A POSITIVITY PRESERVING INEXACT NODA ITERATION
FOR COMPUTING THE SMALLEST EIGENPAIR OF A LARGE
IRREDUCIBLE M-MATRIX

ZHONGXIAO JIA, WEN-WEI LIN, AND CHING-SUNG LIU

Abstract. In this paper, based on the Noda iteration, we present inexac t
Noda iterations (INI), to find the smallest eigenvalue and the associated pos-
tive eigenvector of a large irreducible nonsingular M-matrix. The positivity
of approximations is critical in applications, and if the approximations lose
the positivity then they will be physically meaningless. We propose two dif-
ferent inner tolerance strategies for solving the inner linear systems involved,
and prove that the resulting INI algorithms are globally linear and superlin-
ear with the convergence order \(\frac{1+\sqrt{5}}{2}\), respectively. The proposed INI
algorithms are structure preserving and maintains the positivity of approximate
eigenvectors. We also revisit the exact Noda iteration and establish a new
quadratic convergence result. All the above is first done for the problem of
computing the Perron root and the positive Perron vector of an irreducible
nonnegative matrix and then adapted to that of computing the smallest eigen-
pair of the irreducible nonsingular M-matrix. Numerical examples illustrate
that the proposed INI algorithms are practical, and they always preserve the
positivity of approximate eigenvectors. We compare them with the positivity
non-preserving Jacobi–Davidson method and implicitly restarted Arnoldi
method, which often compute physically meaningless eigenvectors, and illus-
trate that the overall efficiency of the INI algorithms is competitive with and
can be considerably higher than the latter two methods.

1. Introduction

Irreducible nonsingular M-matrices are one of the most important classes of ma-
trices in applications, such as discretized PDEs, Markov chains \([1, 23]\) and electric
circuits \([26]\), and have been studied extensively in the literature; see, for instance,
In most applications, one is interested in finding the smallest eigenvalue $\lambda$ and the associated eigenvector $x$ of an irreducible nonsingular $M$-matrix $A \in \mathbb{R}^{n \times n}$.

$M$-matrices are closely related to nonnegative matrices. For instance, an $M$-matrix $A$ can be expressed in the form $A = \sigma I - B$ with a nonnegative matrix $B \geq 0$ and some constant $\sigma > \rho(B)$, the spectral radius of $B$; $A^{-1}$ is nonnegative. For more properties and a systematic account of $M$-matrices and nonnegative matrices, see [10, 4].

Nonnegative matrices have important applications in many areas [4], including economics, statistics and network theory. For a nonnegative matrix $B$, at least one of the eigenvalues of maximal magnitude is nonnegative and hence equal to the spectral radius $\rho(B)$. The corresponding eigenvectors $x$ satisfy $Bx = \rho(B)x$ and are called the Perron vectors of $B$ if they are nonnegative. The nonnegative $B$ always has at least one Perron vector. In applications, the Perron vectors play an important role, and they describe, e.g., an equilibrium, a probability distribution and an optimal network property [4]. Furthermore, if $B$ is irreducible, then the famous Perron-Frobenius theorem [10] states that $\rho(B)$ is simple and there is a positive eigenvector $x$ associated with $\rho(B)$. One is often interested in verifying the uniqueness and strict positivity of the Perron vector of the irreducible nonnegative matrix $B$ [10, 4]. The well-known PageRank vectors are special Perron vectors of very large Google matrices whose largest eigenvalues are equal to one [14].

From the relation $A = \sigma I - B$, we see that the smallest eigenvalue $\lambda$ of the irreducible nonsingular $M$-matrix $A$ is simple and equal to $\sigma - \rho(B) > 0$, and the positive vector $x$ is the unique associated eigenvector up to scaling. Consequently, if $\sigma$ is available, then the computation of the smallest eigenpair $(\lambda, x)$ of $A$ amounts to that of the largest eigenpair of $B$.

For a general large and sparse $A$, there are a number of general numerical methods for computing a small number of eigenpairs of it. Krylov type methods including the power method applied to $A$ directly are suitable for exterior eigenvalues, i.e., some eigenvalues close to the exterior of the spectrum, and the associated eigenvectors, but they face serious challenges when the desired eigenvalues are inside the interior of the spectrum. Many methods, such as inverse iteration, Rayleigh quotient iteration (RQI) and shift-invert Arnoldi, have been developed to overcome these difficulties; see [24, 21, 30]. However, they require the solution of a possibly ill-conditioned large linear system involving $A$ minus a shift, called inner linear system, at each iteration. This is generally very difficult and even impractical by a direct solver since a factorization of a shifted $A$ may be expensive or prohibited. When inner linear systems are approximately solved by iterative methods, these methods are “inner-outer” iteration methods, also called inexact eigensolvers. The inner iteration means that the inner linear system at each step is approximately solved iteratively, while the outer iteration is the update of the approximate eigenpair(s). Throughout this paper, we always make the underlying hypothesis that a direct solver is not viable for large sparse linear systems and only iterative solvers can be used.

There has been ever growing and intensive interest in inexact eigensolvers over last two decades. Among them, inexact inverse iteration, e.g., [16, 5] and inexact RQI, e.g., [27, 24, 11, 12] are the simplest and the most basic ones. In addition, they are often key ingredients of other sophisticated and practical inexact methods,
such as inverse subspace iteration [22], the Jacobi–Davidson method [28, 30] and the shift-invert residual Arnoldi method [15, 13]. In the mentioned papers and the references therein as well as some others, a number of theoretical results have been established on these methods.

For the computation of the Perron vector of the nonnegative $B$ and the eigenvector of the $M$-matrix $A$ associated with the smallest eigenvalue, a central concern is "how to preserve the strict positivity of approximate eigenvectors". For nonnegative matrices and $M$-matrices, such positivity is crucial in applications since if all the components of an approximate eigenvector do not have the same sign then it is physically meaningless. Unfortunately, both exact and inexact solvers mentioned above are not structure preserving and cannot guarantee the desirable positivity of approximations since it is possible that a converged approximation of $x$ may well have negative components, as is the typical case that the unit length $x$ has very small components. Theoretically, the power method fits into this purpose for nonnegative matrices and naturally preserves the strict positivity of approximate eigenvectors, provided that the starting vector is positive. Due to the equivalence of nonnegative matrix and $M$-matrix eigenproblems, the power method can be adapted to the computation of the smallest eigenpair of an irreducible nonsingular $M$-matrix. However, the power method, though globally convergent, is generally very slow and impractical. Therefore, it is very appealing in both theory and practice to develop both efficient and reliable positivity preserving numerical methods for the nonnegative matrix and $M$-matrix eigenproblems. We will devote ourselves to this topic in this paper.

In 1971, Noda [18] introduced an inverse iteration method with variable shifts for nonnegative matrix eigenvalue problems. This iteration method is a structure preserving method and was motivated by the works of Collatz in 1942 and Wielandt in 1950. For a description and historic overview, we refer the reader to [31, p. 37, 39 and 59] and [8, p. 373] and [4, p. 55]. However, it was Noda who proposed its inverse iteration form with variable shifts different from Rayleigh quotients. We, therefore, call the iteration the Noda iteration (NI). Given a positive starting vector, the NI iteration naturally preserves the strict positivity of approximate eigenvectors at all iterations. It has been adapted to the computation of the smallest eigenpair of an irreducible nonsingular $M$-matrix [35, 3]. The pursue of [35, 3] is to compute the smallest eigenvalue of such $M$-matrix with high relative accuracy. There, it has been shown that NI is practical and effective for the pursue, in which the linear systems involved are solved accurately by a special direct solver, called the GTH like algorithm.

Unlike RQI whose correct convergence to a desired eigenpair is conditional and requires that the initial vector is well selected and reasonably accurate, a major advantage of NI is that, for any given positive initial vector, it converges unconditionally and computes the desired eigenpair correctly. Precisely, the convergence of NI is proven to be globally convergent and asymptotically superlinear [18] and actually quadratic [7]. As it will be clear, NI always generates a monotonically decreasing sequence of approximate eigenvalues whose convergence to $\rho(B)$ is guaranteed; for the smallest eigenvalue $\lambda$ of the irreducible nonsingular $M$-matrix $A$, it always generates a monotonically increasing sequence of approximate eigenvalues that converge to $\lambda$ unconditionally. In other words, the approximate eigenvalues converge to $\rho(B)$ from above or converge to the smallest eigenvalue of the $M$-matrix
from below. In contrast, for symmetric matrices and any initial vector, the Rayleigh quotients always lie inside the spectrum interval. As a result, if RQI converges correctly, the sequence of approximations approach the Perron root of the nonnegative matrix and the smallest eigenvalue of the $M$-matrix from inside.

In this paper, since it is supposed that only iterative solvers are viable to solve the above linear system approximately at each step, based on NI, we first propose an inexact Noda iteration (INI) to find the largest eigenvalue and the associated eigenvector of an irreducible nonnegative matrix $B$. It will be easily adapted to the computation of the smallest eigenvalue of an irreducible nonsingular $M$-matrix $A$. As an inexact eigensolver, our major contribution is to propose two practical inner tolerance strategies for solving the inner linear systems involved inexactly, so that the resulting two INI algorithms are structure preserving and globally converge. The first inner tolerance strategy uses $\gamma \min(\mathbf{x}_k)$ as a stopping criterion for inner iterations, where $\gamma$ is a constant smaller than one and $\mathbf{x}_k$ is the current positive approximate eigenvector. The second inner tolerance strategy solves the inner linear systems with certain decreasing tolerances for inner iterations, which will be described in the context. We establish a rigorous convergence theory of INI with these two inner tolerance strategies, proving that the convergence of the former iteration is globally linear and that of the latter is asymptotically superlinear with the convergence order $\frac{1+\sqrt{5}}{2}$, respectively. In order to derive this superlinear convergence order, we establish a close relationship between the eigenvalue error and the eigenvector error obtained by NI and INI, which is interesting in its own right. We also revisit the convergence of NI and establish a new quadratic convergence result different from that in [7].

Finally, we stress that, different from [35, 3], our aim is the positivity preserving computation of the desired eigenvectors, while their concern is the relative high accuracy computation of the Perron root and the smallest eigenvalue of an $M$-matrix.

The rest of this paper is organized as follows. In Section 2, we introduce the Noda iteration and some preliminaries. In Sections 3 and 4, we propose two INI algorithms for irreducible nonnegative matrix and $M$-matrix eigenvalue problems and establish their convergence theory. In Section 5, numerical results are reported on a few practical problems to illustrate the convergence theory and the effectiveness of INI. We also compare INI with the popular Jacobi–Davidson method and the implicitly restarted Arnoldi method [25], which are not positivity preserving. The experiments indicate that the proposed INI algorithms always preserve the positivity of approximate eigenvectors, while the Jacobi–Davidson method and implicitly restarted Arnoldi method often delivered physically meaningless results. Also, numerically we confirm that the INI algorithms are efficient, competitive with and can be considerably more efficient than the latter two methods. Finally, we summarize the paper with some concluding remarks in Section 6.

2. Preliminaries, Notation and the Noda Iteration

2.1. Preliminaries and Notation. For any matrix $B = [b_{ij}] \in \mathbb{R}^{n \times n}$, we denote $|B| = \|[b_{ij}]\|$. If the entries of $B$ are all nonnegative (positive), then we write $B \geq 0$ ($> 0$). For real matrices $B$ and $C$ of the same size, if $B - C$ is a nonnegative matrix, we write $B \geq C$. A nonnegative (positive) vector is similarly defined. A nonnegative matrix $B$ is said to be reducible if there exists a permutation matrix
where $E$ and $G$ are square matrices; otherwise it is irreducible. Here the superscript $T$ denotes the transpose of a vector or matrix. Throughout the paper, we use a 2-

norm for vectors and matrices, and all matrices are $n \times n$ unless specified otherwise.

We review some fundamental properties of nonnegative matrices and $M$-matrices.

**Theorem 2.1** ([4]). Let $A \in \mathbb{R}^n \times n$ be a nonsingular $M$-matrix. Then the following statements are equivalent:

1. $A = (a_{ij})$, $a_{ij} \leq 0$ for $i \neq j$, and $A^{-1} \geq 0$;  
2. $A = \sigma I - B$ with some $B \geq 0$ and $\sigma > \rho(B)$.

For a pair of vectors $v$ and $w$ with $v \succ 0$, define

$$
\max \left( \frac{w}{v} \right) = \max_i \left( \frac{w^{(i)}}{v^{(i)}} \right), \quad \min \left( \frac{w}{v} \right) = \min_i \left( \frac{w^{(i)}}{v^{(i)}} \right),
$$

where $v = [v^{(1)}, v^{(2)}, \ldots, v^{(n)}]^T$. The following theorem appears in [4, 8, 10, 31] and gives bounds for the spectral radius of a nonnegative matrix $B$.

**Theorem 2.2** ([10, p. 493]). Let $B$ be an irreducible nonnegative matrix. If $v > 0$ is not an eigenvector of $B$, then

$$
(2.1) \quad \min \left( \frac{Bv}{v} \right) < \rho(B) < \max \left( \frac{Bv}{v} \right).
$$

Suppose that $A$ is an irreducible nonsingular $M$-matrix and $(\lambda, x)$ is the smallest eigenpair of it. Then if $v > 0$ is not an eigenvector of $A$, it is easily justified from (2.1) and $A = \sigma I - B$ that

$$
(2.2) \quad \min \left( \frac{Av}{v} \right) \lambda < \max \left( \frac{Av}{v} \right).
$$

For an irreducible nonnegative matrix $B$, recall that the largest eigenvalue $\rho(B)$ of $B$ is simple. Let $x$ be the unit length positive eigenvector corresponding to $\rho(B)$. Then for any orthogonal matrix $[ x \ V ]$ it holds (cf. [8]) that

$$
(2.3) \quad \begin{bmatrix} x^T \\ V^T \end{bmatrix} B \begin{bmatrix} x \\ V \end{bmatrix} = \begin{bmatrix} \rho(B) \\ 0 \end{bmatrix} \begin{bmatrix} c^T \\ L \end{bmatrix}
$$

with $L = V^T B V$ whose eigenvalues constitute the other eigenvalues of $B$. If $\mu$ is not an eigenvalue of $L$, the sep function for $\mu$ and $L$ is defined as

$$
(2.4) \quad \text{sep}(\mu, L) = ||(\mu I - L)^{-1}||^{-1},
$$

which is well defined as $\mu \to \rho(B)$ since $\rho(B)$ is simple. Throughout the paper, we will denote by $\angle(w, z)$ the acute angle of any two nonzero vectors $w$ and $z$.

2.2. The Noda iteration. In [18], Noda presented an inverse iteration with variable shifts for computing the Perron root and vector of an irreducible nonnegative matrix. Given an initial guess $x_0 > 0$ with $||x_0|| = 1$, the Noda iteration (NI)
consists of three steps

\[(\lambda_k I - B) y_{k+1} = x_k, \tag{2.5}\]
\[x_{k+1} = y_{k+1} / \|y_{k+1}\|, \tag{2.6}\]
\[\lambda_{k+1} = \max \left( \frac{B x_{k+1}}{x_{k+1}} \right). \tag{2.7}\]

The main step is to compute a new approximation \(x_{k+1}\) to \(x\) by solving the inner linear system (2.5). Theorem 2.2 shows that \(\lambda_k > \rho(B)\) as long as \(x_k\) is not a scalar multiple of eigenvector \(x\). Furthermore, since \(\lambda_k I - B\) is an irreducible nonsingular \(M\)-matrix, its inverse is irreducible nonnegative. Therefore, \(y_{k+1} > 0\) and \(x_{k+1} > 0\), meaning that the above iteration scheme preserves the strict positivity of approximate eigenvector sequence \(\{x_k\}\). NI is different from RQI, where, in the symmetric case, the Rayleigh quotient of \(B\) with respect to any vector lies in the spectrum interval of \(B\) and thus no more than \(\rho(B)\), and the approximate eigenvectors obtained by RQI do not preserve the strict positivity.

After variable transformation, we get the relation

\[\lambda_{k+1} = \lambda_k - \min \left( \frac{x_k}{y_{k+1}} \right),\]

so \(\lambda_k\) is monotonically decreasing. The algorithm that we will develop here is based on the above inverse iteration with variable shifts. This process is summarized as Algorithm 1.

**Algorithm 1 Noda iteration (NI)**

1. Given an initial guess \(x_0 > 0\) with \(\|x_0\| = 1\) and \(\text{tol} > 0\), compute \(\lambda_0 = \max \left( \frac{B x_0}{x_0} \right)\).
2. for \(k = 0, 1, 2, \ldots\)
3. Solve the linear system \((\lambda_k I - B) y_{k+1} = x_k\).
4. Normalize the vector \(x_{k+1} = y_{k+1} / \|y_{k+1}\|\).
5. Compute \(\lambda_{k+1} = \max \left( \frac{B x_{k+1}}{x_{k+1}} \right)\).
6. until convergence: \(\|B x_{k+1} - \lambda_{k+1} x_{k+1}\| < \text{tol}\).

3. The inexact Noda iteration and convergence theory

Since it is supposed that only iterative solvers are viable to solve the linear system (2.5) at step 3 of Algorithm 1, in this section we shall propose an inexact Noda iteration (INI) for computing the spectral radius \(\rho(B)\) of the irreducible nonnegative matrix \(B\). We provide two different types of practical inner tolerance strategies for solving the inner linear system (2.5) inexact at each step of INI. We will make a detailed and rigorous analysis and prove that the convergence of two different INI is globally linear at least and superlinear with the convergence order \(\frac{1 + \sqrt{5}}{2}^r\), respectively.
3.1. **The inexact Noda iteration.** In INI we compute an approximate solution $y_{k+1}$ in step 3 of Algorithm 1 inexactly such that

$$ (\lambda_k I - B) y_{k+1} = x_k + f_k, \quad x_{k+1} = y_{k+1} / \|y_{k+1}\|, $$

where $f_k$ is the residual vector such that $x_k + f_k$ is positive. Here, the residual norm (inner tolerance) $\xi_k := \|f_k\|$ is allowed to vary at each iteration step $k$.

**Lemma 3.1.** Let $B$ be an irreducible nonnegative matrix and $0 \leq \gamma < 1$ be a fixed constant. For $x_k > 0$, if the residual vector $f_k$ in (3.1) satisfies

$$ |(\lambda_k I - B) y_{k+1} - x_k| = |f_k| \leq \gamma x_k, $$

then the new approximation $x_{k+1} > 0$ and the sequence $\{\lambda_k\}$ with $\lambda_k = \max(Bx_k x_k)$ is monotonically decreasing and bounded below by $\rho(B)$, i.e.,

$$ \lambda_k > \lambda_{k+1} \geq \rho(B). $$

**Proof.** Since $\lambda_k I - B$ is a nonsingular $M$-matrix and $|f_k| \leq \gamma x_k$, the vector $y_{k+1}$ satisfies

$$ y_{k+1} = (\lambda_k I - B)^{-1} (x_k + f_k) > 0. $$

This implies $x_{k+1} = y_{k+1} / \|y_{k+1}\| > 0$ and $\min\left(\frac{x_k + f_k}{y_{k+1}}\right) > 0$.

From (3.1) and the definition of $\lambda_k$ it follows that

$$ \lambda_{k+1} = \max\left(\frac{Bx_{k+1}}{x_{k+1}}\right) = \max\left(\frac{By_{k+1}}{y_{k+1}}\right) = \max\left(\frac{\lambda_k y_{k+1} - x_k - f_k}{y_{k+1}}\right) $$

$$ = \lambda_k - \min\left(\frac{x_k + f_k}{y_{k+1}}\right) < \lambda_k. $$

By Theorem 2.2 we have $\lambda_k > \lambda_{k+1} \geq \rho(B)$. $\square$

Based on (3.1)–(3.2) and Lemma 3.1 we describe INI as Algorithm 2

**Algorithm 2 Inexact Noda Iteration (INI)**

1. Given an initial guess $x_0 > 0$ with $\|x_0\| = 1$, $0 \leq \gamma < 1$ and $\text{tol} > 0$, compute $\lambda_0 = \max\left(\frac{Bx_0}{x_0}\right)$.
2. for $k = 0, 1, 2, \ldots$
3. Solve $(\lambda_k I - B) y_{k+1} = x_k$ approximately by an iterative solver such that

$$ |(\lambda_k I - B)y_{k+1} - x_k| = |f_k| \leq \gamma x_k. $$

4. Normalize the vector $x_{k+1} = y_{k+1} / \|y_{k+1}\|$.
5. Compute $\lambda_{k+1} = \max\left(\frac{Bx_{k+1}}{x_{k+1}}\right)$.
6. until convergence: $\|Bx_{k+1} - \lambda_{k+1} x_{k+1}\| < \text{tol}$.

Note that if $\gamma = 0$, i.e., $f_k = 0$ in (3.1) for all $k$ then Algorithm 2 becomes the standard NI. It follows from Lemma 3.1 that Algorithm 2 generates the positive
vector sequence \( \{x_k\} \), so it is a positivity preserving algorithm. In what follows we will investigate convergence conditions of INI for \( \lambda_k \rightarrow \rho(B) \) as \( k \rightarrow \infty \).

**Lemma 3.2.** Let \( x > 0 \) be the unit length eigenvector of \( B \) associated with \( \rho(B) \). For any vector \( z > 0 \) with \( \|z\| = 1 \), it holds that \( \cos \angle(z, x) > \min(x) \) and

\[
\inf_{\|z\|=1, z>0} \cos \angle(z, x) = \min(x).
\]

**Proof.** Since \( x > 0 \) and \( z > 0 \) with \( \|x\| = \|z\| = 1 \), we have \( \cos \angle(z, x) = z^T x > 0 \). Therefore, the infimum of \( z^T x \) must attain at \( z = e_i \) for some \( i \), where \( e_i \) is the \( i \)th column of the \( n \times n \) identity matrix, that is,

\[
\inf_{\|z\|=1, z>0} \cos \angle(z, x) = \min_i \{\cos \angle(e_i, x)\} = \min_i \{x^{(i)}\} = \min(x).
\]

By \( \|x\| = 1 \), we trivially have an upper bound \( \min(x) \leq n^{-1/2} \), which is attained when all the components of \( x \) are equal. On the other hand, we remark that Lynn and Timlake [17] Theorem 2.1 derived a compact lower bound

\[
\min(x) \geq \frac{\|x\|_1 \min_{i,j} b_{ij}}{\rho(B) - \min_i \sum_j b_{ij} + n \min_{i,j} b_{ij}} > 0
\]

with \( \| \cdot \|_1 \) the vector 1-norm, which is attained and equals the upper bound \( n^{-1/2} \) when \( B \) is a generalized stochastic matrix, i.e., all the row sums of \( B \) are equal to a positive constant. For such \( B \), its row sum is just \( \rho(B) \). So for \( \|x\| = 1 \), the upper bound indicates that \( \min(x) \) is always modestly small for \( n \) large, and the lower bound means that \( \min(x) \) can be very small if \( \min_{i,j} b_{ij} \) is very small.

Let \( \{x_k\} \) be generated by Algorithm 2. We decompose \( x_k \) into the orthogonal direct sum

\[
x_k = x \cos \varphi_k + p_k \sin \varphi_k, \quad p_k \in \text{span}(V) \perp x
\]

with \( \|p_k\| = 1 \) and \( \varphi_k = \angle(x_k, x) \) the acute angle between \( x_k \) and \( x \). Now define

\[
\varepsilon_k = \lambda_k - \rho(B), \quad B_k = \lambda_k I - B.
\]

Then from [23] we have

\[
\begin{bmatrix}
x^T \\
V^T
\end{bmatrix} B_k \begin{bmatrix}
x \\
V
\end{bmatrix} = \begin{bmatrix}
\varepsilon_k \\
0
\end{bmatrix},
\]

where \( L_k = \lambda_k I - L \). For \( \lambda_k \neq \rho(B) \), it is easy to verify that

\[
\begin{bmatrix}
x^T \\
V^T
\end{bmatrix} B_k^{-1} \begin{bmatrix}
x \\
V
\end{bmatrix} = \begin{bmatrix}
\frac{1}{\varepsilon_k} \\
0
\end{bmatrix} b_k^T
\]

with \( b_k^T = -\frac{x^T L_k^{-1}}{\varepsilon_k} \), from which we get

\[
B_k^{-1}V = xb_k^T + VL_k^{-1} = -x^T \frac{L_k^{-1}}{\varepsilon_k} + VL_k^{-1}.
\]

From Lemma 3.1, since \( \{\lambda_k\} \) is monotonically decreasing and bounded by \( \rho(B) \) from below, we must have \( \lim_{k \rightarrow \infty} \lambda_k = \alpha \geq \rho(B) \), i.e., the limit \( \alpha = \rho(B) \) or \( \alpha > \rho(B) \). We next consider these two possibilities, respectively, and present some
basic results which will play an important role in proving the convergence of INI when further restrictions are imposed on inner tolerance $\xi_k = \|f_k\|$.

**Theorem 3.3.** Let $B$ be an irreducible nonnegative matrix, and assume that $(\rho(B), x)$ is the largest eigenpair of $B$ with $x > 0$ and $\|x\| = 1$. If $x_k$, $\lambda_k$, $y_k$ and $f_k$ are generated by Algorithm 2, then the following statements are equivalent:

(i) $\lim_{k \to \infty} x_k = x$;  
(ii) $\lim_{k \to \infty} \lambda_k = \rho(B)$;  
(iii) $\lim_{k \to \infty} \|y_k\|^{-1} = 0$.

**Proof.** (i)$\Rightarrow$(ii): By the definition of $\lambda_k$, we get

$$
\lim_{k \to \infty} \lambda_k = \lim_{k \to \infty} \max \left( \frac{Bx_k}{x_k} \right) = \max \left( \lim_{k \to \infty} \frac{Bx_k}{x_k} \right) = \rho(B).
$$

(i)$\Rightarrow$(iii): Since $|f_k| \leq \gamma x_k$, from (3.3) we have

$$
y_{k+1} = B_k^{-1}(x_k + f_k)
$$

(3.12)

$$
\geq B_k^{-1}(1 - \gamma)x_k
$$

(iii): Since $x_k = x$, we get

$$
y_{k+1} = (1 - \gamma)(\varepsilon_k^{-1}x \cos \varphi_k + B_k^{-1}p_k \sin \varphi_k).
$$

Since $p_k \in \text{span}(V)$, we can write

$$
p_k = Vd_k
$$

with $\|d_k\| = 1$. From (3.11) and (3.12), we get

(3.13)

$$
\varepsilon_k^{-1}x \cos \varphi_k + B_k^{-1}p_k \sin \varphi_k = \varepsilon_k^{-1}(\cos \varphi_k - c^TL_k^{-1}d_k \sin \varphi_k)x + VL_k^{-1}d_k \sin \varphi_k.
$$

From Lemma 3.2 it follows that $\cos \varphi_k > \min(x)$ for all $k$. On the other hand, (i) means that $\cos \varphi_k \to 1$ and $\sin \varphi_k \to 0$ as $k \to \infty$. Since $\lim_{k \to \infty} x_k = x$, it holds that $\lim_{k \to \infty} \lambda_k = \rho(B)$. As a result, from definition (2.4) and $\|V\| = 1$, as $k \to \infty$, the second term in the right-hand side of (3.13) is bounded by

$$
\|VL_k^{-1}d_k \sin \varphi_k\| \leq \|V\|\|L_k^{-1}\|\|d_k\| \sin \varphi_k = \sin \varphi_k \frac{\sin \varphi_k}{\text{sep}(\lambda_k, L)} \to 0,
$$

and

$$
|c^TL_k^{-1}d_k \sin \varphi_k| \leq \|c\| \|L_k^{-1}\| \sin \varphi_k \to 0.
$$

Combining (3.12) with the above and exploiting the triangle inequality of norms, we have

$$
\|y_{k+1}\| \geq (1 - \gamma) \left( (\lambda_k - \rho(B))^{-1} |\cos \varphi_k - c^TL_k^{-1}d_k \sin \varphi_k| - \|VL_k^{-1}d_k \sin \varphi_k\| \right)
$$

$$
\geq (1 - \gamma) \left( \frac{|\cos \varphi_k - c^TL_k^{-1}d_k | \sin \varphi_k|}{|\lambda_k - \rho(B)|} - \frac{\sin \varphi_k}{\text{sep}(\lambda_k, L)} \right)
$$

$$
\geq (1 - \gamma) \left( \frac{|\min(x) - c^TL_k^{-1}d_k | | \sin \varphi_k|}{|\lambda_k - \rho(B)|} - \frac{\sin \varphi_k}{\text{sep}(\lambda_k, L)} \right) \to \infty.
$$

(iii)$\Rightarrow$(ii): From $|f_k| \leq \gamma x$ with $0 \leq \gamma < 1$ and

$$
y_{k+1} = (\lambda_k I - B)^{-1}(x_k + f_k),
$$
we get 
\[ \|y_{k+1}\| \leq \| (\lambda_k I - B)^{-1} (\|x_k\| + \|f_k\|) \| < 2 \| (\lambda_k I - B)^{-1} \|, \]
meaning that 
\[ \frac{1}{\|y_{k+1}\|} > \frac{1}{2 \| (\lambda_k I - B)^{-1} \|} \geq 0. \]
Since \( \|y_k\| \to 0 \) as \( k \to \infty \), the sequence \( \{\lambda_k I - B\} \) tends to a singular matrix, meaning that \( \{\lambda_k\} \) converges to an eigenvalue of \( B \). From Lemma 3.1 we must have \( \lim_{k \to \infty} \lambda_k = \rho(B) \) because \( \{\lambda_k\} \) is monotonically decreasing with the lower bound \( \rho(B) \). Otherwise, if \( \lim_{k \to \infty} \lambda_k = \alpha > \rho(B) \), then \( \alpha \) were an eigenvalue of \( B \) that is larger than the spectral radius \( \rho(B) \) of \( B \), which is impossible. □

**Theorem 3.4.** For Algorithm 2 if \( \lim_{k \to \infty} \lambda_k = \alpha > \rho(B) \), then (i) \( \|y_k\| \) is bounded; (ii) \( \lim_{k \to \infty} \min(x_k + f_k) = 0 \); (iii) \( \sin \angle (x, x_k) \geq \zeta > 0 \) with \( \zeta \) some constant.

**Proof.** (i) Since \( \|f_k\| \leq \gamma x_k \), we get
\[ \|y_{k+1}\| = \| (\lambda_k I - B)^{-1} (x_k + f_k) \| < 2 \| (\lambda_k I - B)^{-1} \| = \frac{2}{\text{sep}(\lambda_k, B)} \to \frac{2}{\text{sep}(\alpha, B)} < \infty. \]

(ii) From (3.5) it follows that
\[ \lim_{k \to \infty} \min \left( \frac{x_k + f_k}{y_{k+1}} \right) = \lim_{k \to \infty} \frac{(\lambda_k - \lambda_{k+1})}{\lambda_k} = 0. \]
On the other hand, from (3.14) and (3.15) we have
\[ \min \left( \frac{x_k + f_k}{y_{k+1}} \right) \geq \min(x_k + f_k) \geq \frac{\min(x_k + f_k)}{\|y_{k+1}\|} > \frac{\min(x_k + f_k) \text{sep}(\lambda_k, B)}{2} \]
Thus, it is holds that
\[ \lim_{k \to \infty} \min(x_k + f_k) = 0. \]

(iii) Suppose there is a subsequence \( \{\sin \angle (x, x_k)\} \) that converges to zero. Then from Theorem 3.3 there is a subsequence \( \{\lambda_k\} \) that converges to \( \rho(B) \). This is a contradiction. □

Keep in mind that in Theorems 3.3, 3.4 we only assume the condition \( \|f_k\| \leq \gamma x_k \) with \( 0 < \gamma < 1 \), under which the sequence \( \{\lambda_k\} \) must converge to either \( \rho(B) \) or \( \alpha > \rho(B) \). So only this condition is not enough to guarantee that INI computes the desired eigenpair \( (\rho(B), x) \) of \( B \), and we have to impose some stronger conditions on \( f_k \) to ensure \( \lim_{k \to \infty} \lambda_k = \rho(B) \).

3.2. Convergence Analysis. In order to make INI converge correctly and practically, we now propose the following two practical inner tolerance strategies for the inexactness of step 3 of Algorithm 2.

- **INI1**: the residual norm satisfies \( \xi_k = \|f_k\| \leq \gamma \min(x_k) \) for some constant \( 0 < \gamma < 1 \);
- **INI2**: the residual vector satisfies \( \|f_k\| \leq d_k x_k \) with \( d_k = 1 - \lambda_k/\lambda_{k-1} \) for \( k \geq 1 \) and \( \|f_0\| \leq \gamma \min(x_0) \) for some constant \( 0 < \gamma < 1 \).
It is easily seen that the residual vectors $f_k$ of INI$_1$ and INI$_2$ must satisfy $|f_k| \leq \gamma x_k$ with $0 < \gamma < 1$. As a result, it is known from Lemma 3.1 that each of INI$_1$ and INI$_2$ generates a monotonically decreasing sequence \{x\} bounded by $\rho(B)$ and a sequence of positive vectors \{x\}. INI$_1$ and INI$_2$ now require stronger conditions on $f_k$ than the previous INI, and we shall prove that, for any given initial guess $x_0 > 0$, INI$_1$ and INI$_2$ guarantee the global convergence \{x\} → $\rho(B)$ and \{x\} → $x$.

We decompose $x_k+1$ in the same manner as (3.3):

\begin{equation}
\tag{3.16}
x_{k+1} = x \cos \phi_{k+1} + p_{k+1} \sin \phi_{k+1}, \quad p_{k+1} \in \text{span}(V) \perp x
\end{equation}

with $\|p_{k+1}\| = 1$. So by definition, $\cos \phi_{k+1} = x^T x_{k+1}$ and $\sin \phi_{k+1} = \|V^T x_{k+1}\|$. Obviously, $x_k \to x$ if and only if $\tan \phi_k \to 0$, i.e., $\sin \phi_k \to 0$.

From (3.10) we have

$\|L_k^{-1}V \| \leq \| \sin \phi_k + \|f_k\| \| \theta \| \|f_k\| \| \cos \phi_k \|$, 

with the last second inequality holding by assuming that $1 + \|\|L_k^{-1}\||\|f_k\|/\|\cos \phi_k\| < 1 - \|\cos \phi_k\|$. This assumption must be satisfied provided that $\|f_k\|$ is suitably small, because, by Lemma 3.2, $\cos \phi_k > \text{min}(x)$ for all $k$.

Particularly, if $f_k = 0$, i.e., $\gamma = 0$, we get

\begin{equation}
\tag{3.17}
\tan \phi_{k+1} \leq \frac{\|L_k^{-1}\| \|f_k\|}{1 - \|\|L_k^{-1}\||\|f_k\|/\|\cos \phi_k\|} \tan \phi_k := \beta_k \tan \phi_k
\end{equation}

for NI. Since NI is quadratically convergent [7], for $k$ large enough we must have

$\beta_k = O(\tan \phi_k) \to 0$.

Therefore, for any given positive constant $\beta < 1$, we have

\begin{equation}
\tag{3.18}
\beta_k < \beta < 1
\end{equation}

for $k \geq N$ with $N$ large enough, meaning that

$\tan \phi_{k+1} < \beta \tan \phi_k$ for $k \geq N$. We remark that $\beta_k = \|L_k^{-1}\| \|f_k\|/\|B\|$ for $B$ a normal matrix since $c = 0$. 

Theorem 3.5. Let $B$ be an irreducible nonnegative matrix. If the sequence \( \{ \lambda_k \} \) is generated by INL$_1$, then \( \{ \lambda_k \} \) is monotonically decreasing and \( \lim_{k \to \infty} \lambda_k = \rho(B) \).

Proof. By the assumption on INL$_1$, since $\xi_k = \| f_k \| \leq \gamma \min(x_k)$, it holds that $|f_k| \leq \gamma x_k$ with $0 < \gamma < 1$, which satisfies the condition in Lemma 3.1. So the sequence \( \{ \lambda_k \} \) is bounded and monotonically decreasing, and we must have either \( \lim_{k \to \infty} \lambda_k = \rho(B) \) or \( \lim_{k \to \infty} \lambda_k = \alpha > \rho(B) \), as already clarified before Theorem 3.4. Next we prove that, for INL$_1$, \( \lim_{k \to \infty} \lambda_k = \rho(B) \) and \( \lim_{k \to \infty} \lambda_k = \alpha > \rho(B) \) is impossible.

Assume that \( \lim_{k \to \infty} \lambda_k = \alpha > \rho(B) \). By \( |f_k| \leq \gamma x_k \), we get
\[
\min(x_k + f_k) \geq (1 - \gamma) \min(x_k).
\]

It follows from (ii) of Theorem 3.4 that
\[
0 = \lim_{k \to \infty} \min(x_k + f_k) \geq (1 - \gamma) \lim_{k \to \infty} \min(x_k) \geq 0.
\]

Thus, we have
\[
\lim_{k \to \infty} \min(x_k) = 0.
\]

From Lemma 3.2 and (iii) of Theorem 3.4 we know that \( \sin \varphi_k \) and \( \cos \varphi_k \) are uniformly bounded below by a positive constant. Therefore, there is an $m > 0$ such that $m \leq \sin \varphi_k$ and $m \leq \cos \varphi_k$, leading to
\[
1 = \frac{\tan \varphi_k}{m} \quad \text{and} \quad \min(x_k) \leq \frac{\sin \varphi_k}{m} \min(x_k).
\]

Using (3.17) and (3.21), we obtain
\[
\tan \varphi_{k+1} \leq \left| L_k \right|^{-1} \varepsilon_k \frac{\tan \varphi_k + \gamma \min(x_k) \tan \varphi_k / m}{1 - c^T L_k^{-1} V^T x_k / m - \gamma (1 + \| c \| \left| L_k^{-1} \right|) \min(x_k) / m}
\]
\[
\leq \left| L_k \right|^{-1} \varepsilon_k \frac{\tan \varphi_k + \gamma \min(x_k) \tan \varphi_k / m}{1 - c^T L_k^{-1} V^T x_k / m - \gamma (1 + \| c \| \left| L_k^{-1} \right|) \min(x_k) / m} \tan \varphi_k.
\]

Define
\[
\beta_k' = \left| L_k \right|^{-1} \varepsilon_k \frac{1 + \gamma \min(x_k) / m}{1 - c^T L_k^{-1} V^T x_k / m - \gamma (1 + \| c \| \left| L_k^{-1} \right|) \min(x_k) / m}.
\]

Note that \( \left| L_k \right|^{-1} \to \left| (\alpha I - L) \right|^{-1} \) is uniformly bounded. Obviously, $\beta_k'$ is a continuous function with respect to $\min(x_k)$ for $0 < \gamma < 1$, and thus $\beta_k' \to \beta_k$ as $\min(x_k) \to 0$. Particularly, for $0 < \gamma < 1$ and any small positive number $\delta$, it holds that $\beta_k' \leq (1 + \delta) \beta_k$ provided that $\min(x_k)$ is suitably small. Therefore, from (3.19), for $k \geq N$ large enough we can choose a sufficiently small $\delta$ such that
\[
\beta_k' \leq (1 + \delta) \beta_k \leq \beta < 1
\]
for $\min(x_k)$ sufficiently small. As a result, for $k \geq N$ large enough and $\min(x_k)$ sufficiently small, we have
\[
\tan \varphi_{k+1} \leq \beta \tan \varphi_k.
\]

It then follows from this that $\tan \varphi_k \to 0$, that is, $x_k \to x$. From Lemma 3.1 this means that \( \{ \lambda_k \} \) converges to $\rho(B)$ monotonically, a contradiction to the assumption that $\lim_{k \to \infty} \lambda_k = \alpha > \rho(B)$. \qed
Theorem 3.6. Let $B$ be an irreducible nonnegative matrix. If $\overline{\lambda}_k$ is generated by INL$_2$, then $\lim_{k \to \infty} \overline{\lambda}_k = \rho(B)$.

Proof. Since $|f_k| \leq d_k x_k$ with $d_k = (\overline{\lambda}_{k-1} - \overline{\lambda}_k) / \overline{\lambda}_{k-1} < 1$ and $\|f_0\| \leq \gamma \min(x_0)$ with $0 < \gamma < 1$, $f_k$ satisfies the condition of Lemma 3.1. Therefore, the sequence $\{\overline{\lambda}_k\}$ generated by INL$_2$ is monotonically decreasing with the lower bound $\rho(B)$ and converges to some limit $\alpha \geq \rho(B)$. This shows that $\xi_k = \|f_k\| \to 0$ as $k \to \infty$. Assume that $\lim_{k \to \infty} \overline{\lambda}_k = \alpha > \rho(B)$. Therefore, there exists a positive integer $N_1$ such that $\xi_k \leq \gamma < 1$ with $\gamma$ defined in INL$_1$. For $\delta$ and $\beta'_k$ as well as $N$ in the proof of Theorem 3.5 take $N_2 = \max\{N_1, N\}$. Then for $k \geq N_2$ we get
\[
\tan \varphi_{k+1} \leq \beta \tan \varphi_k.
\]
Hence, it holds that $\lim_{k \to \infty} \overline{\lambda}_k = \rho(B)$.

For INL$_2$, we further have the following result.

Theorem 3.7. For INL$_2$, it holds that
\[
\lim_{k \to \infty} \varepsilon_k y_{k+1} = x,
\]
where $\varepsilon_k$ is defined by (3.10).

Proof. From (3.10), we have
\[
B_k^{-1} = \frac{1}{\varepsilon_k} xx^T - \frac{xc^T L_k^{-1} V^T}{\varepsilon_k} + VL_k^{-1} V^T.
\]

Therefore, we get
\[
\varepsilon_k y_{k+1} = \varepsilon_k B_k^{-1} (x_k + f_k) = (xx^T - xc^T L_k^{-1} V^T + \varepsilon_k VL_k^{-1} V^T) (x_k + f_k).
\]

Since $L_k^{-1} \to (\rho(B)I - L)^{-1}$ and $\varepsilon_k \to 0$, we have $\varepsilon_k VL_k^{-1} V^T \to 0$, from which it follows that
\[
\lim_{k \to \infty} \varepsilon_k \|VL_k^{-1} V^T (x_k + f_k)\| = 0.
\]

From Theorem 3.3 and the proof of Theorem 3.6 we know that $x_k \to x$ and $f_k \to 0$, which lead to $\lim_{k \to \infty} (x_k + f_k) = x$. Note that $V^T x = 0$. We then get
\[
\lim_{k \to \infty} xc^T L_k^{-1} V^T (x_k + f_k) = xc^T (\rho(B)I - L)^{-1} V^T x = 0.
\]

A combination of the above relations shows that
\[
\lim_{k \to \infty} \varepsilon_k y_{k+1} = \lim_{k \to \infty} (xx^T - xc^T L_k^{-1} V^T + \varepsilon_k VL_k^{-1} V^T) (x_k + f_k) = xx^T x = x.
\]

3.3. Convergence Rates. Theorems 3.3–3.6 have proved the global convergence of INL$_1$ and INL$_2$, but the results are only qualitative and do not tell us anything on how fast the INI methods converge. In this subsection, we will show that the convergence of INL$_1$ and INL$_2$ is globally linear at least and superlinear, respectively. We will precisely derive the asymptotic linear convergence factor of INL$_1$ and the superlinear convergence order of INL$_2$. 
From (3.23) and (3.34), we have
\[(3.24) \quad \varepsilon_{k+1} = \varepsilon_k \left(1 - \min \left(\frac{x_k + f_k}{\varepsilon_k y_{k+1}}\right)\right) = \varepsilon_k \rho_k.\]

Since \(\overline{\lambda}_k - \overline{\lambda}_{k+1} < \overline{\lambda}_k - \rho(B)\), from (3.24), (3.34) and (3.31) we always have
\[(3.25) \quad \rho_k = 1 - \min \left(\frac{x_k + f_k}{\varepsilon_k y_{k+1}}\right) = 1 - \frac{\overline{\lambda}_k - \overline{\lambda}_{k+1}}{\overline{\lambda}_k - \rho(B)} < 1.\]

**Theorem 3.8.** For INL_1, it holds that \(\lim_{k \to \infty} \rho_k \leq \frac{2\gamma}{1+\gamma} < 1\), i.e., the convergence of INL_1 is globally linear at least.

*Proof.* Since \(\xi_k = \|f_k\| \leq \gamma \min(x_k)\) in INL_1, it holds that \(|f_k| \leq \gamma x_k\). Therefore, we have
\[(1 - \gamma) x_k \leq x_k + f_k \leq (1 + \gamma) x_k.\]
As \(B_k^{-1} \geq 0\), it follows from the above relation that
\[(1 - \gamma) B_k^{-1} x_k \leq y_{k+1} \leq (1 + \gamma) B_k^{-1} x_k.\]
Therefore, we have
\[(3.26) \quad \min \left(\frac{x_k + f_k}{\varepsilon_k y_{k+1}}\right) \geq \min \left(\frac{(1 - \gamma) x_k}{(1 + \gamma) \varepsilon_k B_k^{-1} x_k}\right) = \frac{1 - \gamma}{1 + \gamma} \min \left(\frac{x_k}{\varepsilon_k B_k^{-1} x_k}\right).\]
From (3.22), we get
\[\varepsilon_k B_k^{-1} x_k = x x^T x_k - x c^T L_k^{-1} V^T x_k + \varepsilon_k V L_k^{-1} V^T x_k.\]
Since, from Theorem 3.5 \(\lim_{k \to \infty} x_k = x\) and \(\lim_{k \to \infty} \overline{\lambda}_k = \rho(B)\), we have \(\varepsilon_k \to 0\) and \(L_k^{-1} \to (\rho(B) I - L)^{-1}\). On the other hand, note that \(\lim_{k \to \infty} V^T x_k = V^T x = 0\). Then it is easily verified that
\[\lim_{k \to \infty} \varepsilon_k B_k^{-1} x_k = x.\]
Consequently, we obtain
\[\lim_{k \to \infty} \min \left(\frac{x_k + f_k}{\varepsilon_k y_{k+1}}\right) \geq \frac{1 - \gamma}{1 + \gamma} \min \left(\frac{x_k}{\varepsilon_k B_k^{-1} x_k}\right) = \frac{1 - \gamma}{1 + \gamma} > 0.\]
Hence
\[(3.27) \quad \lim_{k \to \infty} \rho_k \leq 1 - \frac{1 - \gamma}{1 + \gamma} = \frac{2\gamma}{1 + \gamma} < 1.\]

\(\square\)

It is seen from (3.27) that if \(\gamma\) is small, then INL_1 must ultimately converge fast.

**Theorem 3.9.** Define the residual \(r_k = (\overline{\lambda}_k I - B)x_k\). Then for INL_2, the following results hold:

(i) \(\lim_{k \to \infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = 0\); (ii) \(\lim_{k \to \infty} \frac{\overline{\lambda}_k - \overline{\lambda}_{k+1}}{\overline{\lambda}_k - 1} = 0\); (iii) \(\lim_{k \to \infty} \frac{\|r_{k+1}\|}{\|r_k\|} = 0\),

that is, the convergence of INL_1 is superlinear.
Proof. (i): Theorem 3.6 has proved that \( \lim_{k \to \infty} \lambda_k = \rho(B) \) and \( \lim_{k \to \infty} x_k = x \). Since, from the proof of Theorem 3.6, \( \xi_k = \|f_k\| \to 0 \), it follows from Theorem 3.7, (3.24) and Theorem 3.6 that

\[
\lim_{k \to \infty} \rho_k = \lim_{k \to \infty} \frac{\xi_{k+1}}{\xi_k} = 1 - \lim_{k \to \infty} \min_{k \to \infty} \left( \frac{x_k + f_k}{\xi_k y_{k+1}} \right) = 1 - \min \left( \lim_{k \to \infty} \frac{x_k + f_k}{\xi_k y_{k+1}} \right) = 1 - \min \left( \lim_{k \to \infty} \frac{x}{y} \right) = 0.
\]

(3.28)

(ii): From (3.24) and (3.28), we have

\[
\lim_{k \to \infty} \lambda_k - \rho_k = \lim_{k \to \infty} \frac{\xi_k - \xi_{k+1}}{\xi_k} = \lim_{k \to \infty} \frac{\xi_k (1 - \rho_k)}{\xi_{k-1} (1 - \rho_{k-1})} = \lim_{k \to \infty} \frac{(1 - \rho_k) \rho_{k-1}}{(1 - \rho_{k-1})} = 0.
\]

(iii): From (3.31), we have

\[
\|r_k\| = \left\| \left( \lambda_k I - B \right) x_k \right\| = \left\| \left( \lambda_k I - B \right) y_k \right\| = \frac{\xi_{k-1} \|x_k + f_k\|}{\xi_{k-1} \|y_k\|} = \frac{\xi_{k-1} \|x_k + f_k\|}{\xi_{k-1} \|y_k\|} = \frac{\xi_{k-1} \|x_k + f_k\|}{\xi_{k-1} \|y_k\|}.
\]

Then it follows from Theorem 3.7, (3.24) and (3.28) that

\[
\lim_{k \to \infty} \frac{\|r_{k+1}\|}{\|r_k\|} = \lim_{k \to \infty} \frac{\xi_k \|x_k + f_k\| \|y_{k+1}\|}{\xi_k \|x_{k-1} + f_{k-1}\| \|y_k\|} = \lim_{k \to \infty} \frac{\rho_{k-1} \|x\| \|x\|}{\|x\| \|x\|} = 0.
\]

\[\square\]

Although Theorem 3.9 has established the superlinear convergence of INL_2, it does not reveal the convergence order. Our next concern is to derive the precise convergence order of INL_2. This is more informative and instructive to understand how fast INL_2 converges.

Elsner [7] proved the (asymptotic) quadratic convergence of the sequence \( \{\lambda_k - \rho(B)\} \), but the constant factor (multiplier) in his quadratic convergence result appears hard to quantify or estimate. Below we establish an intimate and quantitative relationship between the eigenvalue error \( \varepsilon_k = \lambda_k - \rho(B) \) and the eigenvector error \( \tan \varphi_k \), which holds for both NI and INI. This result plays a crucial role in deriving the precise convergence order of INL_2 and proving the quadratic convergence of NI in terms of \( \tan \varphi_k \), with the constant factor in the quadratic convergence result given explicitly.

**Theorem 3.10.** For NI, INI_1 and INI_2, we have

\[
\varepsilon_k \leq \frac{2\|B\|}{\min(x)} \tan \varphi_k + O(\tan^2 \varphi_k)
\]

for \( k \) large enough. For NI we have

\[
\tan \varphi_{k+1} \leq \frac{2\|B\|}{\min(x) \text{sep}(\lambda_k, L)} \tan^2 \varphi_k + O(\tan^3 \varphi_k),
\]

i.e., NI is quadratically convergent.
Proof. Since NI, INI\(_1\) and INI\(_2\) converge, for \(k\) large enough we must have
\[ |p_k| \tan \varphi_k \ll x, \text{ i.e., } |p_k| \sin \varphi_k \ll x \cos \varphi_k. \]
Therefore, from (3.31), the nonnegativity of \(B\) and \(\|p_k\| = 1\), for \(k\) large enough we obtain
\[
\lambda_k = \max \left( \frac{Bx_k}{x_k} \right) = \max \left( \frac{B(x \cos \varphi_k + p_k \sin \varphi_k)}{x \cos \varphi_k + p_k \sin \varphi_k} \right)
\leq \max \left( \frac{\rho(B)x \cos \varphi_k + Bp_k \sin \varphi_k}{x \cos \varphi_k - |p_k| \sin \varphi_k} \right)
\leq \max \left( \frac{\rho(B)x + Bp_k \tan \varphi_k}{x - |p_k| \tan \varphi_k} \right)
\leq \max \left( \frac{\rho(B)(x - |p_k| \tan \varphi_k) + \rho(B)|p_k| \tan \varphi_k + B|p_k| \tan \varphi_k}{x - |p_k| \tan \varphi_k} \right)
\leq \frac{\rho(B) + \|B\|}{\min(x)} \tan \varphi_k + O(\tan^2 \varphi_k)
\leq \rho(B) + \frac{2\|B\|}{\min(x)} \tan \varphi_k + O(\tan^2 \varphi_k).
\]
Therefore, we get
\[
\varepsilon_k = \lambda_k - \rho(B) \leq \frac{2\|B\|}{\min(x)} \tan \varphi_k + O(\tan^2 \varphi_k).
\]

Since \(\|V^T x_k\| = \sin \varphi_k\), we have
\[
|c^T L_k^{-1} V^T x_k / \cos \varphi_k| \leq \|c\| \|L_k^{-1}\| \tan \varphi_k \to 0
\]
as \(k\) increases. Note that \(\text{sep}(\lambda_k, L) = \frac{1}{\|L_k^{-1}\|}\). Therefore, from (3.13), for \(k\) large enough we obtain
\[
\tan \varphi_{k+1} \leq \frac{\varepsilon_k}{\text{sep}(\lambda_k, L)} \frac{1}{1 - \|c\| \|L_k^{-1}\| \tan \varphi_k} \tan \varphi_k
\]
\[
= \frac{\varepsilon_k}{\text{sep}(\lambda_k, L)} (1 + \|c\| \|L_k^{-1}\| \tan \varphi_k + O(\tan^2 \varphi_k)) \tan \varphi_k,
\]
from which and (3.29) it follows that (3.30) holds.

Note that \(\text{sep}(\lambda_k, L) \to \text{sep}(\rho(B), L)\). The above relation proves the (asymptotic) quadratic convergence of NI. \(\square\)

**Theorem 3.11.** For \(k\) large enough, the inner tolerance \(\xi_k\) in INI\(_2\) satisfies
\[
(3.31) \quad \xi_k = \|f_k\| = O(\tan \varphi_{k-1}),
\]
and INI\(_2\) converges superlinearly in the form
\[
(3.32) \quad \tan \varphi_{k+1} \leq C \tan^\alpha \varphi_k
\]
with the convergence order \( \alpha = \frac{1 + \sqrt{5}}{2} \approx 1.618 \) and \( C \) a constant.

Proof. By the condition of INI\(^2\) and \( \lambda_{k-1} > \lambda_k \geq \rho(B) \), we have

\[
\xi_k = \|f_k\| \leq d_k = \frac{\lambda_{k-1} - \lambda_k}{\lambda_{k-1}} \leq \frac{\lambda_{k-1} - \rho(B)}{\lambda_{k-1}} \leq \frac{\lambda_{k-1} - \rho(B)}{\rho(B)}.
\]

Note that (3.29) holds for INI. Therefore, for \( k \) large enough we have

\[
\|f_k\| = O(\tan \varphi_{k-1}),
\]

which is just (3.31).

It is known from Lemma [3.2] that \( \cos \varphi_k > \min(\lambda) \) for all \( k \). For \( k \) large enough, note that \( \operatorname{sep}(\lambda_k, L) \to \operatorname{sep}(\rho(B), L) \), and make the Taylor expansion of the reciprocal of the denominator in (3.17). Then substituting (3.29) into (3.17) and amplifying the term

\[
|c^T L_k^{-1} v^T x_k / \cos \varphi_k| \leq \|c\| L_k^{-1} \| \tan \varphi_k,
\]

by some elementary manipulation we get

\[
\tan \varphi_{k+1} \leq C_1 \tan^2 \varphi_k + C_2 \tan \varphi_k \tan \varphi_{k-1}
\]

with \( C_1 \) and \( C_2 \) certain positive constants. Since \( \tan \varphi_k < \tan \varphi_{k-1} \) for \( k \) large enough, the above inequality can be written as

\[
\tan \varphi_{k+1} \leq C \tan \varphi_k \tan \varphi_{k-1}
\]

with \( C \) a positive constant. Taking the equality sign in the above relation, by the theory of linear difference equation [33, p. 436-7], for \( k \) sufficiently large we obtain

\[
\tan \varphi_{k+1} \leq C\alpha^{-1} \tan^\alpha \varphi_k := C \tan^\alpha \varphi_k
\]

with \( \alpha = \frac{1 + \sqrt{5}}{2} \).

We comment that if \( f_k = 0 \) then \( C_2 = 0 \) in the above proof, in which case INI becomes NI and (3.30), the quadratic convergence of NI, is thus recovered. (3.31) indicates that, at outer iteration step \( k \), the inner tolerance \( \|f_k\| \) in INI\(^2\) decreases like \( \tan \varphi_{k-1} \), so we need to solve the inner linear systems more and more accurately as iterations proceed.

4. Computing the Smallest Eigenpair of an \( M \)-Matrix

In this section, we consider how to compute the smallest eigenpair of an irreducible nonsingular \( M \)-matrix \( A \). For the derivation of INI, suppose that \( A \) is expressed in the form \( A = \sigma I - B \), and let \( (\lambda, x) \) be the smallest eigenpair of it. As have been proved previously, each of INI\(^1\) and INI\(^2\) generates a monotonically decreasing sequence \( \{\lambda_k\} \) that converges to \( \rho(B) \) with \( \lambda_k > \rho(B) \). We denote \( \lambda_k = \sigma - \lambda_k \). It follows that \( \lambda_k < \lambda \) and \( \{\lambda_k\} \) forms a monotonically increasing sequence that converges to \( \lambda \). From Algorithm [2] at iteration \( k \) the approximate solution \( y_{k+1} \) exactly solves the linear system

\[
(\lambda_k I - B) y_{k+1} = x_k + f_k,
\]

where \( \lambda_k = \max \left( \frac{\langle y_k, x_k \rangle}{\|x_k\|^2} \right) \). We then update the next iterate as \( x_{k+1} = y_{k+1} / \|y_{k+1}\| \). Since

\[
\lambda_k I - B = (\sigma I - B) + (\lambda_k - \sigma) I = A - \lambda_k I,
\]

we have

\[
(\lambda_k I - B) y_{k+1} = x_k + f_k
\]

and

\[
x_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|}
\]

for all \( k \).
(4.1) is equivalent to

\[(A - \lambda_k I) y_{k+1} = x_k + f_k,\]

where

\[\lambda_k = \sigma - \max \left( \frac{B x_k}{x_k} \right) = \min \left( \frac{A x_k}{x_k} \right).\]

Since \((A - \lambda_k I)\) is an irreducible nonsingular \(M\)-matrix, it holds that \(y_{k+1} > 0\) provided that \(x_k + f_k > 0\). Thus, we get the relation

\[\lambda_{k+1} = \min \left( \frac{A x_{k+1}}{x_{k+1}} \right) = \lambda_k + \min \left( \frac{x_k + f_k}{y_{k+1}} \right).\]

Therefore, Algorithm 2 can be modified to compute the smallest eigenpair of the irreducible nonsingular \(M\)-matrix \(A\), which is described as Algorithm 3, where, as in Section 3.2, we define

- **INI Type1**: the residual norm satisfies \(\xi_k \leq \gamma \min(x_k)\) for some \(0 < \gamma < 1\).
- **INI Type2**: the residual vector satisfies \(|f_k| \leq d_k x_k\) with \(d_k = 1 - \frac{\lambda_{k-1}}{\lambda_k}\) for \(k \geq 1\) and \(\|f_0\| \leq \gamma \min(x_0)\) with some \(0 < \gamma < 1\).

We should point out that Algorithm 3 itself neither involves \(\sigma\) nor requires that \(A\) be expressed in the form \(A = \sigma I - B\), which is purely for the algorithmic derivation.

**Algorithm 3** INI for \(M\)-matrices

1. Given an initial guess \(x_0 > 0\) with \(\|x_0\| = 1\) and \(\text{tol} > 0\), compute \(\lambda_0 = \min \left( \frac{A x_0}{x_0} \right)\).
2. for \(k = 0, 1, 2, \ldots\)
3. \(\text{Solve } (A - \lambda_k I) y_{k+1} = x_k \) inexactly with Type1 or Type2.
4. \(\text{Normalize the vector } x_{k+1} = y_{k+1}/\|y_{k+1}\|\).
5. \(\text{Compute } \lambda_{k+1} = \min \left( \frac{A x_{k+1}}{x_{k+1}} \right)\).
6. until convergence: \(\|A x_{k+1} - \lambda_{k+1} x_{k+1}\| < \text{tol}\).

Due to the equivalence of Algorithm 3 and Algorithm 2, the previous convergence results for the irreducible nonnegative matrix eigenvalue problem naturally hold for the irreducible nonsingular \(M\)-matrix eigenvalue problem. We summarize the main results as follows.

**Theorem 4.1.** Let \(A\) be an irreducible nonsingular \(M\)-matrix. If \(\lambda_k\) and \(x_k\) are generated by Algorithm 3, then \(\left\{ \lambda_k \right\} \to \lambda\), the smallest eigenvalue of \(A\), monotonically from below as \(k \to \infty\), and \(\lim_{k \to \infty} x_k = x\) with \(x_k > 0\) for all \(k > 0\). Furthermore, the convergence of INI Type1 and INI Type2 is globally linear with the asymptotic linear convergence factor \(\frac{2\gamma}{1+\gamma}\) and superlinear with the convergence order \(\frac{1+\sqrt{5}}{2}\), respectively.

5. Numerical experiments

In this section we present numerical experiments to support our theoretical results on INI and NI, and illustrate the effectiveness of the proposed INI algorithms, compared to NI. In the meantime, we compare INI with the positivity non-preserving algorithms JDQR [29], JDRPCG [19] and the implicitly restarted
Arnoldi method [25], i.e., the Matlab function eigs. We use JDQR for the nonnegative matrix and $M$-matrix eigenvalue problems but use JDRPCG only for symmetric nonsingular $M$-matrices since JDRPCG is designed to compute a few number of smallest eigenpairs of a symmetric matrix. We show that the NI and INI algorithms are always reliable to compute positive eigenvectors while the other algorithms often fail to do so. We also demonstrate that the INI algorithms are efficient, and they are competitive with and can be considerably efficient than the other algorithms. All numerical tests were performed on an Intel (R) Core (TM) i5 CPU 750@ 2.67GHz with 4 GB memory using Matlab 7.11.0 with the machine precision $\epsilon = 2.22 \times 10^{-16}$ under the Microsoft Windows 7 64-bit.

5.1. INI for Nonnegative Matrices. We present two examples to illustrate numerical behavior of NI, INI$_1$ and INI$_2$ for nonnegative matrices. At each outer iteration step we solve the linear system (3.1), i.e.,

$$\begin{align*}
(\lambda_k I - B) y_{k+1} &= x_k + f_k,
\end{align*}$$

using the following stopping criteria:

- for NI: $\|f_k\| \leq 10^{-14}$;
- for INI$_1$: $\|f_k\| \leq \gamma \min(x_k)$ with some $0 < \gamma < 1$;
- for INI$_2$: $\|f_k\| \leq \min\{\gamma \min(x_k), \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}\}$ for $k \geq 1$ and $\|f_0\| \leq \gamma \min(x_0)$ with some $0 < \gamma < 1$.

We explain the inner tolerance used in INI$_2$. Recall that we always have $\min(x_k) \leq n^{-1/2}$, which is reasonably small for $n$ large. Therefore, starting with a general positive vector $x_0$, the sequence $\{\lambda_k\}$ by INI$_2$ with $\|f_k\| \leq \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}$ may not satisfy $\|f_k\| \leq \gamma x_k$, the condition of Lemma [3.1]. Consequently, the global convergence of $\{\lambda_k\}$ is not guaranteed. However, we must have $\|f_k\| \leq \gamma x_k$ for the above-proposed inner tolerance for INI$_2$, such that $\{\lambda_k\}$ generated by INI$_2$ is globally convergent. Furthermore, once $\lambda_k$ has converged with certain accuracy, we will have $\frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}} < \gamma \min(x_k)$ for $k$ enough large, so that, by Theorem [3.11] INI$_2$ will asymptotically converge superlinearly.

In implementations, it is necessary to impose some guards on inner tolerances so as to avoid them being too small in finite precision arithmetic, i.e., below the level of $\epsilon$. For all examples, the stopping criteria for inner iterations are always taken as $\|f_k\| \leq \max\{\gamma \min(x_k), 10^{-13}\}$ for INI$_1$ and $\|f_k\| \leq \max\{\min\{\gamma \min(x_k), \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}\}, 10^{-13}\}$ for INI$_2$.

For each example, we test two $\gamma = 0.8, 0.1$ to observe the effect of $\gamma$ on the convergence of INI$_1$.

In the experiments, the stopping criterion for outer iterations is

$$\frac{\|Bx_k - \lambda_k x_k\|}{\|B\|_1 \|B\|_\infty} \leq 10^{-13},$$

where we use the cheaply computable $\frac{\|B\|_1 \|B\|_\infty}{\|B\|_1 \|B\|_\infty}$ to estimate the 2-norm $\|B\|_2$, which is more reasonable than the individual $\|B\|_1$ or $\|B\|_\infty$. 


For INI and NI, since the coefficient matrices $\lambda_k I - B$ are always positive definite for all $k$ when $B$ is symmetric, we use the conjugate gradient method to solve inner linear systems. For $B$ unsymmetric, we use BICGSTAB as inner solver. In implementations, we use the standard Matlab functions \texttt{bicgstab} and \texttt{pcg}. The outer iteration starts with the normalized vector of $(1, \ldots, 1)^T$ for NI and INI.

The codes of JDQR and JDRPCG use the absolute residual norms to decide the convergence, which is not robust for a general purpose. By setting the stopping criteria “TOL= $10^{-13}(\|B\|_1 \|B\|_\infty)^{1/2}$” for outer iterations in them, we will get the same stopping criteria as that for NI and INI, the parameters “sigma=LM”, and the inner solver “OPTIONS.LSolver=bicgstab” in the unsymmetric case and “OPTIONS.LSolver=minres” in the symmetric case, where the Matlab function \texttt{minres} is the minimal residual method. All the other options are set to be defaults. We do not use any preconditioning in all the iterative solvers for inner linear systems. For \texttt{eigs}, we set the stopping criteria “OPTS.tol= $10^{-13}/(\|B\|_1 \|B\|_\infty)^{1/2}$”. For the computation of Perron roots and vectors, we remind that \texttt{eigs} is directly applicable without shift-invert technique involved, so we do not solve inner linear systems.

We denote by $I_{\text{outer}}$ the number of outer iterations to achieve the convergence and by $I_{\text{inner}}$ the total number of inner iterations. Note that each outer iteration of these algorithms needs one matrix-vector product formed with $B$, while each iteration of BICGSTAB uses two matrix-vector products formed with $B$ and $B^T$, respectively, and each iteration of MINRES or CG uses one matrix-vector product with $B$. We denote by $I_{\text{total}}$ the total matrix-vector products formed with $B$ and possibly $B^T$, which can measure the overall efficiency of NI, INI, JDQR and \texttt{eigs} when the subspace dimensions used by JDQR and \texttt{eigs} are small at each restart.

In view of the above, we have $I_{\text{total}} = I_{\text{outer}} + 2I_{\text{inner}}$ for $B$ unsymmetric and $I_{\text{total}} = I_{\text{outer}} + I_{\text{inner}}$ for $B$ symmetric for our test algorithms. Particularly, for \texttt{eigs} we have $I_{\text{total}} = I_{\text{outer}}$ since no inner linear system is solved.

We remind that for unsymmetric problems JDQR and \texttt{eigs} may compute physically meaningless complex approximate Perron vectors. If an approximate Perron vector is complex, we are only concerned with the signs of components in its real part. Without loss of generality, we let the maximal component in magnitude of the computed Perron vector be positive. In the tables, “Positivity” records “Does the finally computed Perron vector preserve the strict positivity property”. If the answer is “No”, then the percentage in the brace indicates the proportion that the computed Perron vector has the components with positive real part. We also report the CPU time of each algorithm, which measures the overall efficiency too.

Example 5.1. From DIMACS10 test set [6], we consider the nonsymmetric non-negative matrix \texttt{web-Google}. The matrix data was released in 2002 by Google as a part of Google Programming Contest. The \texttt{web-Google} is a directed graph with nodes representing web pages and directed edges representing hyperlinks between them. This matrix is a binary matrix (a matrix each of whose elements is 0 or 1) of order $n = 916,428$ and has 5,105,039 nonzero entries.

Table 1 reports the results obtained by NI, INI$_1$ with $\gamma = 0.8$ and $\gamma = 0.1$, INI$_2$, JDQR and \texttt{eigs}. Figures 1–2 illustrate how the residual norms of outer iterations evolve versus the sum of inner iterations and versus the outer iterations for NI, INI$_1$ with two $\gamma$ and INI$_2$, and JDQR, respectively. Figure 2 shows that the residual norms computed by the five algorithms decreased monotonically. It also indicates
that the NI and INI algorithms converged slowly and similarly in the beginning of outer iterations. Then they started converging fast. We find that INI$_1$ and INI$_2$ achieved the same superlinear (quadratic) convergence and exhibited very similar convergence behavior to NI, and all of them used nine outer iterations to achieve the convergence. These results show that our theory on both INI$_1$ and INI$_2$ can be conservative. Strikingly, however, we see from Table 1 that although eigs was the fastest and JDQR was as efficient as INI$_1$ and INI$_2$ in terms of $I_{total}$ and the CPU time, they were not positivity preserving. They computed the Perron root reliably, but the finally converged Perron vectors did not positivity. More precisely, as we counted, the the computed Perron vector by JDQR had 544894 negative components, and the one by eigs had 352493 negative components. This means that about 59% and 39% of the components of the computed eigenvectors were physically meaningless, respectively, that is, only 41% and 61% of the components were positive, as is also indicated by Table 1.

The physically meaningless results by JDQR and eigs for a desired positive eigenvector are not unusual: Because some component(s) of the Perron vector must be very small, it is quite possible that the JDQR and eigs cannot guarantee the strict positivity of approximate eigenvectors and, particularly, those very small or tiny true positive components may change signs and become negative in the approximations, even though the approximations have already converged to the positive eigenvector $x$ in the conventional sense and attains its maximum accuracy, namely, the errors of them and $x$ attain the level of $\epsilon$.

| Method     | $I_{outer}$ | $I_{inner}$ | $I_{total}$ | CPU time | Positivity |
|------------|-------------|-------------|-------------|----------|------------|
| NI         | 9           | 115.5       | 240         | 33.4     | Yes        |
| INI$_1$ with $\gamma = 0.8$ | 9           | 55          | 119         | 14.2     | Yes        |
| INI$_1$ with $\gamma = 0.1$ | 9           | 61.5        | 132         | 17.1     | Yes        |
| INI$_2$    | 9           | 56          | 121         | 14.4     | Yes        |
| JDQR       | 4           | 52.5        | 109         | 14.2     | No (41%)   |
| eigs       | 80          | ——          | 80          | 7.3      | No (61%)   |

We see from Table 1 that the INI$_1$ and INI$_2$ improved the overall efficiency of NI considerably. Actually, the computing times of INI$_1$ with $\gamma = 0.8$ and $\gamma = 0.1$ were 43% and 51% of NI, respectively, and that of INI$_2$ was 43% of NI. In terms of either $I_{total}$ or the CPU time, INI$_1$ with $\gamma = 0.8$, INI$_2$ and JDQR were twice as fast as NI. We also see that eigs was the fastest, but unfortunately it could not produce positive approximate Perron vector. It is worthwhile to point out that, as we have observed, NI needed more inner iterations than INI$_1$ and INI$_2$ at each outer iteration, and inner iterations used by each algorithm increased as outer iterations proceeded. This is due to the fact that $\lambda_k$ was closer to $\lambda$ and inner linear system (2.5) was worse conditioned with increasing $k$ and was solved with possibly higher accuracy, as required by INI$_2$ and JDQR. Even so, Figure 5.1 illustrates that INI and JDQR converged superlinearly as inner iterations increase.

**Example 5.2.** Consider the symmetric nonnegative matrix `delaunay_n20` from DIMACS10 test set [6]. The matrix is generated by Delaunay triangulations of random
points in the unit square. It is a binary matrix of order \( n = 2^{20} = 1,048,576 \) with 6,291,372 nonzero entries.

Table 2 and Figures 3–4 report the results and convergence processes. First of all, Table 2 shows that JDQR and \texttt{eigs} were unreliable to compute the Perron vector and were unable to produce positive computed solutions, though they were good for computing the Perron root. \texttt{eigs} was the fastest, and it was about 1.2 \( \sim \) 1.5 times as fast as INI in terms of \( I_{\text{total}} \) and the CPU time. However, it was fatal that \texttt{eigs} was not positivity preserving and produced a physically meaningless approximation of the Perron vector. JDQR was the slowest and unreliable, and NI was the second most expensive in terms of both \( I_{\text{total}} \) and the CPU time. INI1 with two \( \gamma \) and INI2 cost only about 50\% of NI, 33\% of JDQR, a considerable improvement. Table 2 says that only 82\% and 60\% of the components of the finally converged approximations by JDQR and \texttt{eigs} were positive, respectively. Precisely, our records show that the finally computed solutions by them had 194155 and 425103 negative components, respectively, meaning that about 18\% and 40\% of the components of the computed eigenvectors were physically meaningless.

Neglecting the positivity preservation of the computed Perron vector, Table 2 and Figure 4 indicate that NI, INI, INI1 with two \( \gamma \) and JDQR used exactly the same, i.e., ten outer iterations. As is seen from Figure 4, all the algorithms converged slowly until the sixth outer iteration, then they speeded up very considerably and converged superlinearly. Furthermore, from the sixth outer iteration onwards, Figure 4 clearly demonstrates that each algorithm actually achieved the quadratic convergence. It is worthwhile to point out that, unlike NI and INI which converged monotonically, JDQR exhibited irregular convergence behavior: the residual norms of JDQR decreased at the first four outer iterations, increased at the two outer iterations followed, and then converged regularly from the sixth outer iteration upwards until the convergence. Besides, Figure 3 indicates that NI, INI and JDQR converged slowly and linearly in the initial stage, and then they converged increasingly faster, i.e., superlinearly as inner iterations increase.
Table 2. The total outer and inner iterations in Example 5.2

| Method            | \(I_{\text{outer}}\) | \(I_{\text{inner}}\) | \(I_{\text{total}}\) | CPU time | Positivity |
|-------------------|-----------------------|-----------------------|-----------------------|----------|------------|
| NI                | 10                    | 524                   | 534                   | 86       | Yes        |
| INI\(_1\) with \(\gamma = 0.8\) | 10                    | 259                   | 269                   | 40       | Yes        |
| INI\(_1\) with \(\gamma = 0.1\) | 10                    | 291                   | 301                   | 46       | Yes        |
| INI\(_2\)        | 10                    | 261                   | 271                   | 40       | Yes        |
| JDQR with Minres  | 9                     | 750                   | 759                   | 121      | No (82%)   |
| eigs              | 200                   | —                     | 200                   | 28       | No (60%)   |

5.2. INI for M-matrices. In this subsection, we use NI and INI to find the smallest eigenpair of an irreducible nonsingular M-matrix and illustrate the effectiveness of INI. For NI, INI\(_1\) and INI\(_2\), the stopping criteria for inner and outer iterations were the same as those for nonnegative matrices. For JDQR, we also take the same stopping criterion for outer iterations, and the parameter “sigma=SM”. All the other parameters are the same as those for nonnegative matrix eigenvalue problems. For eigs, since we now compute the smallest eigenvalue and the associated eigenvector of an irreducible nonsingular M-matrix \(A\), we use \(\text{eigs}(\sigma \text{speye}(n) - A,'LM',\text{OPTS})\) to replace with \(\text{eigs}(A,'SM',\text{OPTS})\) for some easily chosen \(\sigma > \rho(A)\). The reason of this strategy is twofold: First, \(\sigma I - A\) becomes an irreducible nonnegative matrix, which means that we only need to compute the largest eigenpair of \(\sigma I - A\) without solving any linear system and we then recover the smallest eigenpair of \(A\) trivially. Second, if \(\text{eigs}(A,'SM',\text{OPTS})\) is used, we have to solve a sequence of inner linear systems by first making a sparse LU factorization of \(A\) and then solving lower and upper triangular systems. This is assumed to be prohibited for a very large \(A\) throughout this paper.

Example 5.3. We consider a nonsymmetric sparse M-matrix from a 3D Human Face Mesh [9] with a small noise. This is a matrix of order \(n = 42,875\) with 693,875 nonzero entries.

We choose \(\sigma\) in such a way: let \(d = \max_{1 \leq i \leq n} \{a_{ii}\}\), which equals 2771, and take \(\sigma = 3000\). We then get a nonnegative matrix \(B = \sigma I - A\). Table 3 reports the
results obtained, and Figures 5–6 depict the convergence processes of NI, INI and JDQR.

For this problem, the good news is that all the algorithms worked well and all the computed eigenvectors preserved the positivity. We see from the figures that NI, INI and JDQR converged smoothly and fast as inner iterations increase and outer iterations proceed. Except INI\_1 with $\gamma = 0.8$, NI and INI used comparable outer iterations to attain the desired accuracy. However, for the overall efficiency, NI was the most expensive, and INI\_2 improved considerably and used 65% of $I_{total}$ and 67% of the CPU time of NI. INI\_1 with two $\gamma$ improved the performance of NI more substantially. INI\_1 with $\gamma = 0.8$ was the most efficient in terms of $I_{total}$ and the CPU time, and eigs used almost the same CPU time as it but considerably more $I_{total}$. In addition, we have observed that all the other INI were competitive with JDQR.

Finally, we see that for this $M$-matrix two different $\gamma$ led to distinct convergence behavior. As Figure 5 indicates, NI and INI\_2 typically converged superlinearly, and INI\_1 with $\gamma = 0.1$ exhibited very similar superlinear convergence to NI, while INI\_1 with $\gamma = 0.8$ typically converged linearly. These confirm our theory, demonstrating that our theoretical results can be realistic and pronounced.

**Table 3.** The total outer and inner iterations in Example 5.3

| Method         | $I_{outer}$ | $I_{inner}$ | $I_{total}$ | CPU time | Positivity |
|----------------|-------------|-------------|-------------|----------|------------|
| NI             | 7           | 1400        | 2807        | 19.1     | Yes        |
| INI\_1 with $\gamma = 0.8$ | 10          | 521         | 1052        | 7.2      | Yes        |
| INI\_1 with $\gamma = 0.1$ | 7           | 773         | 1553        | 10.2     | Yes        |
| INI\_2         | 7           | 912         | 1831        | 12.9     | Yes        |
| JDQR           | 6           | 506         | 1018        | 10.9     | Yes        |
| eigs           | 1420        | —           | 1420        | 7.0      | Yes        |

**Figure 5.** Example 5.3

The outer residual norms versus sum of inner iterations.

**Figure 6.** Example 5.3

The outer residual norms versus the outer iterations.

**Example 5.4.** Consider the $M$-matrix nicolo\_da\_uzzano from AIM@SHAPE Shape Repository [2] by using the technique in [32]. This is a symmetric matrix and
describes the full resolution shape (2 millions triangles) edited to remove the errors introduced by scanning and reconstruction phases. The matrix \( \text{nico} \) is generated by the barycentric mapping method [32], is of order \( n = 943,870 \) and has 8,960,880 nonzero entries.

For \( \text{eigs} \), we obtain a nonnegative matrix \( B = \sigma I - A \) in this way: let \( d = \max_{1 \leq i \leq n} \{a_{ii}\} \), which equals 2421, and then take \( \sigma = 2500 \). Table 4 reports the numerical results obtained by NI, INI, JDQR, JDRPCG and \( \text{eigs} \). Figures 7–8 describe the convergence processes of NI and INI. We omit the convergence curves of JDQR and JDRPCG because they performed very poorly on this problem, as is seen from Table 4.

As far as \( I_{\text{outer}} \) is concerned, NI and INI worked very well and used only four outer iterations to achieve the desired accuracy. As Figure 8 show, they had indistinguishable convergence curves. Furthermore, all of them were reliable and positivity preserving at each outer iteration. For the overall performance, INI and INI were equally efficient, and they used about 60% cost of NI in terms of both \( I_{\text{total}} \) and the CPU time.

In contrast, JDQR, JDRPCG and \( \text{eigs} \) performed poorly for this problem. At much more expenses than NI and INI, JDQR finally failed to compute a correct eigenvalue, and it misconverged and computed some other eigenpair. JDRPCG and \( \text{eigs} \) improved JDQR substantially, and they both computed the desired eigenvalue but delivered physically meaningless eigenvectors. Precisely, our records show that the final converged eigenvectors obtained by JDRPCG and \( \text{eigs} \) were not positive, and had 410651 and 549218 negative components, respectively, meaning that about 43% and 58% of the components of the computed eigenvectors became meaningless.

Neglecting the reliability or the positivity, for the overall efficiency, JDRPCG was competitive with INI in terms of the CPU time and \( I_{\text{total}} \). However, \( \text{eigs} \) used at least as twice \( I_{\text{total}} \) as INI, and it consumed 1.6 times of the CPU time of INI.

| Method       | \( I_{\text{outer}} \) | \( I_{\text{inner}} \) | \( I_{\text{total}} \) | CPU time | Positivity |
|--------------|-------------------------|-------------------------|-------------------------|----------|------------|
| NI           | 4                       | 697                     | 701                     | 54       | Yes        |
| INI1 with \( \gamma = 0.8 \) | 4                       | 400                     | 404                     | 31       | Yes        |
| INI1 with \( \gamma = 0.1 \) | 4                       | 440                     | 444                     | 33       | Yes        |
| INI2         | 4                       | 405                     | 409                     | 31       | Yes        |
| JDQR with Minres | 53                      | 10458                   | 10511                   | 1480     | Wrong eigenpair |
| JDRPCG       | 146                     | 279                     | 425                     | 23       | No (57%)   |
| \( \text{eigs} \) | 940                     | ——                     | 940                     | 53       | No (42%)   |

Remark 5.5. We have tested all the above examples using the power method, which preserves the positivity of the approximate eigenvectors at each iteration and is linearly convergent for any positive starting vector. For Examples 5.1 and 5.2, we have found that the power method needed 3 \( \sim \) 6 times of the CPU cost of INI1 because of the eigenvalue clustering. For \( M \)-matrices in Examples 5.3 and 5.4, we applied the power method to the nonnegative matrix \( B \) of the form \( B = \sigma I - A \) with \( \sigma \) chosen previously, for finding \( \rho(B) \) and the Perron vector. The power method needed more than ten times of the CPU cost of INI1. So INI was much more
efficient than the power method for these four examples. We omit the details on numerical results obtained by the power method.

6. Conclusions

For the efficient computation of the smallest eigenpair of a large irreducible non-singular $M$-matrix, we have proposed a positivity preserving inexact Noda iteration method with two practical inner tolerance strategies provided for solving the linear systems involved. We have analyzed the convergence of the method in detail, and have established a number of global linear and superlinear convergence results, with the linear convergence factor and the superlinear convergence order derived explicitly. Precisely, we have proved that INI with any inner solver converges at least linearly for $\xi_k = \|f_k\| \leq \gamma \min(x_k)$ and superlinearly with the convergence order $1 + \sqrt{2}$ for the decreasing $\|f_k\| \leq \frac{\lambda_k - \lambda_k^{-1}}{2}$, respectively. We have also revisited the convergence of NI and proved its quadratic convergence in a form different from [7]. The results on INI clearly show how inner tolerance affects the convergence of outer iterations.

Numerically, the proposed INI algorithms have been shown to be practical for large nonnegative matrix and $M$-matrix eigenvalue problems, and they can reduce the total computational cost of NI substantially. In the meantime, the numerical experiments have shown that the INI algorithms are superior to the positivity non-preserving Jacobi–Davidson method and eigs. They always preserve the positivity of approximate eigenvectors and produce physically meaningful approximations, while the Jacobi–Davidson method and eigs often fail to produce positive approximate eigenvectors and thus deliver physically meaningless results. Moreover, the overall efficiency of the proposed INI algorithms is competitive with and can be considerably higher than the Jacobi–Davidson method and eigs for some practical problems.

References

[1] J. Abate, L. Choudhury and W. Whitt, Asymptotics for steady-state tail probabilities in structured Markov queueing models, Commun. Statist-Stochastic Models, 1 (1994), pp. 99–143.
[2] AIM@SHAPE: Advanced and innovative models and tools for the development of semantic-based systems for handling, acquiring, and processing knowledge embedded in multidimensional digital objects. Available at http://shapes.aim-at-shape.net.

[3] A. S. Alfa, J. Xue and Q. Ye, Accurate computation of the smallest eigenvalue of a diagonally dominant $M$-matrix, Math. Comp., 71 (2002), pp. 217–236.

[4] A. Berman and R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Vol. 9 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1994.

[5] J. Berns-Müller, I. G. Graham and A. Spence, Inexact inverse iteration for symmetric matrices, Linear Algebra Appl., 416 (2006), pp. 389–413.

[6] DIMACS10 test set and the University of Florida Sparse Matrix Collection. Available at www.cise.ufl.edu/research/sparse/dimacs10/.

[7] L. Elsner, Inverse iteration for calculating the spectral radius of a non-negative irreducible matrix, Linear Algebra and Appl., 15 (1976), pp. 235–242.

[8] G. H. Golub and C. F. Van Loan, Matrix Computations, The John Hopkins University Press, Baltimore, London, 4th ed., 2012.

[9] X. D. Gu, T.-M. Huang, W.-Q. Huang, S.-S. Lin, W.-W. Lin and S.-T. Yau, High performance computing for spherical conformal and Riemann mappings, Technical Report, Department of Applied Mathematics, National Chiao Tung University, Taiwan, 2013.

[10] R. A. Horn and C. R. Johnson, Matrix Analysis, The Cambridge University Press, Cambridge, UK, 1985.

[11] Z. Jia, On convergence of the inexact Rayleigh quotient iteration with MINRES, J. Comput. Appl. Math., 236 (2012) pp. 4276–4295.

[12] Z. Jia, On convergence of the inexact Rayleigh quotient iteration with the Lanczos method used for solving linear systems, Sci. China Math., 56 (2013), pp. 2145–2160.

[13] Z. Jia and C. Li, On inner iterations in the shift-invert residual Arnoldi method and the Jacobi–Davidson method, Sci. China Math., 2014. doi: 10.1007/s11425-014-4791-5.

[14] A. Langville and C. Meyer, Google’s PageRank and Beyond: the Science of Search Engine Rankings, The Princeton University Press, Princeton, 2006.

[15] C. Lee, Residual Arnoldi method: theory, package and experiments, Ph D thesis, TR-4515, Department of Computer Science, University of Maryland at College Park, 2007.

[16] Y.-L. Lai, K.-Y. Lin and W.-W. Lin, An inexact inverse iteration for large sparse eigenvalue problems, Numer. Linear Algebra Appl., 4 (1997), pp. 425–437.

[17] M. S. Lynn and W. P. Timlake, Bounds for Perron eigenvectors and subdominant eigenvalues of positive matrices, Linear Algebra Appl., 2 (1969), pp. 143–152.

[18] T. Noda, Note on the computation of the maximal eigenvalue of a non-negative irreducible matrix, Numer. Math., 17 (1971), pp. 382–386.

[19] Y. Notay, JDCG and JDRPCG codes, available at [http://homepages.ulb.ac.be/~ynotay/]

[20] A. M. Ostrowski, On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. V. (Usual Rayleigh quotient for non-Hermitian matrices and linear elementary divisors), Arch. Rational Mech. Anal., 3 (1959), pp. 472–481.

[21] B. N. Parlett, The Symmetric Eigenvalue Problem, Vol. 20 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1998.

[22] M. Robbé, M. Sadkane and A. Spence, Inexact inverse subspace iteration with preconditioning applied to non-Hermitian eigenvalue problems, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 92–113.

[23] G. Robertazzi, Computer Networks and Systems: Queueing Theory and Performance Evaluation, Springer-Verlag, 1990.

[24] Y. Saad, Numerical Methods for Large Eigenvalue Problems, Revised Version, Vol. 66 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 2011.

[25] D. C. Sorensen, Implicit application of polynomial filters in a k-step Arnoldi method, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 92357–385.

[26] P. Shivakumar, J. Williams, Q. Ye and C. Marinov, On two-sided bounds related to weakly diagonally dominant $M$-matrices with applications to digital circuit dynamics, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 298–312.

[27] V. Simoncini, L. Eldn, Inexact Rayleigh quotient-type methods for eigenvalue computations, BIT 42 (2002), pp. 159–182.

[28] G. L. G. Sleijpen and H. A. Van der Vorst, A Jacobi–Davidson iteration method for linear eigenvalue problems, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.
[29] G. L. G. Sleijpen, *JDQR and JDQZ codes*, available at [http://www.math.uu.nl/people/sleijpen](http://www.math.uu.nl/people/sleijpen).

[30] G. W. Stewart, *Matrix Algorithms Vol. II*, SIAM, Philadelphia, PA, 2001.

[31] R. S. Varga. *Matrix Iterative Analysis*, 2nd-ed., Springer-Verlag Berlin Heidelberg, 2000.

[32] M. Q. Wei, M. Y. Pang and C. L. Fan, *Survey on planar parameterization of triangular meshes*, Proceedings of the 2010 International Conference on Measuring Technology and Mechatronics Automation, 2010, pp. 702–705.

[33] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[34] F. Xue, D.B. Szyld, *Efficient preconditioned inner solves for inexact Rayleigh quotient iteration and their connections to the single-vector Jacobi–Davidson method*, SIAM J. Matrix Anal. Appl. 32 (2011), pp. 993–1018.

[35] J. Xue, *Computing the smallest eigenvalue of an M-matrix*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 748–762.