal discrete Laplacian A_{idl} defined in a cube. The construction is done using the standard 7-point finite difference scheme, with 25 uniformly distributed nodes in each direction, and with Dirichlet boundary conditions. A_{idl} is of order 15625, with a semi-simple eigenvalue $\lambda_{idl} = 0.2026661316$ of multiplicity 6. We introduce a small symmetric perturbation to A_{idl}, achieved by the MATLAB command

$$\texttt{A} = \texttt{A\_idl} + \texttt{1e-6} * \texttt{sprandsym(A\_idl)},$$

Table 4 The tight cluster of six eigenvalues of a practical discrete Laplacian

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
0.2026660102	0.2026660903	0.2026661136	0.2026661687	0.2026661947	0.2026662343

to mimic the break of domain symmetry incurred by the discretization by a standard mesh generator. Table 4 shows the six clustered eigenvalues of A.

Using MATLAB notation, let `w = V*randn(6,1)` be a vector lying in the eigenspace $\mathcal{V} = \text{span}\{v_1, \ldots, v_6\}$ associated with $\{\lambda_1, \ldots, \lambda_6\}$, `z = randn(n,1)`, and `x_0 = w+1e-3*z` be the initial RQI iterate. We run RQI 50 times with such a random x_0, and we terminate RQI once the eigenvalue approximation error $\min_{i=1,\ldots,6} |\rho_k - \lambda_i| < 10^{-15}$.

Figure 1 illustrates the convergence of RQI, including the fastest, the slowest and a typical convergence history obtained in the 50 experiments. The eigenvalue approximation errors e_k (y-axis) are plotted against the RQI iteration step k (x-axis). The top part of Fig. 1 shows the local convergence of RQI towards the semi-simple eigenvalue λ_{idl} of A_{idl}. In a vast majority cases, RQI converges cubically in two iterations, with $e_0 \approx 10^{-2}$, $e_1 \approx 10^{-6}$ and $e_2 \approx 10^{-16}$. On the other hand, the bottom part of Fig. 1 shows that RQI generally converges much slower when used to resolve one of the tightly clustered simple eigenvalues of A. It converges cubically in the first step, but then it usually exhibits slow convergence for a few iterations before entering the phase of fast asymptotic convergence. The two scenarios are almost identical in the first iteration, but they differ considerably thereafter. To maintain a consistently fast convergence for this cluster of eigenvalues, one could use block variants of Newton-like methods to resolve the whole eigenspace, at the cost of solving six linear systems per iteration step.

4 Convergence for defective eigenvalues

We have shown in Sect. 3 that Newton-like methods converge towards semi-simple eigenvalues at least quadratically. In this section, we study the computation of defective eigenvalues. This is naturally more challenging. We show that the local convergence of Newton-like methods towards a defective eigenpair is generally only linear.

4.1 Inverse iteration

The proof of the linear convergence of inverse iteration for defective eigenvalues is similar to that for semi-simple eigenvalues. Let λ be a defective eigenvalue of $T(\cdot)$ with geometric multiplicity $geo_T(\lambda) = J$. This means that there are exactly J right and J left Jordan chains associated with λ, namely

$$\{\varphi_{1,0}, \ldots, \varphi_{1,m_1-1}\}, \{\varphi_{2,0}, \ldots, \varphi_{2,m_2-1}\}, \ldots, \{\varphi_{J,0}, \ldots, \varphi_{J,m_J-1}\}, \text{ and}$$
$$\{\psi_{1,0}, \ldots, \psi_{1,m_1-1}\}, \{\psi_{2,0}, \ldots, \psi_{2,m_2-1}\}, \ldots, \{\psi_{J,0}, \ldots, \psi_{J,m_J-1}\}. \quad (30)$$

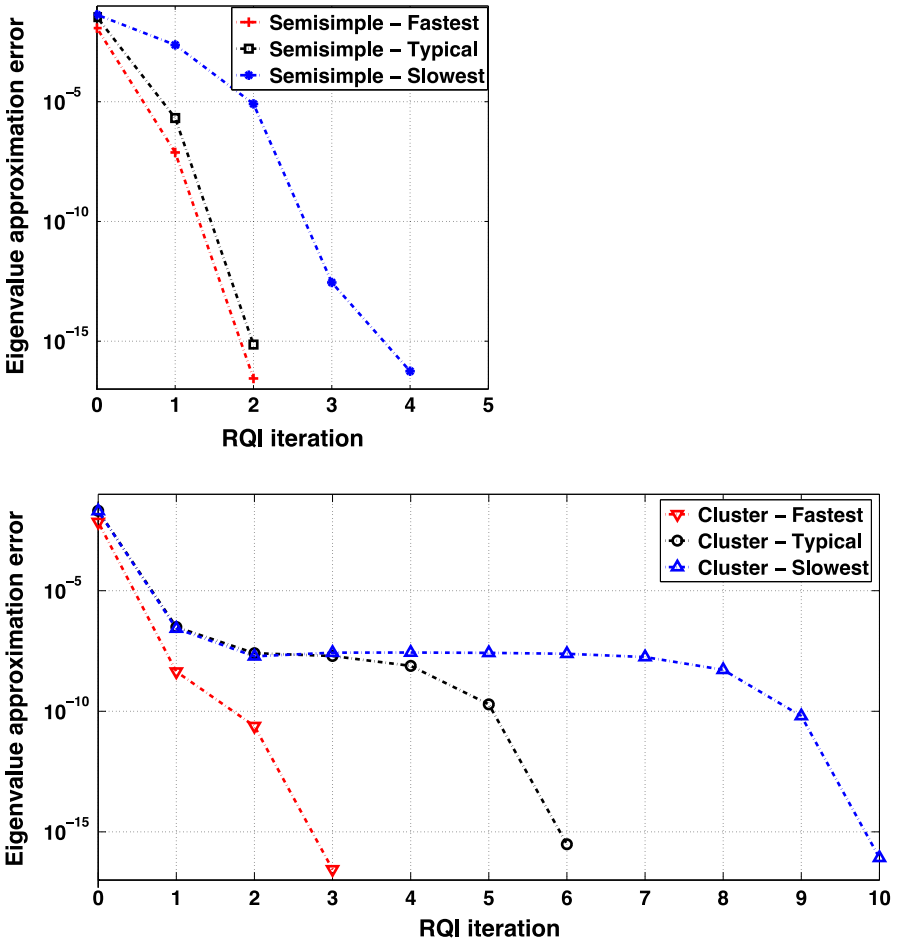

Fig. 1 The convergence history of RQI with different initial iterate (semilog plots of the eigenvalue approximation error $\min_{1\le i\le 6}|\rho_k-\lambda_i|$). *Top* towards the semi-simple λ_{idl} of the ideal Laplacian. *Bottom* towards one simple eigenvalue of the cluster $\{\lambda_1,\ldots,\lambda_6\}$ of a practical Laplacian

Consider a singular value decomposition of $T(\lambda)$ in the following form

$$Y^*T(\lambda)X=\begin{bmatrix}0_J & 0\\ 0 & \Sigma_{n-J}\end{bmatrix},$$

where $X=[X_J\ X_{n-J}]$ and $Y=[Y_J\ Y_{n-J}]$ are unitary matrices, $X_J\in\mathbb{C}^{n\times J}$ and $Y_J\in\mathbb{C}^{n\times J}$ have orthonormal columns forming a basis of $\mathrm{span}\{\varphi_{1,0},\ldots,\varphi_{J,0}\}$ and $\mathrm{span}\{\psi_{1,0},\ldots,\psi_{J,0}\}$, respectively, and Σ_{n-J} has all nonzero singular values of $T(\lambda)$ on its diagonal. Therefore, there exist nonsingular $K_J, M_J\in\mathbb{C}^{J\times J}$ such that $X_J=[\varphi_{1,0}\ \ldots\ \varphi_{J,0}]K_J$ and $Y_J=[\psi_{1,0}\ \ldots\ \psi_{J,0}]M_J$. To simplify the analysis, assume that v is a candidate right eigenvector that can be represented as a linear combination of $\{\varphi_{\cdot,0}\}$ associated with Jordan chains of length ≥ 2. It turns out that this assumption

is important for the development of a clear conclusion of the linear convergence of inverse iteration for defective eigenvalues. From (10), we have

$$Y_J^* T'(\lambda) v = M_J^* [\psi_{1,0} \ \dots \ \psi_{J,0}]^* T'(\lambda) v = 0,$$

and therefore

$$\begin{bmatrix} Y^* & \\ & 1 \end{bmatrix} \begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix} \begin{bmatrix} X & \\ & 1 \end{bmatrix} = \begin{bmatrix} Y^*T(\lambda)X & Y^*T'(\lambda)v \\ u^*X & 0 \end{bmatrix} = \begin{bmatrix} 0_J & 0 & 0 \\ 0 & \Sigma_{n-J} & Y_{n-J}^* T'(\lambda)v \\ u^*X_J & u^*X_{n-J} & 0 \end{bmatrix}.$$

Let $h = [h_a^T \ h_b^T \ h_c]^T$ be a vector in the null space of the above square matrix, where $h_a \in \mathbb{C}^J$, $h_b \in \mathbb{C}^{n-J}$ and $h_c \in \mathbb{C}$. Then

$$\begin{bmatrix} 0_J & 0 & 0 \\ 0 & \Sigma_{n-J} & Y_{n-J}^* T'(\lambda)v \\ u^*X_J & u^*X_{n-J} & 0 \end{bmatrix} \begin{bmatrix} h_a \\ h_b \\ h_c \end{bmatrix} = \begin{bmatrix} 0 \\ \Sigma_{n-J} h_b + Y_{n-J}^* T'(\lambda) v h_c \\ u^*X_J h_a + u^*X_{n-J} h_b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

It follows from the second block row that $h_b = -\Sigma_{n-J}^{-1} Y_{n-J}^* T'(\lambda) v h_c$, and thus the last row is equivalent to $u^*X_J h_a - u^*X_{n-J}\Sigma_{n-J}^{-1} Y_{n-J}^* T'(\lambda) v h_c = 0$. Since u specifies the scaling of $v = [\varphi_{1,0} \ \dots \ \varphi_{J,0}] d_v \in \text{range}(X_J)$ such that $u^*v = 1$, we have $u^*X_J \neq 0$. Without loss of generality, assume that $u^*X_J = [\gamma_1 \ \dots \ \gamma_J]$ where $\gamma_J \neq 0$. Then, depending on whether $h_c = 0$, we can determine h_a, h_b and h as follows.

$$h = \begin{bmatrix} h_a \\ h_b \\ h_c \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{\gamma_1}{\gamma_J} \\ 0_{n-J} \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{\gamma_{J-1}}{\gamma_J} \\ 0_{n-J} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \chi \\ f \\ 1 \end{bmatrix} \right\}, \tag{31}$$

where $\chi = \frac{1}{\gamma_J} u^*X_{n-J}\Sigma_{n-J}^{-1} Y_{n-J}^* T'(\lambda) v \in \mathbb{C}$ and $f = -\Sigma_{n-J}^{-1} Y_{n-J}^* T'(\lambda) v \in \mathbb{C}^{n-J}$. Therefore the null space $\mathcal{N}_1$ of the Jacobian at (λ, v) is of dimension J — *one dimension larger* than it is in the semi-simple case; see (21). In fact,

$$\mathcal{N}_1 \equiv \text{null}\left(\begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix}\right) = \begin{bmatrix} X & \\ & 1 \end{bmatrix} \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{\gamma_1}{\gamma_J} \\ 0_{n-J} \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{\gamma_{J-1}}{\gamma_J} \\ 0_{n-J} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \chi \\ f \\ 1 \end{bmatrix}\right\}$$

$$= \text{span}\left\{\begin{bmatrix} X\mathbf{e_1} - \frac{\gamma_1}{\gamma_J} X\mathbf{e_J} \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} X\mathbf{e_{J-1}} - \frac{\gamma_{J-1}}{\gamma_J} X\mathbf{e_J} \\ 0 \end{bmatrix}\right\}$$

$$\oplus \text{ span}\left\{\begin{bmatrix} \chi X\mathbf{e_J} + X_{n-J} f \\ 1 \end{bmatrix}\right\}. \tag{32}$$

The complementary space $\mathcal{M}_1$ of $\mathcal{N}_1$ can be defined as follows

$$\mathcal{M}_1 = \text{span}\left\{\begin{bmatrix} X\mathbf{e_J} \\ 0 \end{bmatrix}\right\} \oplus \text{range}\left(\begin{bmatrix} X_{n-J} \\ 0 \end{bmatrix}\right), \tag{33}$$

so that $\dim(\mathcal{M}_1) = n - J + 1$, and $\mathbb{C}^{n+1} = \mathcal{N}_1 \oplus \mathcal{M}_1$. From (32) and (33), it follows by the definitions of $\mathcal{N}_1$ and $\mathcal{M}_1$ that

$$\mathcal{M}_2 = \text{range}\left(\begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix}\right) = \begin{bmatrix} T(\lambda) & T'(\lambda)v \\ u^* & 0 \end{bmatrix} \mathcal{M}_1$$

$$= \text{span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} \oplus \text{range}\left(\begin{bmatrix} T(\lambda)X_{n-J} \\ u^* X_{n-J} \end{bmatrix}\right), \ (u^* X\mathbf{e_J} = \gamma_J \neq 0 \text{ by assumption}) \tag{34}$$

so that $\dim(\mathcal{M}_2) = n - J + 1$. There is considerable freedom to choose the complementary space $\mathcal{N}_2$; for example, one can define

$$\mathcal{N}_2 = \text{range}\left(\begin{bmatrix} \big(T(\lambda)X_{n-J}\big)^{\perp} \\ 0 \end{bmatrix}\right), \tag{35}$$

where $\big(T(\lambda)X_{n-J}\big)^{\perp}$ consists of J vectors orthogonal to range$\big(T(\lambda)X_{n-J}\big)$, such that $\mathcal{M}_2$ and $\mathcal{N}_2$ are orthogonal complements of each other.

Similar to the analysis for semi-simple eigenvalues, let $e_k = [x_k; \mu_k] - [v; \lambda]$ be the error between the Newton iterate and the *particular eigenpair* (λ, v) in the kth step, $\mathcal{P}_{\mathcal{N}_1}$ be the projector onto $\mathcal{N}_1$ along $\mathcal{M}_1$, and $\mathcal{P}_{\mathcal{M}_1} = I - \mathcal{P}_{\mathcal{N}_1}$. To complete the analysis, we make the following assumption.

Assumption 7 For any sequence of Newton iterate errors $\{e_k\}$ whose components lying in $\mathcal{N}_1$ converge to zero linearly, their components lying in any one-dimensional subspace of $\mathcal{N}_1$, if any, also converge to zero linearly.

The assumption states that the Newton iterate errors lying in any subspace of the kernel of the Jacobian exhibit qualitatively the same behavior, and no "special" subspace exists in which the iterate errors converge at a rate of higher order than they do in other subspaces. This assumption of isotropism is consistent with our observations of the Newton iterate errors in experiments.

Since the Jacobian is singular at (λ, v), Theorem 4 shows that $\|\mathcal{P}_{\mathcal{N}_1}(e_k)\|$ and $\|\mathcal{P}_{\mathcal{M}_1}(e_k)\|$, respectively, converge to zero linearly and quadratically. By Assumption 7, since span $\left\{\begin{bmatrix} \chi X\mathbf{e_J} + X_{n-J} f \\ 1 \end{bmatrix}\right\}$ is a subspace of $\mathcal{N}_1$ (see (32)), the iterate errors $\{e_k\}$ have a component lying in this one-dimensional space that converges linearly. The last entry "1" of this basis vector represents the eigenvalue approximation error, and $X_{n-J} f$ represents an eigenvector approximation error (range(X_J) is the desired eigenspace). As a result, inverse iteration converges towards λ and its eigenspace linearly.

4.2 Numerical experiments for inverse iteration

Numerical results are provided in this section to illustrate the linear convergence of inverse iteration for defective eigenvalues. We choose two problems from NLEVP, and we constructed four problems artificially, because benchmark problems in the literature with defective eigenvalues are rather limited.

A description of these problems is given in Table 5. For example, the problem df_art_m1m2, which defines a function $T(\mu) \in \mathbb{C}^{256\times 256}$, is an artificially constructed, truly nonlinear problem. It has a defective eigenvalue $\lambda = 0$, and the algebraic and the geometric multiplicities of λ are $alg_T(\lambda) = 5$ and $geo_T(\lambda) = 4$, respectively; the length of the shortest and the longest Jordan chains associated with λ are $\min\{m_i\} = 1$ and $\max\{m_i\} = 2$, respectively.

The four artificially constructed problems are as follows. For $jordan3$,

$$T(\mu) = G_A^* D(\mu) G_B, \qquad \text{where}$$
$$D(\mu) = \begin{bmatrix} \mu-2 & -1 & & & \\ & \mu-2 & -1 & & \\ & & \mu-2 & & \\ & & & \mu I_{253} - diag([0;\ 1;\ c_1]) & \end{bmatrix},$$

Table 5 Description of the test problems for defective eigenvalues

Problem	Source	Type	Size	Eigenvalue	$alg_T(\lambda)$	$geo_T(\lambda)$	$\min\{m_i\}$	$\max\{m_i\}$
$time\_delay$	NLEVP	nep	3	$3\pi i$	2	1	2	2
$jordan3$	Artificial	lep	256	2	3	1	3	3
df_art_m1m2	Artificial	nep	256	0	5	4	1	2
df_art_m1m3	Artificial	nep	256	0	5	3	1	3
df_art_m2m3	Artificial	nep	256	0	5	2	2	3
$mirror$	NLEVP	pep	9	0	9	7	1	2

and for the other three problems,

$$T(\mu) = G_A^* \begin{bmatrix} D_5(\mu) & \\ & diag([c_0 + \mu c_1 + \mu^2 c_2]) \end{bmatrix} G_B, \qquad \text{where}$$

$$D_5(\mu) = \begin{bmatrix} e^{\mu}-1 & 1-\tan^{-1}(2\mu) & & & & \\ & 2\sin(\mu) & & & & \\ & & -5\ln(1+\mu) & & & \\ & & & 8\mu & & \\ & & & & \tan^{-1}(3\mu) & \end{bmatrix}$$

for df_art_m1m2,

$$D_5(\mu) = \begin{bmatrix} e^{\mu}-1 & 1-\tan^{-1}(2\mu) & & & \\ & 2\sin(\mu) & \cos(\mu^2) & & \\ & & -5\ln(1+\mu) & & \\ & & & 8\mu & \\ & & & & \tan^{-1}(3\mu) \end{bmatrix}$$

for df_art_m1m3, and

$$D_5(\mu) = \begin{bmatrix} e^{\mu}-1 & 1-\tan^{-1}(2\mu) & & & \\ & 2\sin(\mu) & \cos(\mu^2) & & \\ & & -5\ln(1+\mu) & & \\ & & & 8\mu & e^{-2\mu} \\ & & & & \tan^{-1}(3\mu) \end{bmatrix}$$

df_art_m2m3, respectively, and G_A, G_B and c_i $(i = 0, 1, 2)$ are random matrices described in Sect. 3.2. Among the six problems, $time\_delay$, $jordan3$ and df_art_m2m3 have Jordan chains of minimum length ≥ 2, for which we showed in the previous section that inverse iteration converges linearly. The other three problems with a Jordan chain of length 1 are not directly covered by our analysis. Nevertheless, our experiments show that inverse iteration exhibits linear convergence for these problems as well.

To perform the test, we first generated an initial eigenpair approximation (μ_0, x_0) in the same manner as we did for semi-simple eigenvalues. We then ran inverse iteration with (μ_0, x_0), and we found that the algorithm does converge linearly until the eigenvalue and eigenvector approximation errors decrease to the magnitude around $\epsilon^{1/\max\{m_i\}}$ (ϵ is the machine precision)—the highest precision one can achieve for defective eigenvalues; see [17,18] for details. An estimate of the order of convergence was obtained by applying the least-squares line formula to the sequence $(\log \|r_k\|, \log \|r_{k+1}\|)$ for $k = 1, 2, \ldots$, where $r_k = T(\mu_k)x_k$ is the kth step eigenresidual. We did not use r_0, because the residual norm $\|r_1\|$ is usually significantly smaller than $\|r_0\|$, and thus the inclusion of r_0 produces considerable noise in our estimate. The linear convergence of inverse iteration for all the six defective problems are summarized in Table 6.

Table 6 Estimated order of convergence η of inverse iteration for a defective λ

Problem	$\angle(x_0, v)$	# Iters	Estimated η
time_delay	10^{-3}	20	1.001
jordan3	5×10^{-3}	14	0.984
df_art_m1m2	5×10^{-2}	17	1.001
df_art_m1m3	10^{-2}	14	0.978
df_art_m2m3	10^{-2}	15	0.969
mirror	10^{-3}	12	1.005

4.3 Rayleigh functional iteration and single-vector JD

The linear convergence of inverse iteration for defective eigenpairs is much less satisfactory than the quadratic convergence for semi-simple eigenpairs. Unfortunately, as we show in this section, other classical Newton-like methods, including standard RFI/JD and their two-sided variants, also converge linearly in general in the defective case.

To see this, we first assume that the defective eigenvalue λ has only one right Jordan chain $\{\varphi_{1,0}, \ldots, \varphi_{1,m-1}\}$ so that $alg_T(\lambda) = m$ and $geo_T(\lambda) = 1$. Let (μ, x) be an eigenpair approximation, and $p = T(\mu)^{-1}T'(\mu)x$ the unnormalized new eigenvector approximation computed by any classical Newton-like method. Recalling the structure of $T(\mu)^{-1}$ from Theorem 3, we have

$$\begin{aligned} p = T(\mu)^{-1}T'(\mu)x &= \sum_{h=0}^{m-1} \frac{\sum_{s=0}^{h} \langle T'(\mu)x, \psi_{1,s}\rangle \varphi_{1,h-s}}{(\mu-\lambda)^{m-h}} + Q(\mu)T'(\mu)x \\ &= \sum_{i=0}^{m-1} \frac{\langle T'(\mu)x, \psi_{1,0}\rangle}{(\mu-\lambda)^{m-i}}\varphi_{1,i} + \sum_{i=0}^{m-2} \frac{\langle T'(\mu)x, \psi_{1,1}\rangle}{(\mu-\lambda)^{m-1-i}}\varphi_{1,i} + \cdots \\ &\quad + \sum_{i=0}^{1} \frac{\langle T'(\mu)x, \psi_{1,m-2}\rangle}{(\mu-\lambda)^{2-i}}\varphi_{1,i} + \frac{\langle T'(\mu)x, \psi_{1,m-1}\rangle}{(\mu-\lambda)}\varphi_{1,0} + Q(\mu)T'(\mu)x \\ &= \sum_{j=0}^{m-1}\sum_{i=0}^{m-j-1} \frac{\langle T'(\mu)x, \psi_{1,j}\rangle}{(\mu-\lambda)^{m-j-i}}\varphi_{1,i} + Q(\mu)T'(\mu)x. \end{aligned} \tag{36}$$

To analyze the direction of p, assume that the eigenvector $\varphi_{1,0}$ and the generalized eigenvector $\varphi_{1,1}$ are not parallel. Then for every $j < m-1$, consider the following component shown in the last equation of (36)

$$\sum_{i=0}^{m-j-1} \frac{\langle T'(\mu)x, \psi_{1,j}\rangle}{(\mu-\lambda)^{m-j-i}}\varphi_{1,i} = \frac{\langle T'(\mu)x, \psi_{1,j}\rangle}{(\mu-\lambda)^{m-j}} \sum_{i=0}^{m-j-1} (\mu-\lambda)^i \varphi_{1,i}. \tag{37}$$

The above expression shows clearly that for every $j = 0, 1, \ldots, m-1$ in (36), the ratio between the magnitude of the error component (in the direction of $\varphi_{1,1}$) and the magnitude of the eigenvector component (in the direction of $\varphi_{1,0}$) in (37) is of the

order of $\mathcal{O}(\mu-\lambda)$. In particular, assume that for $j=1$, $\langle T'(\lambda)\varphi_{1,0},\psi_{1,1}\rangle \neq 0$, so that $\langle T'(\mu)x,\psi_{1,1}\rangle \approx \langle T'(\lambda)\varphi_{1,0},\psi_{1,1}\rangle = \mathcal{O}(1)$ for μ sufficiently close to λ and x sufficiently close to $\varphi_{1,0}$ in direction. Then it can be shown without difficulty that p has a significant component of the form

$$\frac{\varphi_{1,0}+(\mu-\lambda)\varphi_{1,1}+\mathcal{O}\left((\mu-\lambda)^2\right)}{(\mu-\lambda)^{m-1}} \tag{38}$$

up to a scaling factor.

As a result, the local convergence rate of Newton-like methods depends on $|\mu-\lambda|$, which is usually connected to the accuracy of the eigenvector approximation x. Let the accuracy of x be represented by $\sin\angle(x,\varphi_{1,0})$, and assume that $|\mu-\lambda|=\mathcal{O}(\sin^{\eta}\angle(x,\varphi_{1,0}))$. Then the convergence of the Newton-like method is of order η. In particular, for RFI where $\mu=\rho_F(x;T,y)$ is the Rayleigh functional value, the following Proposition shows that $\eta=1$ in general for defective eigenvalues, and thus RFI converges only linearly.

Proposition 8 *Let λ be a defective eigenvalue of the holomorphic operator $T(\cdot): U\to\mathbb{C}^{n\times n}$, and φ and ψ be a corresponding unit right and a unit left eigenvector, respectively. Assume that either φ or ψ is associated with a Jordan chain of length ≥ 2. Let $x=\varphi\cos\alpha+\varphi_\perp\sin\alpha$ and $y=\psi\cos\beta+\psi_\perp\sin\beta$, where $\alpha,\beta<\frac{\pi}{2}$, $\varphi_\perp\perp\operatorname{null}T(\lambda)$ and $\psi_\perp\perp\operatorname{null}(T(\lambda))^*$ are unit vectors, and thus $\|x\|=\|y\|=1$. Assume that $y^*T'(\lambda)x\neq 0$. Let $\rho=\rho_F(x;T,y)$ be the Rayleigh functional value closest to λ such that $y^*T(\rho)x=0$. Then for sufficiently small α,*

$$\begin{aligned}|\rho-\lambda| &\leq \frac{2\|T(\lambda)\|\,|\sin\alpha\sin\beta|}{|y^*T'(\lambda)x|}\\ &= \frac{2\|T(\lambda)\|\,|\sin\alpha\sin\beta|}{\left|\overline{\cos\beta}\sin\alpha\psi^*T'(\lambda)\varphi_\perp+\overline{\sin\beta}\cos\alpha\psi_\perp^*T'(\lambda)\varphi+\overline{\sin\beta}\sin\alpha\psi_\perp^*T'(\lambda)\varphi_\perp\right|}.\end{aligned} \tag{39}$$

Assume in addition that $\psi^*T'(\lambda)\varphi_\perp$, $\psi_\perp^*T'(\lambda)\varphi$ and $\psi_\perp^*T'(\lambda)\varphi_\perp$ are bounded away from zero. Then, for sufficiently small α, $|\rho-\lambda|\leq\mathcal{O}(\sin\alpha)$.

Proof The first part of the proposition through (39) was shown in [27, Section 4.3] using the Newton-Kantorovich Theorem. The goal here is to show that, in contrast to the scenario for semi-simple eigenvalues, we have $|\rho-\lambda|\leq\mathcal{O}(\sin\alpha)$ (instead of $\mathcal{O}(\sin\alpha\sin\beta)$), whether the symmetry of $T(\lambda)$ exists or not.

First, assume that y is not a good left eigenvector approximation, i.e., $\sin\beta=\mathcal{O}(1)$. Since $\sin\alpha$ is small, we have $2\|T(\lambda)\|\,|\sin\alpha\sin\beta|=\mathcal{O}(\sin\alpha)$, and from (39), $|y^*T'(\lambda)x|=|\overline{\sin\beta}\cos\alpha\psi_\perp^*T'(\lambda)\varphi+\mathcal{O}(\sin\alpha)|=\mathcal{O}(1)$. Therefore it follows that

$$|\rho-\lambda|\leq\frac{2\|T(\lambda)\|\,|\sin\alpha\sin\beta|}{|y^*T(\lambda)x|}=\mathcal{O}(\sin\alpha). \tag{40}$$

Now assume that y is a good left eigenvector approximation such that $\sin\beta=\mathcal{O}(\sin\alpha)$; in particular, suppose the local symmetry of $T(\cdot)$ exists, and thus we choose

$y = x$ or $y = \bar{x}$ to generate the two-sided Rayleigh functional value. In this case, $|\sin\alpha| = |\sin\beta|$, and it follows from (39) that

$$2\|T(\lambda)\||\sin\alpha\sin\beta| = \mathcal{O}(\sin^2\alpha), \text{ and } |y^*T'(\lambda)x| = \mathcal{O}(\sin\alpha).$$

Therefore, (40) still holds. □

Remark. Proposition 8 shows that for a defective λ, the use of a left eigenvector approximation for the Rayleigh functional ρ_F does not generate an eigenvalue approximation of second order accuracy, as is achieved for simple and semi-simple eigenvalues. This lack of high accuracy is attributed to the fact that $\psi^*T'(\lambda)\varphi = 0$. Moreover, the use of a left eigenvector approximation for ρ_F could introduce additional complications. Namely, since $|y^*T'(\lambda)x| = \mathcal{O}(\sin\alpha)$ from (39), it follows that $|y^*T'(\rho)x| \le |y^*T'(\lambda)x| + \mathcal{O}(\rho - \lambda) = \mathcal{O}(\sin\alpha)$. Thus there is a risk that $y^*T'(\rho)x \neq 0$, a critical condition required by the definition of Rayleigh functionals (see [23]), may be violated, at least numerically. As a result, we see from (17) that the small magnitude of $y^*T'(\rho)x$ could introduce numerical instabilities in the projector $\Pi_k^{(1)}$ for single-vector JD. We therefore recommend using y far from a left eigenvector approximation to compute the Rayleigh functional value for a defective λ.

From Proposition 8, it follows immediately from (37) and (38) that RFI converges linearly towards defective eigenvalues. In addition, assuming that the Rayleigh functional value $\rho = \rho_F(x; T, y)$ satisfies $\langle T'(\rho)x, y\rangle \neq 0$, then the single-vector JD (17) also converges linearly in this case, because it is mathematically equivalent to RFI. We summarize this result as follows.

Theorem 9 *Let λ be a defective eigenvalue of the holomorphic operator $T(\cdot) : U \to \mathbb{C}^{n\times n}$ with exactly one left and one right Jordan chains $\{\varphi_{1,0}, \ldots, \varphi_{1,m-1}\}$ and $\{\psi_{1,0}, \ldots, \psi_{1,m-1}\}$, where $\varphi_{1,0}$ and $\varphi_{1,1}$ are not in the same direction, and $\langle T'(\lambda)\varphi_{1,0}, \psi_{1,1}\rangle \neq 0$. Let (ρ_0, x_0) be an initial eigenpair approximation with $\angle(x_0, \varphi_{1,0})$ sufficiently small, and $\rho_k = \rho_F(x_k; T, y_k)$ be the Rayleigh functional value closest to λ. Assume that the upper bound* (40) *in* Proposition 8 *are qualitatively sharp, i.e., $|\rho_k - \lambda| = \mathcal{O}(\sin\angle(x_k, \varphi_{1,0}))$. Then RFI converges locally towards $(\lambda, \varphi_{1,0})$ linearly. The same conclusion applies to single-vector JD* (17) *if, in addition, ρ_k is such that $y_k^*T'(\rho_k)x_k \neq 0$.*

Remark 10 Theorem 9 can be extended without difficulty to a defective λ with multiple Jordan chains. Without loss of generality, let $\{\varphi_{1,0}, \ldots, \varphi_{1,m-1}\}$ and $\{\psi_{1,0}, \ldots, \psi_{1,m-1}\}$ be one of the longest right and corresponding left Jordan chains. Assume that $\varphi_{1,0}$ is not parallel to $\varphi_{1,1}$, and $\langle T'(\lambda)\varphi_{1,0}, \psi_{1,1}\rangle \neq 0$. Given the structure of $T(\mu)^{-1}$ near a defective λ with multiple Jordan chains (11), as RFI proceeds, we see that the longest Jordan chains become increasingly more dominant over shorter Jordan chains in the eigenvector approximation x_k. The above analysis can be used verbatim to show that RFI and single-vector JD converges towards λ linearly.

Table 7 Estimated order of convergence η of RFI and TSRFI for a defective λ

Problem	$\angle(x_0, v)$	RFI		TSRFI	
		# Iters	Estimated η	# Iters	Estimated η
time_delay	10^{-3}	18	1.002	15	0.981
jordan3	5×10^{-3}	19	0.961	13	0.967
df_art_m1m2	5×10^{-2}	12	0.984	8	0.973
df_art_m1m3	10^{-2}	14	0.967	10	0.965
df_art_m2m3	10^{-2}	14	0.974	11	1.013
mirror	10^{-3}	10^{1}	**1.991**	10^{1}	**3.003**

[1] For problem "*mirror*", we used the more descriptive approach discussed in Sect. 3.2 to measure $s_1^{(j)} = \sin\angle(x_1^{(j)}, \text{null}\, T(\lambda))$ for the sequence $(\mu_0^{(j)}, x_0^{(j)})$ $(j = 1, 2, \ldots, 10)$, and find the slope of the linear least squares fit for $\left(\log s_0^{(j)}, \log s_1^{(j)}\right)$

4.4 Numerical experiments for RFI/JD

We illustrate with numerical experiments the typical linear convergence of RFI and single-vector JD for defective eigenvalues. We run RFI with an initial eigenpair approximation (μ_0, x_0), and we analyze $(\log \|r_k\|, \log \|r_{k+1}\|)$, where $r_k = T(\mu_k)x_k$ $(k = 1, 2, \ldots)$, for the estimated order of convergence. Here, note that the local symmetry does not exist for both test problems, and we choose $y_k = T'(\lambda)x_k$ to generate the value of the Rayleigh functional $\rho_F(x_k; T, y_k)$. Table 7 shows the error angle of the initial iterate, the number of iterations taken, and the estimated order of convergence.

We see clearly that RFI converges linearly for all the test problems except "*mirror*", for which RFI converges quadratically. The faster convergence is due to the special spectral structure of this problem. Specifically, it has two Jordan chains $\{\varphi_{1,0}, \varphi_{1,1}\} = \{\mathbf{e_6}, \mathbf{e_6}\}$ and $\{\varphi_{2,0}, \varphi_{2,1}\} = \{\mathbf{e_7}, \mathbf{e_7}\}$ of length 2, and five Jordan chains $\{\mathbf{e_1}\}$, $\{\mathbf{e_2}\}$, $\{\mathbf{e_4}\}$, $\{\mathbf{e_5}\}$ and $\{\mathbf{e_9}\}$ of length 1. For both longest Jordan chains, the generalized eigenvector $\varphi_{\cdot,2}$ equals the corresponding eigenvector $\varphi_{\cdot,1}$, violating the critical assumption in Theorem 9 and Remark 10 that $\varphi_{\cdot,0}$ not be parallel to $\varphi_{\cdot,1}$ for at least one longest Jordan chain.

In addition, we tested the convergence of the two-sided Rayleigh functional iteration (TSRFI), which converges cubically for simple eigenvalues; see, e.g., [22]. Table 7 shows that, except for "*mirror*", this algorithm also converges linearly in this setting, which is consistent with Proposition 8 and Theorem 9. As we have discussed, for a defective eigenvalue λ with a single Jordan chain, we generally have $|\rho_F(x; T, y) - \lambda| = \mathcal{O}\big(\angle(x, v)\big)$ (instead of $\mathcal{O}\big(\angle(x, v)^2\big)$), no matter whether y is a good left eigenvector approximation; consequently, TSRFI exhibits the same order of convergence (linear) as RFI. The unusual cubic convergence for "*mirror*" is again due to the special structure of Jordan chains associated with the defective eigenvalue λ.

5 Conclusion

We studied the local convergence of several classical single-vector Newton-like methods for the solution of a degenerate eigenvalue and a corresponding eigenvector of nonlinear algebraic eigenproblems of the form $T(\lambda)v = 0$. For semi-simple eigenvalues, these algorithms exhibit the same order of convergence as they do for simple ones. The convergence is generally quadratic; in addition, RFI/JD with appropriate use of the two-sided Rayleigh functional can achieve cubic convergence for problems with symmetry. Intuitively, semi-simple eigenvalues can be considered as a group of simple eigenvalues sharing the identical eigenvalue; therefore the convergence results for simple eigenvalues also hold in the semi-simple case. The convergence analysis for defective eigenvalues is more complicated due to their spectral structure. We showed the typical linear convergence of inverse iteration, standard RFI/JD and their two-side variants in the defective case. Our analyses are illustrated by numerical experiments. In a companion paper [29], we study accelerated algorithms that converge quadratically for defective eigenvalues.

Acknowledgments We thank the referees for their comments and suggestions, which helped improve our presentation.

References

1. Asakura, J., Sakurai, T., Tadano, H., Ikegami, T., Kimura, K.: A numerical method for polynomial eigenvalue problems using contour integral. Japan J. Indus. Appl. Math. **27**, 73–90 (2010)
2. Betcke, T., Higham, N.J., Mehrmann, V., Schröder, C., Tisseur, F.: NLEVP: a collection of nonlinear eigenvalue problems. ACM Trans Math Softw, **39** (2013) article No. 7
3. Betcke, T., Voss, H.: A Jacobi-Davidson type projection method for nonlinear eigenvalue problems. Future Gener. Comput. Syst. **20**, 363–372 (2004)
4. Beyn, W.-J.: An integral method for solving nonlinear eigenvalue problems. Linear Alg. Appl. **436**, 3839–3863 (2012)
5. Decker, D.W., Kelley, C.T.: Newton's method at singular points I. SIAM J. Numer. Anal. **17**, 66–70 (1980a)
6. Decker, D.W., Kelley, C.T.: Newton's method at singular points II. SIAM J. Numer. Anal. **17**, 465–471 (1980b)
7. Decker, D.W., Keller, H.B., Kelley, C.T.: Convergence rates for Newton's method at singular points. SIAM J. Numer. Anal. **20**, 296–314 (1983)
8. Hale, N., Higham, N.J., Trefethen, L.N.: Computing A^{α}, $log(A)$, and related matrix functions by contour integrals. SIAM J. Numer. Anal. **46**, 2505–2523 (2008)
9. Jarlebring, E., Michiels, W.: Analyzing the convergence factor of residual inverse iteration. BIT Numer. Math. **51**, 937–957 (2011)
10. Kozlov, V., Maz'ia, V.: Differential Equations with Operator Coefficients. Springer, Berlin (1999)
11. Kressner, D.: A block Newton method for nonlinear eigenvalue problems. Numerische Mathematik **114**, 355–372 (2009)
12. Mackey, D.S., Mackey, N., Mehl, C., Mehrmann, V.: Vector spaces of linearizations for matrix polynomials. SIAM J. Matrix Anal. Appl. **28**, 971–1004 (2006)
13. Mackey, D.S., Mackey, N., Mehl, C., Mehrmann, V.: Structured polynomial eigenvalue problems: good vibrations from good linearizations. SIAM J. Matrix Anal. Appl. **28**, 1029–1051 (2006)
14. The Matrix Market. http://math.nist.gov/MatrixMarket/, NIST, (2007)
15. Mehrmann, V., Schröder, C.: Nonlinear eigenvalue and frequency response problems in industrial practice. J. Math. Indus. 1 (2011), article No. 7
16. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods. Mitteilungen der Gesellschaft für Angewandte Mathematik und Mechanik **27**, 121–151 (2005)

17. Moro, J., Burke, J.V., Overton, M.L.: On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. SIAM J. Matrix Anal. Appl. **18**, 793–817 (1997)
18. Moro, J., Dopico, F.M.: First order eigenvalue perturbation theory and the newton diagram. In: Drmac, Z., Hari, V., Sopta, L., Tutek, Z., Veselic K. (eds.) Applied Mathematics and Scientific Computing, pp. 143–175. Kluwer Academic Publishers (2003)
19. Neumaier, A.: Residual inverse iteration for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. **22**, 914–923 (1985)
20. Osborne, M.R.: Inverse iteration, Newton's method, and non-linear eigenvalue problems. The Contributions of Dr. J. H. Wilkinson to Numerical Analysis, Symposium Proceedings Series, 19, pp. 21–53, The Institute of Mathematics and its Applications, Southend-on-Sea, Essex (1978)
21. Ruhe, A.: Algorithms for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. **10**, 674–689 (1973)
22. Schreiber, K.: Nonlinear eigenvalue problems: Newton-type methods and nonlinear Rayleigh functionals, Ph.D thesis, Department of Mathematics, TU Berlin, (2008)
23. Schwetlick, H., Schreiber, K.: Nonlinear Rayleigh functionals. Linear Alg. Appl. **436**, 3991–4016 (2012)
24. Smith, B.C., Knyazev, A.V.: Sparse (1–3)d Laplacian on a rectangular grid with exact eigenpairs, MATLAB Central File Exchange, http://www.mathworks.com/matlabcentral/fileexchange/27279-laplacian-in-1d-2d-or-3d/content/laplacian.m
25. Spence, A., Poulton, C.: Photonic band structure calculations using nonlinear eigenvalue techniques. J. Comput. Phys. **204**, 65–81 (2005)
26. Su, Y., Bai, Z.: Solving rational eigenvalue problems via linearization. SIAM J. Matrix Anal. Appl. **32**, 201–216 (2011)
27. Szyld, D.B., Xue, F.: Local convergence analysis of several inexact Newton-type algorithms for general nonlinear eigenvalue problems. Numerische Mathematik **123**, 333–362 (2013)
28. Szyld, D.B., Xue, F.: Several properties of invariant pairs of nonlinear algebraic eigenvalue problems. IMA J Numer Anal (2014). doi:10.1093/imanum/drt026
29. Szyld, D.B., Xue, F.: Local convergence of Newton-like methods for degenerate eigenvalues of nonlinear eigenproblems. II. Accelerated algorithms. Numer Math. (2014). doi:10.1007/s00211-014-0640-2
30. Tisseur, F., Meerbergen, K.: The quadratic eigenvalue problem. SIAM Rev. **43**, 234–286 (2001)