On semilocal convergence analysis for two-step Newton method under generalized Lipschitz conditions in Banach spaces

Yonghui Ling\textsuperscript{1} · Juan Liang\textsuperscript{1} · Weihua Lin\textsuperscript{1}

Received: 8 October 2020 / Accepted: 11 September 2021 / Published online: 6 October 2021 © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
In the present paper, we consider the semilocal convergence issue of two-step Newton method for solving nonlinear operator equation in Banach spaces. Under the assumption that the first derivative of the operator satisfies a generalized Lipschitz condition, a new semilocal convergence analysis for the two-step Newton method is presented. The Q-cubic convergence is obtained by an additional condition. This analysis also allows us to obtain three important spatial cases about the convergence results based on the premises of Kantorovich, Smale and Nesterov-Nemirovskii types. As applications of our convergence results, a nonsymmetric algebraic Riccati equation arising from transport theory and a two-dimensional nonlinear convection-diffusion equation are provided.

Keywords Two-step Newton method · Generalized Lipschitz conditions · Semilocal convergence · Algebraic Riccati equation

Mathematics Subject Classification (2010) 49M15 · 47J25 · 65J15

1 Introduction
In this paper, we aim to study the semilocal convergence of iterative methods for approximating the solution of nonlinear operator equation

\[ F(x) = 0, \]  

where \( F \) is a given Fréchet differentiable nonlinear operator which maps from some open convex subset \( \mathbb{D} \) in a Banach space \( \mathbb{X} \) to another Banach space \( \mathbb{Y} \). Newton’s method is an efficient and presumably the most important iterative method known to

\textsuperscript{1} Department of Mathematics, Minnan Normal University, Zhangzhou, 363000, China
solve such equation. For a given initial guess \( x_0 \in \mathbb{D} \), the sequence \( \{x_k\} \) constructed by Newton’s method proceeds the following update rule:

\[
x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \quad k = 0, 1, 2, \ldots
\]

(2)

A crucial semilocal convergence result for Newton’s method (2) is the famous Kantorovich theorem [39], which gives a simple and clear criterion guaranteeing the existence of a solution of (1), uniqueness of this solution in an appropriate ball and the quadratic convergence of Newton’s method (2). The hypotheses used mainly focus on the nonsingularity of the first Fréchet derivative of \( F \) at initial guess \( x_0 \) and the behavior of the Fréchet derivative of \( F \) on an prescribed ball of the initial guess \( x_0 \) contained in \( \mathbb{D} \) (such as that the second Fréchet derivative of \( F \) is bounded or the first Fréchet derivative of \( F \) is Lipschitz continuous). There are a large number of works on the weakness and/or extension of the hypotheses made on the operator \( F \) and its derivative (see, for example, [4, 6, 19, 20, 24, 29, 32, 35]). More studies and applications of the Kantorovich-like theorem, one can see the recent survey paper by Kelley in [40] for more details.

Another crucial semilocal convergence result concerning Newton’s method (2) is the well-known Smale \( \alpha \)-theory (with analytic \( F \)) [60] (see also [59, 61]), where the notion of approximate zeros was proposed and the rules were established to determine if an initial guess \( x_0 \) is an approximate zero. Since then, numerous remarkable improvements and extensions have been made along this line of research (see, for example, [12, 15–18, 44, 58, 64, 66] and references therein). In particular, to drop the analytic assumption, Wang and Han in [65] (see also [63]) introduced the notion of weak \( \gamma \)-condition for nonlinear twice continuously differentiable operator \( F \) between Banach spaces. This notion was also used in [46] and [43] to extend and improve the corresponding results in [16] and [14], respectively. Recently, the main results obtained in [43] were improved further by the corresponding ones in [45].

Wang in [63] introduced some generalized Lipschitz condition called Lipschitz condition with \( L \)-average, under which Kantorovich-like convergence theorem and Smale’s \( \alpha \)-theory can be investigated together. The generalized Lipschitz condition was also used to study the semilocal convergence for various iterative methods. For example, Xu and Li in [68] investigated the semilocal convergence of Gauss-Newton’s method for singular systems with constant rank derivatives under the assumption that the first derivative of operator \( F \) satisfy the generalized Lipschitz condition. This notion was extended to study the convergence of Gauss-Newton’s method for one kind of special singular systems in [41] and improved further in [5]. Besides, Alvarez et al. in [1] extended this notion to a Riemannian context and established a unified convergence result for Newton’s method. Recently, Wang et al. [62] developed the majorizing function technique for the convergence analysis of an extended Newton method for multiobjective optimization under the \( L \)-average Lipschitz condition. Moreover, Ferreira and Svaiter in [27] presented a new, simple and clear semilocal convergence analysis for Newton’s method by using Kantorovich’s majorants principle. More extensions of this idea are referred to [9, 26, 28]. It is worth pointing out that the majorant condition used in [27] is equivalent to the above generalized Lipschitz condition (see [67]).
Euler’s method and Halley’s method are two of the most important cubic generalization for Newton’s method. Kantorovich-like theorem and Smale-like $\alpha$-theory for these two iterative methods are referred to [10, 11, 21–23, 33, 49] and references therein. For the applications in the field of matrix functions, the efficiency (when properly implemented) of these two iterative methods have been shown in [36, 48] for computing matrix $p$th root and in [54, 55] for computing the polar decomposition of a matrix.

For general nonlinear (1), however, the preceding classical third-order methods need to evaluate the second Fréchet derivative which is very time consuming. The order of convergence of the classical two-step Newton method has also three, but without evaluating any second Fréchet derivative. The two-step Newton method with initial guess $x_0$ is defined via the following update rule:

$$
\begin{align*}
  y_k &= x_k - \frac{1}{F'(x_k)} F(x_k), \\
  x_{k+1} &= y_k - \frac{1}{F'(y_k)} F(y_k),
\end{align*}
$$

(3)

The results concerning semilocal convergence (including existence, uniqueness and convergence) of this iterative method have been studied under the assumptions of Newton-Kantorovich type. By applying the majorizing function technique used in the previous work of Zabrejko and Nguyen [69] on Newton’s method, Appell et al. [3] established the semilocal convergence and error estimate under the hypothesis that the first Fréchet derivative of $F$ satisfies the Lipschitz condition. Amat et al. [2] investigated the convergence behavior based fundamentally on a generalization required to the second Fréchet derivative of $F$. Recently, to weak the conditions used in [2], Magreñán Ruiz and Argyros in [53] presented new convergence analysis (including semilocal and local convergence) under the assumptions that the first Fréchet derivative of $F$ satisfies Lipschitz and Lipschitz-like conditions.

The motivation of this paper is based on the following two applications for the two-step Newton method (3). The first one stems from [47, 50] for solving a non-symmetric algebraic Riccati equation (NSARE) arising in transport theory, where the monotone convergence guaranteeing the implementation of the two-step Newton based algorithm was showed. The second one stems from [13, 52] for the inverse eigenvalue problems (IEP) and inverse singular value problems (ISVP), where the two-step Newton method (3) was used to present effective algorithms for solving the solutions of the IEP and ISVP.

The goal of this paper is to establish a general semilocal convergence result for the two-step Newton method (3) under the assumption that the first derivative of $F$ satisfies some generalized Lipschitz condition, which was introduced by Wang in [63] for Newton’s method (2). When the unified convergence criterion given in [63] is satisfied, we show that the existence and uniqueness of a solution, together with the Q-superquadratic convergence of the two-step Newton method (3). The main novelty in our convergence analysis is the relationships between majorizing function $h$ and nonlinear operator $F$ are made clear by using the convexity of $h'$. Moreover, we show also that the two-step Newton method (3) is Q-cubically convergent under a slightly stronger condition. In particular, this convergence analysis allows us to obtain some important special cases, which include Kantorovich-type
convergence result under the Lipschitz condition, Smale-type convergence results under the \( \gamma \)-condition and convergence result for self-concordant functions under the Nesterov-Nemirovskii condition. We also adapt our convergence result to compute the approximation of minimal positive solution of a NSARE arising from transport theory. Numerical experiments confirm our convergence result. Note that, the \( L \)-average Lipschitz condition implies the classical Lipschitz condition with the Lipschitz constant being the supremum of the function \( L(\cdot) \) in the involved ball. One of the main advantages of adopting the \( L \)-average Lipschitz condition is shown in Section 4.2, in which a two-dimensional nonlinear convection-diffusion equation is considered. That is, when the convergence result under the classical Lipschitz condition is not applicable, it provides the possibility to choose some suitable nonnegative and monotonically increasing function \( L(\cdot) \) such that the convergence result under the \( L \)-average Lipschitz condition is applicable for ensuring the convergence of the two-step Newton method.

It is worth noting that the majorizing function technique used in the present paper has been widely used in the convergence analysis of Newton-type methods for nonlinear operator equation [25, 26, 63], of the Gauss-Newton method for convex composite optimization [42] and of extended Newton method for multiobjective optimization [62]. This analysis tool enables us to establish the more precise convergence criterion and estimate of convergence radius for the iterative methods.

The rest of this paper is organized as follows. In Section 2, we introduce some preliminary notions and properties of the majorizing function and majorizing sequences. The semilocal convergence analysis of the two-step Newton method is provided in Section 3, under the \( L \)-average Lipschitz condition. In Section 4, applications to a nonsymmetric algebraic Riccati equation and a two-dimensional nonlinear convection-diffusion equation are presented. We conclude with some final remarks in Section 5.

2 Preliminaries

Let \( X \) and \( Y \) be Banach spaces. For \( x \in X \) and real \( r > 0 \), throughout the whole paper, we use \( B(x, r) \) to stand for the open ball with center \( x \) and radius \( r \), and \( \bar{B}(x, r) \) denote its closure. Moreover, \( I \) denotes the identity operator.

We assume that \( L(\cdot) \) is a positive nondecreasing function on \( [0, R) \), where \( R > 0 \) satisfies

\[
\frac{1}{R} \int_0^R L(u)(R - u) \, du = 1. \tag{4}
\]

Let \( \beta > 0 \). The majorizing function \( h : [0, R] \to \mathbb{R} \) is defined by

\[
h(t) = \beta - t + \int_0^t L(u)(t - u) \, du, \quad t \in [0, R]. \tag{5}
\]

This majorizing function was introduced by Wang in [63] to study the semilocal convergence of Newton’s method (2). It was proved in [67, Proposition 1.2] that this majorizing function is equivalent to the one used in [27]. The main reason we use this majorizing function in the present paper is that it may provide the more concrete
convergence criterion and error estimate for the two-step Newton method. Clearly, we have
\[ h'(t) = -1 + \int_0^t L(u) \, du, \quad t \in [0, R) \]
and
\[ h''(t) = L(t) > 0 \quad \text{for a.e. } t \in [0, R). \]
Then, we obtain that
\[ \int_s^t L(u) \, du = h'(t) - h'(s) \quad \text{for any } s, t \in [0, R) \text{ with } s < t. \]
This simple equality will be frequently applied to our convergence analysis for the two-step Newton method (3). Assume that \( r_0 \) satisfies
\[ \int_0^{r_0} L(u) \, du = 1. \quad (6) \]
It follows that \( h(t) \) is strictly convex, \( h'(t) \) increasing, convex and \(-1 \leq h'(t) < 0\) for any \( t \in [0, r_0) \).

The following lemma gives some properties about elementary convex analysis (see Theorem 4.1.1 and Remark 4.1.2 in [34, p.21]) and will also be frequently used in our convergence analysis for the two-step Newton method (3).

**Lemma 1** Let \( f : (0, R) \to \mathbb{R} \) be a continuously differentiable and convex function, where \( R > 0 \). Then,

\[
\begin{align*}
(1) & \quad (1 - \theta) f'(\theta t) \leq f(t) - f(\theta t) \leq (1 - \theta) f'(t) \quad \text{for all } t \in (0, R) \text{ and } 0 \leq \theta \leq 1. \\
(2) & \quad \frac{f(u) - f(\theta u)}{u} \leq \frac{f(v) - f(\theta v)}{v} \quad \text{for all } u, v \in (0, R), u < v \text{ and } 0 \leq \theta \leq 1.
\end{align*}
\]
In particular, if \( f \) is strictly convex, then the above inequalities are strict.

Define
\[ b := \int_0^{r_0} L(u) u \, du, \quad (7) \]
where \( r_0 \) is defined by (6). The following lemma is taken from [63, Lemma 1.2] which gives some basic properties for the majorizing function \( h \) defined by (5).

**Lemma 2** [63] If \( 0 < \beta < b \), then \( h \) is decreasing on \([0, r_0]\) and increasing on \([r_0, R]\), and
\[ h(\beta) > 0, \quad h(r_0) = \beta - b < 0, \quad h(R) = \beta > 0. \quad (8) \]
Moreover, \( h \) has a unique zero in each interval, denoted by \( t^* \) and \( t^{**} \). They satisfy
\[ \beta < t^* < \frac{r_0}{b} \cdot \beta < r_0 < t^{**} < R. \quad (9) \]
Choose initial point \( t_0 = 0 \). Let \( \{s_k\} \) and \( \{t_k\} \) denote the corresponding sequences generated by the two-step Newton method for the majorizing function \( h \) given in (5),
that is,
\[
\begin{cases}
  s_k = t_k - \frac{h(t_k)}{h'(t_k)}, \\
  t_{k+1} = s_k - \frac{h(s_k)}{h'(t_k)},
\end{cases}
\quad k = 0, 1, 2, \ldots. \tag{10}
\]

**Remark 1** Suppose that $0 < \beta \leq b$. Applying Lemmas 1 and 2, and standard analytical techniques (see, for example, [49]), it is easy to show that the sequences \{s_k\} and \{t_k\} generated by (10) satisfy the following relations:
\[
0 \leq t_k < s_k < t_{k+1} < t^* \quad \text{for all } k \geq 0,
\tag{11}
\]
and converge increasingly to the same point $t^*$, where $t^*$ is the unique zero of $h$ on $[0, r_0]$, $r_0$ is given by (6). Moreover, we can obtain
\[
t^* - t_{k+1} \leq \frac{1}{2} \left( \frac{h''(t^*)}{h'(t^*)} \right)^2 (t^* - t_k)^3, \quad k \geq 0.
\tag{12}
\]
In particular, if $2 + t^* h''(t^*) / h'(t^*) \geq 0$, then we have
\[
s_k - t_k \geq (t^* - t_k) + \frac{h''(t^*)}{2h'(t^*)} (t^* - t_k)^2, \quad k \geq 0.
\tag{13}
\]
The above convergence properties of the sequences \{s_k\} and \{t_k\} will be applied to our convergence analysis for the two-step Newton (3).

we conclude this section with the notions of generalized Lipschitz condition and Q-order of convergence.

**Definition 1** Let $x_0 \in \mathbb{D}$ be such that $F'(x_0)^{-1}$ is nonsingular and $r > 0$ such that $B(x_0, r) \subseteq \mathbb{D}$. Then, $F'$ is said to satisfy the $L$-average Lipschitz condition on $B(x_0, r)$ if, for any $x, y \in B(x_0, r)$ with $\|x - x_0\| + \|y - x\| < r$,
\[
\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \leq \int_{\|x-x_0\|+\|y-x\|} L(u) \, du.
\tag{14}
\]

The preceding generalized Lipschitz condition was first introduced by Wang in [63] where the terminology of “the center Lipschitz condition in the inscribed sphere with $L$-average” was used. Subsequently, to study the convergence behavior of Gauss-Newton, some modified versions were introduced by Li and Ng in [42] for convex composite optimization and Li et al. in [41] for singular systems of equations.

Obviously, according to the above definition, the $L$-average Lipschitz condition on $B(x_0, r)$ implies the classical Lipschitz condition with Lipschitz constant being $L(r)$. The introduction of the $L$-average Lipschitz condition is beneficial to provide the more precise convergence criterion and estimate of convergence radius for Newton’s method. See [63] for more details.

**Definition 2** Let sequence \{x_k\} $\subset \mathbb{X}$. We say that \{x_k\} converges to $x^*$ with Q-superquadratic if, for any $c > 0$, there exists a constant $N_c \geq 0$ such that $\|x_{k+1} - x^*\| \leq c \cdot \|x_k - x^*\|$.\]
\(x^* \parallel \leq c \|x_k - x^*\|^2\) holds for all \(k \geq N_c\), or equivalently

\[\lim_{k \to \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} = 0.\]

In addition, we say that \(\{x_k\}\) converges to \(x^*\) with Q-cubic if there exist two constants \(c \geq 0\) and \(N_c \geq 0\) such that \(\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^3\) holds for all \(k \geq N_c\), or equivalently

\[\lim \sup_{k \to \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^3} < \infty.\]

Q-order of convergence is well-known concept that measures the speed of convergence of sequences. One can see [37, 57] for more properties on this notion.

3 The main theorem and corollaries

In this section, we begin with some technical lemmas about error estimates for the majorizing sequences \(\{s_k\}\) and \(\{t_k\}\) defined by (10), and about the relationships between majorizing function \(h(t)\) defined by (5) and nonlinear operator \(F\). Then, we provide the convergence analysis of the two-step Newton method (3) under \(L\)-average Lipschitz condition (14) in Banach spaces.

3.1 Technical lemmas

Let \(x_0 \in \mathbb{D}\) be an initial guess such that the inverse \(F'(x_0)^{-1}\) exists and let \(B(x_0, r_0) \subset \mathbb{D}\), where \(r_0\) satisfies (6). Set

\[\beta := \|F'(x_0)^{-1}F(x_0)\|.\]

Recall that the majorizing function \(h\) is defined by (5), \(b\) is given by (7), \(t^*\) and \(t^{**}\) are the unique zeros of \(h\) on \([0, r_0]\) and \([r_0, R]\), respectively, where \(r_0\) and \(R\) satisfies (6) and (4). Recall also that the sequences \(\{s_k\}\) and \(\{t_k\}\) generated by (10) converge increasingly to \(t^*\) when \(0 < \beta \leq b\), where \(b\) is defined by (7).

The following lemmas, which provide clear relationships between majorizing function \(h\) and nonlinear operator \(F\), will play key roles for the semilocal convergence analysis of the two-step Newton method (3).

**Lemma 3** Assume that \(\|x - x_0\| \leq t < t^*\). If the first derivative \(F'\) satisfies the \(L\)-average Lipschitz condition (14) in \(B(x^*, t)\), then \(F'(x)\) is nonsingular and

\[\|F'(x)^{-1}F'(x_0)\| \leq - \frac{1}{h'(\|x - x_0\|)} \leq - \frac{1}{h'(t)}.\]

In particular, \(F'\) is nonsingular in \(B(x_0, t^*)\).
Proof Take \( x \in B(x_0, t), 0 \leq t < t^* \). By using the \( L \)-average Lipschitz condition (14), we have

\[
\| F'(x_0)^{-1} F'(x) - I \| \leq \int_0^{\| x - x_0 \|} L(u) \, du = h'(\| x - x_0 \|) - h'(0).
\]

Since \( h'(0) = -1 \) and \( h' \) is strictly increasing in \( (0, t^*) \), we obtain

\[
\| F'(x_0)^{-1} F'(x) - I \| \leq h'(t) + 1 < 1,
\]

the last due to \(-1 < h'(t) < 0\) for any \( t \in (0, t^*) \). Therefore, the Banach lemma is applicable to conclude that \( F'(x_0)^{-1} F'(x) \) is nonsingular and (16) holds. The proof is complete. \( \square \)

Lemma 4 Let \( \{s_k\} \) and \( \{t_k\} \) be generated by (10). Assume that \( F' \) satisfies the \( L \)-average Lipschitz condition (14) in \( B(x_0, t^*) \). If \( 0 < \beta \leq b \), then the sequences \( \{x_k\} \) and \( \{y_k\} \) generated by the two-step Newton method (3) with initial guess \( x_0 \) are well defined and contained in \( B(x_0, t^*) \). Moreover, for all \( k = 0, 1, 2, \ldots \), we have

(i) \( F'(x_k)^{-1} \) exists and \( \| F'(x_k)^{-1} F'(x_0) \| \leq -1 / h'(\| x_k - x_0 \|) \) \( \leq -1 / h'(t_k) \).

(ii) \( \| F'(x_0)^{-1} F(x_k) \| \leq h(t_k) \).

(iii) \( \| x_{k+1} - x_k \| \leq s_k - t_k \).

(iv) \( \| x_{k+1} - y_k \| \leq (t_{k+1} - s_k) \cdot \left( \frac{\| y_k - x_k \|}{s_k - t_k} \right)^2 \leq t_{k+1} - s_k \).

(v) \( \| x_{k+1} - x_k \| \leq t_{k+1} - t_k \).

Proof We reason by induction. The case \( k = 0 \) is true obviously for (i)–(iii). Thus \( y_0 \in B(x_0, t^*) \) owing to \( \| y_0 - x_0 \| \leq s_0 - t_0 = s_0 < t^* \). As for (iv) and (v), by (3), we have

\[
F(y_0) = F(y_0) - F(x_0) - F'(x_0)(y_0 - x_0)
= \int_0^1 [F'(x_0 + \tau (y_0 - x_0)) - F'(x_0)](y_0 - x_0) \, d\tau.
\]

Then, the \( L \)-average Lipschitz condition (14) is applicable to deduce that

\[
\| F'(x_0)^{-1} F(y_0) \| \leq \int_0^1 \| F'(x_0)^{-1} F'(x_0 + \tau (y_0 - x_0)) - F'(x_0) \| \| y_0 - x_0 \| \, d\tau
\leq \int_0^1 \left( \int_0^{\| y_0 - x_0 \|} L(u) \, du \right) \| y_0 - x_0 \| \, d\tau
= \int_0^1 [h'(\| y_0 - x_0 \|) - h'(0)] \| y_0 - x_0 \| \, d\tau.
\]

In view of \( h' \) is strictly convex in \([0, r_0] \) and noting that \( \| y_0 - x_0 \| \leq s_0 - t_0 \) by (iii), it follows from Lemma 1 that

\[
h'(\| y_0 - x_0 \|) - h'(0) = \frac{h'(\| y_0 - x_0 \|) - h'(0)}{\| y_0 - x_0 \|} \cdot \| y_0 - x_0 \|
\leq \frac{h'(r_0 - t_0) - h'(0)}{s_0 - t_0} \cdot \| y_0 - x_0 \|.
\]
Then, combining the above inequality and (10), one has that

\[
\|F'(x_0)^{-1}F(y_0)\| \leq \int_0^1 [h'(\tau s_0) - h'(0)]s_0 \, d\tau \cdot \left( \frac{\|y_0 - x_0\|}{s_0 - t_0} \right)^2
\]

This leads to

\[
\|x_1 - y_0\| = \|F'(x_0)^{-1}F(y_0)\| \leq (t_1 - s_0) \cdot \left( \frac{\|y_0 - x_0\|}{s_0 - t_0} \right)^2.
\]

Hence, we have

\[
\|x_1 - x_0\| \leq \|x_1 - y_0\| + \|y_0 - x_0\| \leq (t_1 - s_0) + (s_0 - t_0) = t_1 - t_0.
\]

That is to say, (iv) and (v) hold for the case \(k = 0\), which implies that \(x_1 \in B(x_0, t^*)\). Now we assume that \(x_k, y_k \in B(x_0, t^*)\), \(\|x_k - x_0\| \leq t_k\) and (i)–(v) hold for some \(k \geq 0\). Then, applying the inductive hypothesis (iii), we obtain that \(\|y_k - x_0\| \leq \|y_k - x_k\| + \|x_k - x_0\| \leq s_k\). In addition, we use the inductive hypothesis (v) and (11) to yield

\[
\|x_{k+1} - x_0\| \leq \sum_{i=0}^k \|x_{i+1} - x_i\| \leq \sum_{i=0}^k (t_{i+1} - t_i) = t_{k+1} < t^*,
\]

which implies that \(x_{k+1} \in B(x_0, t^*)\). This together with (16) gives that (i) holds for the case \(k + 1\). For (ii), by (3) again, we have the following identity:

\[
F(x_{k+1}) = F(x_{k+1}) - F(y_k) - F'(x_k)(x_{k+1} - y_k)
\]

It follows from the \(L\)-average Lipschitz condition (14) that

\[
\|F'(x_0)^{-1}F(x_{k+1})\| \leq \int_0^1 \|F'(x_0)^{-1}[F'(y_k + \tau(x_{k+1} - y_k)) - F'(x_k)]\|\|x_{k+1} - y_k\| \, d\tau
\]

In view of \(h'\) is increasing and convex in \([0, r_0]\), by applying Lemma 1 and the inductive hypotheses (iii)–(iv), one has that

\[
\int_{\|x_k - x_0\| + \|y_k - x_k + \tau(x_{k+1} - y_k)\|}^{\|x_k - x_0\| + \|y_k - x_k + \tau(x_{k+1} - y_k)\|} L(u) \, du
\]

\[
= h'(\|x_k - x_0\| + \|y_k - x_k + \tau(x_{k+1} - y_k)\|) - h'(\|x_k - x_0\|)
\]

\[
\leq h'(\|x_k - x_0\| + \|y_k - x_k\| + \|x_{k+1} - y_k\|) - h'(\|x_k - x_0\|)
\]

\[
\leq \frac{h'(s_k + \tau(t_{k+1} - s_k)) - h'(t_k)}{s_k - t_k + \tau(s_k - t_k)} \cdot (\|y_k - x_k\| + \|x_{k+1} - y_k\|)
\]

\[
\leq h'(s_k + \tau(t_{k+1} - s_k)) - h'(t_k).
\]
This allows us to get
\[
\| F'(x_0)^{-1} F(x_{k+1}) \| \leq \int_{0}^{1} \left[ h'(s_k + \tau(t_{k+1} - s_k)) - h'(t_k) \right] \cdot \frac{\| x_{k+1} - y_k \|}{t_{k+1} - s_k} \, d\tau
\]
\[
= h(t_{k+1}) - h(s_k) - h'(t_k)(t_{k+1} - s_k) \cdot \frac{\| x_{k+1} - y_k \|}{t_{k+1} - s_k}
\]
\[
\leq h(t_{k+1}) \cdot \frac{\| x_{k+1} - y_k \|}{t_{k+1} - s_k}
\]
\[
\leq h(t_{k+1}),
\]
(17)
which shows that (ii) holds for the case \( k + 1 \). Combining (16) and (17), we further obtain that
\[
\| y_{k+1} - x_{k+1} \| = \| F'(x_{k+1})^{-1} F(x_{k+1}) \|
\leq \| F'(x_{k+1})^{-1} F'(x_0) \| \cdot \| F'(x_{k+1})^{-1} F(x_{k+1}) \|
\leq -\frac{h(t_{k+1})}{h'(t_{k+1})} = s_{k+1} - t_{k+1}.
\]
(18)
This means that (iii) holds for the case \( k + 1 \). Then, we conclude that \( \| y_{k+1} - x_0 \| \leq \| y_{k+1} - x_{k+1} \| + \| x_{k+1} - x_0 \| \leq s_{k+1} - t^* \) and so \( y_{k+1} \in B(x_0, t^*) \). As for (iv), noting that
\[
x_{k+2} - y_{k+1} = -F'(x_{k+1})^{-1} F(y_{k+1})
\]
\[
= -F'(x_{k+1})^{-1} \left[ F(y_{k+1}) - F(x_{k+1}) - F'(x_{k+1})(y_{k+1} - x_{k+1}) \right]
\]
\[
= -F'(x_{k+1})^{-1} \int_{0}^{1} \left[ F'(x_{k+1}^{\tau}) - F'(x_{k+1}) \right](y_{k+1} - x_{k+1}) \, d\tau,
\]
where \( x_{k+1}^{\tau} := x_{k+1} + \tau(y_{k+1} - x_{k+1}) \), by using (16), the \( L \)-average Lipschitz condition (14), we have
\[
\| x_{k+2} - y_{k+1} \| \leq -\frac{1}{h'(t_{k+1})} \int_{0}^{1} \left[ \int_{\| x_{k+1} - x_0 \|}^{\| y_{k+1} - x_{k+1} \| + \| y_{k+1} - x_{k+1} \|} L(u) \, du \right] \, \| y_{k+1} - x_{k+1} \| \, d\tau.
\]
Taking into account that \( h' \) is increasing and convex in \([0, r_0]\) again, by combining (18) with Lemma 1, one gets that
\[
\int_{\| x_{k+1} - x_0 \|}^{\| y_{k+1} - x_{k+1} \| + \| y_{k+1} - x_{k+1} \|} L(u) \, du
\]
\[
= h'([\| x_{k+1} - x_0 \| + \| y_{k+1} - x_{k+1} \|]) - h'([\| x_{k+1} - x_0 \|])
\]
\[
\leq -\frac{h'(t_{k+1} + \tau(s_{k+1} - t_{k+1})) - h'(t_{k+1})}{s_{k+1} - t_{k+1}} \cdot \| y_{k+1} - x_{k+1} \|.
\]
This permits us to arrive at

\[ \|x_{k+2} - y_{k+1}\| \leq -\frac{1}{h'(t_{k+1})} \int_0^1 \left[ h'(t_{k+1} + \tau(s_{k+1} - t_{k+1})) - h'(t_{k+1}) \right] \cdot \frac{\|y_{k+1} - x_{k+1}\|^2}{s_{k+1} - t_{k+1}} \, d\tau \]

\[ = -\frac{1}{h'(t_{k+1})} \left[ h(s_{k+1}) - h(t_{k+1}) - h'(t_{k+1})(s_{k+1} - t_{k+1}) \right] \cdot \left( \frac{\|y_{k+1} - x_{k+1}\|^2}{s_{k+1} - t_{k+1}} \right)^2 \]

\[ = (t_{k+2} - s_{k+1}) \cdot \left( \frac{\|y_{k+1} - x_{k+1}\|^2}{s_{k+1} - t_{k+1}} \right)^2 . \]

Furthermore, we derive from this together with (18) that

\[ \|x_{k+2} - x_{k+1}\| \leq \|x_{k+2} - y_{k+1}\| + \|y_{k+1} - x_{k+1}\| \leq t_{k+2} - t_{k+1} . \]

Therefore, all the statements in the lemma hold by induction. This completes the proof.

Lemma 5 Under the same assumptions of Lemma 4. Then, the sequence \(\{x_k\}\) converges to a point \(x^* \in B(x_0, t^*)\) with \(F(x^*) = 0\). Moreover, we have

\[ \|x^* - x_k\| \leq t^* - t_k, \quad k \geq 0 , \quad (19) \]

and

\[ \|x^* - y_k\| \leq (t^* - s_k) \left( \frac{\|x^* - x_k\|}{t^* - t_k} \right)^2 , \quad k \geq 0 . \quad (20) \]

Proof We apply Lemma 4 (v) and (11) to obtain that

\[ \sum_{k=N}^{\infty} \|x_{k+1} - x_k\| \leq \sum_{k=N}^{\infty} (t_{k+1} - t_k) = t^* - t_N < +\infty \quad \text{for any} \quad N \in \mathbb{N} . \]

Thus, \(\{x_k\}\) is a Cauchy sequence in \(B(x_0, t^*)\) and so converges to some \(x^* \in B(x_0, t^*)\). For any \(k \geq 0\), the above inequality also means that \(\|x^* - x_k\| \leq t^* - t_k\).

Next, we show that \(F(x^*) = 0\). Thanks to Lemma 3, one has that \(\{\|F'(x_k)\|\}\) is bounded. Then, it follows from Lemma 4 that

\[ \|F(x_k)\| \leq \|F'(x_k)\| \|F'(x_k)^{-1} F(x_k)\| \leq \|F'(x_k)\| (s_k - t_k) . \]

Letting \(k \to \infty\), by noting the fact that \(\{s_k\}\) and \(\{t_k\}\) are converge to the same point \(t^*\) (see Remark 1), we get that \(\lim_{k \to \infty} F(x_k) = 0\). Since \(F\) is continuous in \(B(x_0, t^*)\), \(\{x_k\} \subset B(x_0, t^*)\) and \(\{x_k\}\) converges to \(x^*\), we obtain \(\lim_{k \to \infty} F(x_k) = F(x^*)\), which verifies that \(F(x^*) = 0\). It remains to show the estimate (20). Due to Lemma 4, we have

\[ \|y_k - x_0\| \leq \|y_k - x_k\| + \|x_k - x_0\| \leq s_k . \quad (21) \]

On the other hand, we can derive the following identity:

\[ x^* - y_k = -F'(x_k)^{-1} \int_0^1 \left[ F'(x_k + \tau(x^* - x_k)) - F'(x_k) \right](x^* - x_k) \, d\tau . \]
Then, in view of $h'$ is increasing and convex in $[0, r_0)$, by combining (16), the $L$-average Lipschitz condition (14) and Lemma 1, one gets that

$$\|x^* - y_k\| \leq - \frac{1}{h'(t_k)} \int_0^1 \left( \int_{\|x_k - x_0\|}^{\|x_k - x_0\| + \tau \|x^* - x_k\|} L(u) \, du \right) \|x^* - x_k\| \, d\tau$$

$$= - \frac{1}{h'(t_k)} \int_0^1 \left[ h'(\|x_k - x\| + \tau \|x^* - x_k\|) - h'(\|x_k - x_0\|) \right] \|x^* - x_k\| \, d\tau$$

$$\leq - \frac{1}{h'(t_k)} \int_0^1 \frac{h'(t_k + \tau(t^* - t_k)) - h'(t_k)}{t^* - t_k} \, d\tau \cdot \|x^* - x_k\|^2$$

$$= (t^* - s_k) \cdot \left( \frac{\|x^* - x_k\|}{t^* - t_k} \right)^2,$$

as claimed. The proof of this lemma is complete. \(\square\)

**Lemma 6** Under the same assumptions of Lemma 4 and the assumption that $2 + t^* h''(t^*) / h'(t^*) > 0$, we have

$$\frac{\|y_k - x_k\|}{s_k - t_k} \leq \frac{1}{1 + \frac{h''(t^*)}{2h'(t^*)}} \left( t^* - s_k \right) \cdot \frac{\|x^* - x_k\|}{t^* - t_k}, \quad k \geq 0. \quad (22)$$

**Proof** By using (10), we deduce that

$$t^* - s_k = - \frac{1}{h'(t_k)} \int_0^1 \left[ h'(t_k + \tau(t^* - t_k)) - h'(t_k) \right] (t^* - t_k) \, d\tau.$$

Taking into account the convexity of $h'$ in $[0, r_0)$, it follows from Lemma 1 that, for any $\tau \in (0, 1]$,

$$h'(t_k + \tau(t^* - t_k)) - h'(t_k) = \frac{h'(t_k + \tau(t^* - t_k)) - h'(t_k)}{\tau (t^* - t_k)} \cdot \tau (t^* - t_k)$$

$$\leq \frac{h'(t^*) - h'(t_k)}{t^* - t_k} \cdot \tau (t^* - t_k).$$

Then, in view of the positivity of $-1/h'(t)$, one has from Lemma 1 again that

$$t^* - s_k \leq - \frac{1}{h'(t_k)} \int_0^1 \frac{h'(t^*) - h'(t_k)}{t^* - t_k} \tau \, d\tau \cdot (t^* - t_k)^2 \leq - \frac{1}{2 h'(t^*)} \cdot (t^* - t_k)^2. \quad (23)$$

the last due to $h'$ is strictly increasing. Since $\|y_k - x_k\| \leq \|x^* - y_k\| + \|x^* - x_k\|$, it follows from (19) and (20) that

$$\|y_k - x_k\| \leq \frac{t^* - s_k}{(t^* - t_k)^2} \|x^* - x_k\|^2 + \|x^* - x_k\|$$

$$\leq \left( \frac{t^* - s_k}{t^* - t_k} + 1 \right) \|x^* - x_k\|.$$

Then, by (23), we can get further that

$$\|y_k - x_k\| \leq \left( 1 - \frac{h''(t^*)}{2h'(t^*)} (t^* - t_k) \right) \|x^* - x_k\|.$$

© Springer
Thus, we conclude from (13) that
\[
\frac{\|y_k - x_k\|}{s_k - t_k} \leq \frac{1}{2} \frac{h''(t^*)}{h(t^*)} (t^* - t_k) \left\| x^* - x_k \right\|
\]
\[
\leq \frac{1}{1 + \frac{h''(t^*)}{2h(t^*)} (t^* - t_k)} \cdot \frac{\|x^*-x_k\|}{t^*-t_k},
\]
which yields the desired result.

### 3.2 Convergence results

Our main semilocal convergence result of this paper is presented in the following theorem, in which we provide Q-superquadratic and Q-cubic convergence criterion of the two-step Newton method (3) under \(L\)-average Lipschitz condition (14) in Banach spaces. Then, we obtain three important special cases from this main result. They include the Kantorovich-type convergence result under the Lipschitz condition, Smale-type convergence result for analytical operators and convergence result for self-concordant functions under the Nesterov-Nemirovskii condition.

**Theorem 1** Let \(F : \mathbb{D} \subset \mathbb{X} \rightarrow \mathbb{Y}\) be a continuously Fréchet differentiable nonlinear operator in open convex subset \(\mathbb{D}\). Assume that there exists an initial guess \(x_0 \in \mathbb{D}\), such that \(F'(x_0)^{-1}\) exists and that \(F'\) satisfies the \(L\)-average Lipschitz condition (14) in \(B(x_0, t^*)\). Let \(\{x_k\}\) be the iterates generated by the two-step Newton method (3) with initial guess \(x_0\). If \(0 < \beta \leq b\), then \(\{x_k\}\) is well defined and converges Q-superquadratically to a solution \(x^* \in B(x_0, t^*)\) of (1), and this solution \(x^*\) is unique on \(B(x_0, r)\), where \(t^* \leq r < t^{**}\). Moreover, if

\[
2 + \frac{t^* h''(t^*)}{h(t^*)} \geq \frac{t^* L(t^*)}{1 - \int_0^{t^*} L(u) \, du} > 0,
\]

then the order of convergence is cubic at least and we have the following error bounds

\[
\left\| x^* - x_{k+1} \right\| \leq \frac{1}{2} H_*^2 \cdot \frac{2 - t^* H_*}{2 + t^* H_*} \cdot \left\| x^* - x_k \right\|^3, \quad k \geq 0,
\]

where \(H_* \triangleq h''(t^*)/h'(t^*)\).

**Proof** Thanks to Lemma 4, we conclude that the sequence \(\{x_k\}\) is well defined. By using Lemma 4 (v) and (11), one has that \(\|x_k - x_0\| \leq t_k < t^*\) for any \(k \geq 0\), which implies that \(\{x_k\}\) is contained in the ball \(B(x_0, t^*)\). Moreover, it follows from Lemma 5 that \(\{x_k\}\) converges to \(x^*\), a solution of (1) in \(B(x_0, t^*)\). Next, we will verify the superquadratic and cubic convergence of the iterates. To do this, we apply standard analytical techniques to derive that

\[
x^* - x_{k+1} = x^* - y_k + F'(x_k)^{-1} F(y_k)
\]
\[
= -F'(x_k)^{-1} [F(x^*) - F(y_k) - F'(y_k)(x^* - y_k) + (F'(y_k) - F'(x_k))(x^* - y_k)]
\]
\[
= -F'(x_k)^{-1} \left[ \int_0^{t_k} \left( F'(y_k^*) - F'(y_k) \right) (x^* - y_k) \, d\tau + (F'(y_k) - F'(x_k))(x^* - y_k) \right],
\]
where $y_k^r := y_k + \tau (x^* - y_k)$. By (16) and the $L$-average Lipschitz condition (14), we have

$$
\|x^* - x_{k+1}\| \leq -\frac{1}{h'(t_k)} \left[ \int_0^1 \left( \int_{\|y_0-x_0\|+\tau \|x^*-y_k\|} L(u) \, du \right) \|x^* - y_k\| \, d\tau \right]
+ \int_{\|y_k-x_0\|+\tau \|y_k-x\|} L(u) \, du \cdot \|x^* - y_k\|.
$$

Taking into account $h'$ is increasing and convex in $[0, r_0)$, combining (20), (21), Lemma 1 and Lemma 4 (iii), one can deduce that

$$
\|x^* - x_{k+1}\| \leq -\frac{1}{h'(t_k)} \left[ \int_0^1 \frac{h'(s_k + \tau (t^* - s_k)) - h'(s_k)}{t^* - s_k} \, d\tau \cdot \|x^* - y_k\|^2 
+ \frac{h'(s_k) - h'(t_k)}{t^* - t_k} \cdot \|y_k - x_k\| \|x^* - y_k\| \right]
= -\frac{1}{h'(t_k)} \left[ \left( h(t^*) - h(s_k) - h'(s_k)(t^* - s_k) \right) \left( \frac{\|x^* - y_k\|}{t^* - s_k} \right)^2 
+ \left( h'(s_k) - h'(t_k) \right) (t^* - s_k) \cdot \frac{\|y_k - x_k\| \|x^* - y_k\|}{s_k - t_k} \right].
$$

By Lemma 4 (iii) and (20) again, the above inequality can be derive further that

$$
\|x^* - x_{k+1}\| \leq (t^* - t_{k+1}) \left( \frac{\|x^* - x_k\|}{t^* - t_k} \right)^2.
$$

Then, it follows from (12) that

$$
\frac{\|x^* - x_{k+1}\|}{\|x^* - x_k\|^2} \leq \frac{t^* - t_{k+1}}{(t^* - t_k)^2} \leq \frac{1}{2} \left( \frac{h''(t^*)}{h'(t^*)} \right)^2 (t^* - t_k).
$$

Letting $k \to \infty$ in the above inequalities, by noting that $\{t_k\}$ converges to $t^*$, we have

$$
\lim_{k \to \infty} \frac{\|x^* - x_{k+1}\|}{\|x^* - x_k\|^2} = 0,
$$

which means that $\{x_k\}$ converges Q-superquadratically to $x^*$ (See Definition 2 for the definition). In addition, if the condition (24) is also satisfied, then the estimates (20), (22) and (12) are applicable to conclude from (26) further that

$$
\|x^* - x_{k+1}\| \leq -\frac{1}{h'(t_k)} \left[ \left( h(t^*) - h(s_k) - h'(s_k)(t^* - s_k) \right) 
+ \left( h'(s_k) - h'(t_k) \right) (t^* - s_k) \right] \cdot \frac{1 - \frac{h''(t^*)}{2h'(t^*)} (t^* - t_k)}{1 + \frac{h''(t^*)}{2h'(t^*)} (t^* - t_k)} \cdot \left( \frac{\|x^* - x_k\|}{t^* - t_k} \right)^3
\leq \left( \frac{1}{2} \left( \frac{h''(t^*)}{h'(t^*)} \right)^2 \right)^2 \cdot \frac{1 - \frac{r^* h''(t^*)}{2h'(t^*)} (t^* - t_k)}{1 + \frac{r^* h''(t^*)}{2h'(t^*)} (t^* - t_k)} \cdot \|x^* - x_k\|^3.
$$

This shows the estimate (25) in Theorem 1 and so the order of convergence for the iterates is Q-cubic.

Finally, we show the uniqueness of the solution. We first to prove the solution $x^*$ of (1) is unique on $B(x_0, t^*)$. Assume that there exists another solution $x^{**}$ on
\( \mathbb{B}(x_0, t^*) \). Then, one has that \( \|x^{**} - x_0\| \leq t^* \). Next, we will show by induction that
\[
\|x^{**} - x_k\| \leq t^* - t_k, \quad k = 0, 1, 2, \ldots.
\] (28)
Since \( t_0 = 0 \), the case \( k = 0 \) holds obviously. We suppose that the above inequality holds for some \( k \geq 0 \). As the same process on the estimate \( \|x^* - y_k\| \) in (20), we get
\[
\|x^{**} - y_k\| \leq (t^* - s_k) \left( \frac{\|x^{**} - x_k\|}{t^* - t_k} \right)^2.
\]
In addition, following the same process on the estimate \( \|x^* - x_{k+1}\| \) in (27), one has that
\[
\|x^{**} - x_{k+1}\| \leq (t^* - t_{k+1}) \cdot \left( \frac{\|x^{**} - x_k\|}{t^* - t_k} \right)^2.
\]
Then, by applying the inductive assumption (28) to the above inequality, we have that (28) also holds for the case \( k + 1 \). Since \( \{x_k\} \) converges to \( x^* \) and \( \{t_k\} \) converges to \( t^* \), we conclude from (28) that \( x^{**} = x^* \). Hence, \( x^* \) is the unique zero of (1) on \( \mathbb{B}(x_0, r) \). Suppose for contradiction that \( F \) has at least one zero there. That is to say, there exists \( x^{**} \in D \subset \mathbb{X} \) such that \( t^* < \|x^{**} - x_0\| < r \) and \( F(x^{**}) = 0 \). We will prove that the preceding hypotheses do not hold. Since
\[
F(x^{**}) = F(x_0) + F'(x_0)(x^{**} - x_0) + \int_0^1 [F'(x_0) - F'(x_0)](x^{**} - x_0) \, d\tau,
\] (29)
where \( x_0^\tau := x_0 + \tau (x^{**} - x_0) \). Note that,
\[
\|F'(x_0)^{-1}[F(x_0) + F'(x_0)(x^{**} - x_0)]\| \geq \|x^{**} - x_0\| - \|F'(x_0)^{-1}F(x_0)\|
= \|x^{**} - x_0\| - h(0).
\]
In addition, we use the \( L \)-average Lipschitz condition (14) to yield
\[
\left\| F'(x_0)^{-1} \int_0^1 [F'(x_0^\tau) - F'(x_0)](x^{**} - x_0) \, d\tau \right\|
\leq \int_0^1 \left( \int_0^1 \|x^{**} - x_0\| \, L(u) \, du \right) \|x^{**} - x_0\| \, du
\]
\[
= \int_0^1 \left( \int_0^1 [h'(\|x^{**} - x_0\|) - h'(0)] \|x^{**} - x_0\| \, d\tau \right)
= h(\|x^{**} - x_0\|) - h(0) - h'(0) \cdot \|x^{**} - x_0\|.
\]
In view of \( F(x^{**}) = 0 \) and \( h'(0) = -1 \), it follows from (29) that
\[
h(\|x^{**} - x_0\|) - h(0) - h'(0) \cdot \|x^{**} - x_0\| \geq \|x^{**} - x_0\| - h(0),
\]
which is equivalent to \( h(\|x^{**} - x_0\|) \geq 0 \). We then obtain from Lemma 2 that \( h \) is strictly positive in the interval \( (\|x^{**} - x_0\|, R) \). Thus, we know \( r \leq \|x^{**} - x_0\| \), which is a contradiction to the preceding hypotheses. Consequently, \( F \) does not have zeros in \( \mathbb{B}(x_0, r) \) and \( x^* \) is the unique zero of (1) in \( \mathbb{B}(x_0, r) \). The proof is complete. \( \square \)

**Remark 2** The convergence criterion \( 0 < \beta \leq b \) given in Theorem 1 was obtained by Wang in [63] for studying the quadratic convergence of Newton’s method (2)
under a unified framework. As is stated in Theorem 1, this criterion guarantees only superquadratic convergence for the two-step Newton method (3). To obtain cubic convergence, we also need the condition (24).

In what follows, based on Theorem 1, we will obtain some corollaries by taking various forms of the positive function \( L \).

Firstly, for the case when \( L \) is a positive constant function, then the \( L \)-average Lipschitz condition (14) reduces to the following affine-invariant Lipschitz condition:

\[
\| F'(x_0)^{-1}[F'(y) - F'(x)] \| \leq L \| y - x \|, \quad x, y \in B(x_0, r_0),
\]

where \( r_0 = 1/L \) due to (6). The majorizing function \( h \) defined by (5) now has the form below:

\[
h(t) = \beta - t + \frac{L}{2} t^2, \quad t \in [0, R],
\]

where \( R = 2/L \) by (4). The constant \( b \) defined in (7) reduces to \( b = 1/(2L) \). Moreover, thanks to Lemma 2, if \( L\beta < 1/2 \), then the zeros of \( h \) in \((0, 1/L) \) and \((1/L, 2/L) \) are

\[
t^* = 1 - \sqrt{1 - 2L\beta} \quad \text{and} \quad t^{**} = 1 + \sqrt{1 - 2L\beta},
\]

respectively. Therefore, we have the following Kantorovich-type convergence result from Theorem 1 for two-step Newton method (3) under Lipschitz condition (30).

**Corollary 1** Let \( F : D \subset X \rightarrow Y \) be a continuously Fréchet differentiable nonlinear operator in open convex subset \( D \). Assume that there exists an initial guess \( x_0 \in D \) such that \( F'(x_0)^{-1} \) exists and that \( F' \) satisfies the Lipschitz condition (30). Let \( \{x_k\} \) be the iterates generated by the two-step Newton method (3) with initial guess \( x_0 \). If \( 0 < L\beta \leq 1/2 \), then \( \{x_k\} \) is well defined and converges \( Q \)-superquadratically to a solution \( x^* \in B(x_0, t^*) \) of (1), and this solution \( x^* \) is unique on \( \bar{B}(x_0, r) \), where \( t^* \leq r < t^{**} \), \( t^* \) and \( t^{**} \) are given in (31). Moreover, if \( 0 \leq L\beta < 4/9 \), then the order of convergence is cubic at least and we have the following error bounds

\[
\| x^* - x_{k+1} \| \leq \frac{L^2}{2(1 - 2L\beta)} \cdot \frac{\sqrt{1 - 2L\beta} + 1}{3\sqrt{1 - 2L\beta} - 1} \cdot \| x^* - x_k \|^3, \quad k \geq 0.
\]

**Remark 3** The convergence result in Corollary 1 is obtained under the weaker assumption on the Lipschitz condition in comparison with the one presented in [2], where the assumption that the second derivative \( F'' \) satisfies the Lipschitz condition is needed. To obtain the \( Q \)-cubic convergence, we need the convergence criterion \( 0 < L\beta < 4/9 \), which has slightly stronger than the usual one \( 0 < L\beta < 1/2 \) for ensuring the quadratic convergence of Newton’s method (2).

Secondly, we suppose that \( \gamma > 0 \). Let \( L \) be the positive function defined by

\[
L(u) := \frac{2\gamma}{(1 - \gamma u)^3}, \quad u \in [0, \frac{1}{\gamma}).
\]
Then, the $L$-average Lipschitz condition (14) reduces to
\[
\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \leq \frac{1}{(1 - \gamma \|x - x_0\| - \|y - x\|)^2} - \frac{1}{(1 - \gamma \|x - x_0\|)^2}.
\]
(32)
The majorizing function $h$ defined by (5) reduces to
\[
h(t) = \beta - t + \frac{\gamma t^2}{1 - \gamma t}, \quad t \in [0, \frac{1}{\gamma}).
\]
(33)
The constants $r_0$ and $b$ defined in (6) and (7) are given by
\[
r_0 = 1 - \frac{1}{\sqrt{2}} \frac{1}{\gamma} \quad \text{and} \quad b = (3 - 2\sqrt{2}) \frac{1}{\gamma},
\]
respectively. If $\alpha := \beta \gamma \leq 3 - 2\sqrt{2}$, then the zeros of $h$ are
\[
t^* = \frac{1 + \alpha - \sqrt{(1 + \alpha)^2 - 8\alpha}}{4\gamma} \quad \text{and} \quad t^{**} = \frac{1 + \alpha + \sqrt{(1 + \alpha)^2 - 8\alpha}}{4\gamma},
\]
(34)
respectively. Then, the constant $H_* := h''(t^*)/h'(t^*)$ given in Theorem 1 now has the following concrete form:
\[
H_* = -\frac{32\gamma}{\sqrt{(1 + \alpha)^2 - 8\alpha} \cdot (3 - \alpha + \sqrt{(1 + \alpha)^2 - 8\alpha})^2}.
\]
(35)
Consequently, we have the following Smale-type convergence result from Theorem 1 for two-step Newton method (3) under the condition (32).

**Corollary 2** Let $F : \mathbb{D} \subset \mathbb{X} \rightarrow \mathbb{Y}$ be a continuously Fréchet differentiable nonlinear operator in open convex subset $\mathbb{D}$. Assume that there exists an initial guess $x_0 \in \mathbb{D}$ such that $F'(x_0)^{-1}$ exists and that $F'$ satisfies the condition (32). Let $\{x_k\}$ be the iterates generated by the two-step Newton method (3) with initial guess $x_0$. If $0 < \alpha \leq 3 - 2\sqrt{2}$, then $\{x_k\}$ is well defined and converges $Q$-superquadratically to a solution $x^* \in \overline{B}(x_0, t^*)$ of (1), and this solution $x^*$ is unique on $\overline{B}(x_0, r)$, where $t^* \leq r < t^{**}$, $t^*$ and $t^{**}$ are given in (34). Moreover, if $0 < \alpha < 3 - \sqrt{2} - \sqrt{4}$, then the order of convergence is cubic at least and we have the following error bounds
\[
\|x^* - x_{k+1}\| \leq \frac{q H_*^2}{2} \cdot \|x^* - x_k\|^3, \quad k \geq 0,
\]
(36)
where $H_*$ is given in (35) and
\[
q := \frac{\sqrt{(1 + \alpha)^2 - 8\alpha} \cdot (3 - \alpha + \sqrt{(1 + \alpha)^2 - 8\alpha})^2 + 4(1 + \alpha - \sqrt{(1 + \alpha)^2 - 8\alpha})}{\sqrt{(1 + \alpha)^2 - 8\alpha} \cdot (3 - \alpha + \sqrt{(1 + \alpha)^2 - 8\alpha})^2 - 4(1 + \alpha - \sqrt{(1 + \alpha)^2 - 8\alpha})}.
\]

**Remark 4** If $F$ is twice continuously Fréchet differentiable, then $F'$ satisfies the condition (32) if and only if $F$ satisfies the following condition
\[
\|F'(x_0)^{-1}F''(x)\| \leq \frac{2\gamma}{(1 - \gamma \|x - x_0\|)^2}, \quad x \in B(x_0, 1/\gamma).
\]
(37)
In fact, if $F$ satisfies (32), then (37) is trivially valid. Conversely, if $F$ satisfies (37), then by noting that $h'(t) = -2 + 1/(1 - \gamma t)^2$ and $h''(t) = 2\gamma/(1 - \gamma t)^3$, we have

$$
\|F'(x_0)^{-1}[F'(y) - F'(x)]\| \leq \int_0^1 \|F'(x_0)^{-1}F''(x + \tau(y - x))\||y - x||\,d\tau
\leq \int_0^1 h''(\|x - x_0\| + \|y - x\|)\|y - x\||\,d\tau
= h'(\|x - x_0\| + \|y - x\|) - h'(\|x - x_0\|),
$$

which means that $F$ satisfies (32). The condition (37) is called the $\gamma$-condition which was introduced by Wang and Han in [65] to study the Smale point estimate theory. Based on the above observation, the convergence result stated in Corollary 2 also holds when the condition (32) is replaced by the $\gamma$-condition (37).

One important and typical class of examples satisfying the $\gamma$-condition (37) is the one of analytic operators. Smale [59] studied the convergence and error estimate of Newton’s method (2) under the hypotheses that $F$ is analytic and satisfies

$$
\|F'(x_0)^{-1}F^{(n)}(x_0)\| \leq n!\gamma^{n-1}, \quad n \geq 2,
$$

where $\gamma$ is given by

$$
\gamma := \sup_{k>1} \left\| \frac{F'(x_0)^{-1}F^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}.
\tag{38}
$$

We then obtain from Theorem 1 that another Smale-type convergence result of the two-step Newton method (2) for the analytic operator.

**Corollary 3** Let $F : D \subset X \rightarrow Y$ be an analytic operator in open convex subset $D$. Assume that there exists an initial guess $x_0 \in D$ such that $F'(x_0)$ is nonsingular. Let $\{x_k\}$ be the iterates generated by the two-step Newton method (3) with initial guess $x_0$. If $0 < \alpha := \beta \gamma \leq 3 - 2\sqrt{2}$, where $\gamma$ is given by (38), then $\{x_k\}$ is well defined and converges Q-superquadratically to a solution $x^* \in B(x_0, r^*)$ of (1), and this solution $x^*$ is unique on $B(x_0, r)$, where $t^* \leq r < t^{**}$, $t^*$ and $t^{**}$ are given in (34). Moreover, the order of convergence is cubic at least and the error estimate (36) holds when $0 < \alpha < 3 - \frac{3}{2} - \frac{3}{4}$.

Lastly, we present a semilocal convergence result from Theorem 1 for the two-step Newton method (3) under the condition introduced by Nesterov and Nemirovskii in the seminal work [56].

Let $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ be a strictly convex and three times continuously differentiable function in open convex subset $D$. We denote by $f'(x)$ the gradient of $f$ at $x$, and $f''(x)$ the Hessian matrix of $f$ at $x$. Let $x_0 \in D$ be an initial guess such that the inverse $f''(x_0)^{-1}$ exists. For a given constant $a > 0$, if $f$ satisfies the inequality

$$
|f''(x)[u, u, u]| \leq 2a^{-1/2} \left(f''(x)[u, u]\right)^{3/2}, \quad x \in D, u \in \mathbb{R}^n,
$$

then we say that $f$ is $a$-self-concordant [56]. Here,

$$
f''(x)[u, u, u] := (u^TH_1(x)u, u^TH_2(x)u, \ldots, u^TH_n(x)u)u,
$$

\(\copyright\) Springer
where $H_i(x) (i = 1, 2, \ldots, n)$ is the Hessian matrix of the $i$th component of $f'$ at $x$. For $x \in \mathbb{D}$, we set

$$
\langle u_1, u_2 \rangle_x := a^{-1} \{ f''(x) u_1, u_2 \}, \quad \forall u_1, u_2 \in \mathbb{R}^n,
$$

and define the norm

$$
\| u \|_x := \sqrt{\langle u, u \rangle_x}, \quad \forall u \in \mathbb{R}^n.
$$

Then, $(\mathbb{R}^n, \| \cdot \|_x)$ is a Banach space.

Choose $L(u) = 2/(1-u)^3$ in (5). We obtain the following majorizing function

$$
h(t) = \beta - t + t^2 \frac{1}{1-t}, \quad t \in [0, 1),
$$

which is a special case of the one given in (33); i.e., $\gamma \equiv 1$. Then, there exist two zeros for this majorizing function when $\beta \leq 3 - 2\sqrt{2}$. These two zeros are as follows:

$$
t^* = \frac{1 + \beta - \sqrt{(1 + \beta)^2 - 8\beta}}{4} \quad \text{and} \quad t^{**} = \frac{1 + \beta + \sqrt{(1 + \beta)^2 - 8\beta}}{4}.
$$

For a given vector $x \in \mathbb{R}^n$ and a positive number $r$, we set

$$
B_r(x) := \{ y \in \mathbb{R}^n : \| y - x \|_x < r \} \quad \text{and} \quad \overline{B}_r(x) := \{ y \in \mathbb{R}^n : \| y - x \|_x \leq r \}.
$$

They correspond to the open ball and its closure of center $x$ and radius $r$ when $\mathbb{R}^n$ is endowed with the metric structure induced by the preceding inner product $(\cdot, \cdot)_x$. If $f$ is an $a$-self-concordant function, then we have (see [1, Lemma 5.1])

$$
\| f''(x_0)^{-1} f'''(x) \|_{x_0} \leq \frac{2}{(1 - \| x - x_0 \|_{x_0})^3}, \quad x \in B_1(x_0).
$$

This implies that $f'''$ satisfies the $L$-average Lipschitz condition (14) with $L(u) = 2/(1-u)^3$. Let $X = Y = (\mathbb{R}^n, \| \cdot \|_x)$. Then, by Theorem 1, we have the following semilocal convergence result about the minimization of $a$-self-concordant function for two-step Newton method which defined by

$$
\begin{align*}
\left\{ \begin{array}{l}
y_k = x_k - f''(x_k)^{-1} f'(x_k), \\
x_{k+1} = y_k - f''(x_k)^{-1} f'(y_k),
\end{array} \right. \quad k = 0, 1, 2, \ldots.
\end{align*}
$$

**Corollary 4** Let $f : \mathbb{D} \subset \mathbb{R}^n \to \mathbb{R}$ be an $a$-self-concordant function in open convex subset $\mathbb{D}$. Assume that there exists an initial guess $x_0 \in \mathbb{D}$ such that $f''(x_0)$ is nonsingular. Let $\{x_k\}$ be the vector sequence generated by the two-step Newton method (40) for solving $f'(x) = 0$ with initial guess $x_0$. If $\beta := \| f''(x_0) f'(x_0) \|_{x_0} \leq 3 - 2\sqrt{2}$, then $\{x_k\}$ is well defined and converges $Q$-superquadratically to a point $x^*$ which is the minimizer of $f$ in $\overline{B}_{t^*}(x_0)$, and this minimizer $x^*$ is unique on $\overline{B}_r(x_0)$, where $t^* \leq r < t^{**}$, $t^*$ and $t^{**}$ are given in (39). Moreover, if $0 < \beta < 3 - 3\sqrt{2} - \sqrt{4}$, then the order of convergence is cubic at least and we have the following error bounds

$$
\| x^* - x_{k+1} \| \leq \frac{q H^2}{2} \cdot \| x^* - x_k \|^3, \quad k \geq 0,
$$

where

$$
H_\beta = \frac{32}{\sqrt{(1 + \beta)^2 - 8\beta \cdot (3 - \beta + \sqrt{(1 + \beta)^2 - 8\beta})^2}}
$$

$\square$ Springer
and

\[ q = \frac{\sqrt{(1+\beta)^2 - 8\beta} \cdot (3 - \beta + \sqrt{(1+\beta)^2 - 8\beta})^2 + 4(1+\beta - \sqrt{(1+\beta)^2 - 8\beta})}{\sqrt{(1+\beta)^2 - 8\beta} \cdot (3 - \beta + \sqrt{(1+\beta)^2 - 8\beta})^2 - 4(1+\beta - \sqrt{(1+\beta)^2 - 8\beta})}. \]

Remark 5 Semilocal convergence result on the analysis of self-concordant minimization for Newton’s method \((2)\) has already been presented by Alvarez et al. in \([1]\). In addition, Ferreira and Svaiter \([28]\) provided another semilocal convergence result on self-concordant minimization for Newton’s method with a relative error tolerance.

4 Applications

By virtue of the semilocal convergence results established in the preceding section, this section is devoted to providing two applications. The first one concerns the minimal positive solution of a NSARE arising from transport theory. The second one is a two-dimensional nonlinear convection-diffusion equation, which is used to show the advantage of considering the \(L\)-average Lipschitz condition rather than the classical Lipschitz condition.

4.1 Algebraic Riccati equation

In this subsection, we apply the two-step Newton method \((3)\) to solve a special nonlinear vector equation which is obtained by a NSARE arising from transport theory. Throughout this subsection, we use the following definitions and notations. We call matrix \(A = (a_{ij})_{m \times n} \in \mathbb{R}^{m \times n}\) a positive matrix (nonnegative matrix) if \(a_{ij} > 0 (a_{ij} \geq 0)\) hold for all \(i = 1, 2, \ldots, m, j = 1, 2, \ldots, n\). If all the components of a vector are positive (negative), we call it a positive (negative) vector. For a given vector \(a\), we denote by \(\text{diag}(a)\) the diagonal matrix whose diagonal elements are the components of \(a\). We denote the vectors of all zeros and ones with proper dimension by \(0\) and \(e\), respectively. The norm of a vector or a matrix used in this section is \(\infty\)-norm.

The form of the NSARE is as follows:

\[ XCX - XD - AX + B = 0, \tag{41} \]

where \(A, B, C, D \in \mathbb{R}^{n \times n}\) are known matrices given by

\[ A = \Delta - e\mathbf{q}^T, \quad B = e\mathbf{e}^T, \quad C = \mathbf{q}\mathbf{q}^T, \quad D = \Gamma - \mathbf{q}\mathbf{e}^T, \tag{42} \]

with

\[
\begin{align*}
\Delta &= \text{diag}(\delta_1, \delta_2, \ldots, \delta_n) \text{ with } \delta_i = \frac{1}{c\omega_i (1 + \alpha)} > 0, \\
\Gamma &= \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_n) \text{ with } \gamma_i = \frac{1}{c\omega_i (1 - \alpha)} > 0, \\
\mathbf{q} &= (q_1, q_2, \ldots, q_n)^T \text{ with } q_i = \frac{c_i}{2\omega_i} > 0.
\end{align*}
\]
Here $c \in (0, 1]$ and $\alpha \in [0, 1)$. Moreover, $\{\omega_i\}_{i=1}^n$ and $\{c_i\}_{i=1}^n$ are the sets of the Gauss-Legendre nodes and weights, respectively, on the interval $[0, 1]$, and satisfy

$$0 < \omega_n < \cdots < \omega_2 < \omega_1 < 1 \quad \text{and} \quad \sum_{i=1}^n c_i = 1, c_i > 0, i = 1, 2, \ldots, n.$$  

The NSARE (41) has positive solutions (that is, the solution is a positive matrix), but only the minimal positive solution of it is physically meaningful [38].

Lu [51] first proved that the solution of NSARE (41) must have the following form:

$$X = T \circ (uv^T) = (uv^T) \circ T,$$

where $\circ$ denotes the Hadamard product, $T = (t_{ij})_{n \times n} = \left(\frac{1}{\delta_i + \gamma_j}\right)_{n \times n}$, $u$ and $v$ are vectors satisfying

$$\begin{cases}
    u = u \circ (Pv) + e, \\
v = v \circ (Pu) + e,
\end{cases} \quad (43)$$

with

$$P = (p_{ij})_{n \times n} = \left(\frac{q_j}{\delta_i + \gamma_j}\right)_{n \times n}, \quad \tilde{P} = (\tilde{p}_{ij})_{n \times n} = \left(\frac{q_j}{\gamma_i + \delta_j}\right)_{n \times n}. \quad (44)$$

Define nonlinear operator $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ by

$$f(u, v) = u^T - G(u, v), \quad (45)$$

Then, one can rewrite (43) as $f(u, v) = 0$. Hence, the minimal positive solution of NSARE (41) can be obtained via computing the minimal positive solution of the nonlinear vector (45). There have been a lot of studies about the monotone convergence of various iterative methods for solving the minimal positive solution of (41), one can see [7, 50, 51] and references therein.

Clearly, $f$ is a continuously Fréchet differentiable nonlinear operator in $\mathbb{R}^{2n}$. The Jacobian matrix of $f$ at point $(u, v)$ has the following form:

$$f^r(u, v) = I_{2n} - G(u, v), \quad (46)$$

with

$$G(u, v) = \begin{bmatrix} G_1(v) & H_1(u) \\ H_2(v) & G_2(u) \end{bmatrix}, \quad (47)$$

where $I_{2n}$ is the identity matrix of order $2n$,

$$\begin{cases}
    G_1(v) = \text{diag}(Pv), \\
    G_2(u) = \text{diag}(\tilde{P}u), \\
    H_1(u) = [u \circ p_1, u \circ p_2, \ldots, u \circ p_n], \\
    H_2(v) = [v \circ \tilde{p}_1, v \circ \tilde{p}_2, \ldots, v \circ \tilde{p}_n].
\end{cases}$$

$p_i$ and $\tilde{p}_i$ are the $i$th column of $P$ and $\tilde{P}$ for each $i = 1, 2, \ldots, n$, respectively. Choose initial guess $[u_0^T, v_0^T]^T = 0^T$. Then, we have $f(u_0, v_0) = -e$ and $f'(u_0, v_0) = I_{2n}$. Thus,

$$\beta := \|f'(u_0, v_0)^{-1}f(u_0, v_0)\|_\infty = \|e\|_\infty = 1.$$
Moreover, for any \((u, v)\) and \((u', v')\) \(\in \mathbb{R}^2\), it follows from (46) that

\[
\|f'(u_0, v_0)^{-1}[f'(u, v) - f'(u', v')]\|_\infty = \|G(u, v) - G(u', v')\|_\infty \\
\leq 2 \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^{n} p_{ij}, \sum_{j=1}^{n} \tilde{p}_{ij} \right\} \cdot \frac{\|u - u'\|_\infty}{\|v - v'\|_\infty}.
\]

In [51, Lemma 3], Lu derived that \(\sum_{j=1}^{n} p_{ij} < \frac{c(1-\alpha)}{2}\) and \(\sum_{j=1}^{n} \tilde{p}_{ij} < \frac{c(1+\alpha)}{2}\). By making use of these two estimates, we can obtain that

\[
\|f'(u_0, v_0)^{-1}[f'(u, v) - f'(u', v')]\|_\infty < c(1 + \alpha) \frac{\|u - u'\|_\infty}{\|v - v'\|_\infty}.
\]

That is, the Fréchet derivative of \(f\) satisfies the Lipschitz condition (30) with the Lipschitz constant \(L = c(1 + \alpha)\). Therefore, Corollary 1 is applicable to conclude that the iterative sequence generated by the two-step Newton method (3) for nonlinear operator \(f\) defined by (45) starting from the zero vector \(\theta\) converges Q-superquadratically to the minimal positive solution if

\[
L\beta = c(1 + \alpha) \leq \frac{1}{2}.
\]

Moreover, the order of convergence is cubic at least if \(L\beta = c(1 + \alpha) < \frac{4}{9}\). In particular, we can obtain that the minimal positive \(w_* := [u_*^T, v_*^T]^T\) belongs to the open ball with center \(\theta\) and radius \(r\), where

\[
\frac{1 - \sqrt{1 - 2c(1+\alpha)}}{c(1+\alpha)} \leq r < \frac{1 + \sqrt{1 - 2c(1+\alpha)}}{c(1+\alpha)}.
\]

That is, it satisfies \(0 < \|w_*\|_\infty \leq r\), which coincides with the one given in [7, Theorem 4.1].

We end this subsection with some numerical experiments illustrating the convergence results. The algorithm for implementing the two-step Newton method is summarized by Algorithm 1 as follows.
Algorithm 1 Two-step Newton method for solving (45).

Given $c \in (0, 1)$ and $\alpha \in [0, 1)$. Choose initial point $[u_0^T, v_0^T]^T = 0^T$. Form the matrices $P$ and $\bar{P}$ by (44). For $k = 0, 1, 2, \ldots$ until convergence, do:

Step 1. Form the matrix $G(u_k, v_k)$ by (47).

Step 2. Compute $\bar{v}_k$ from the system of linear equations below:

$$[I_n - G_2(u_k) - H_2(v_k)(I_n - G_1(v_k))^{-1}H_1(u_k)]\bar{v}_k$$
$$= H_2(v_k)(I_n - G_1(v_k))^{-1}[e - H_1(u_k)v_k] + e - H_2(v_k)u_k.$$  

Step 3. Compute $\tilde{u}_k$ from the following formula:

$$\tilde{u}_k = (I_n - G_1(v_k))^{-1}[H_1(u_k)(\bar{v}_k - v_k) + e].$$

Step 4. Compute $v_{k+1}$ from the system of linear equations below:

$$[I_n - G_2(u_k) - H_2(v_k)(I_n - G_1(v_k))^{-1}H_1(u_k)]v_{k+1}$$
$$= H_2(v_k)(I_n - G_1(v_k))^{-1}[\tilde{u}_k \circ P(\bar{v}_k - v_k) - H_1(u_k)\bar{v}_k + e]$$
$$+ \bar{v}_k \circ \bar{P}(\tilde{u}_k - u_k) - H_2(v_k)\tilde{u}_k + e.$$

Step 5. Compute $u_{k+1}$ from the following formula:

$$u_{k+1} = (I_n - G_1(v_k))^{-1}[\tilde{u}_k \circ P(\bar{v}_k - v_k) + e + H_1(u_k)(v_{k+1} - \bar{v}_k)].$$

As Example 5.2 of [30], the constants $c_i$ and $\omega_i$ are given by a numerical quadrature formula on the interval $[0, 1]$ which is obtained by dividing $[0, 1]$ into $n/4$ subintervals of equal length and applying Gauss-Legendre quadrature with four nodes to each subinterval. Our numerical experiments were carried out in MATLAB version R2014a running on a PC with Intel(R) Core(TM) i3-3110M of 2.40 GHz CPU and 12GB memory. In our implementations, the iterations in the algorithm are stopped when the following condition is satisfied:

$$\text{Res} := \max \left\{ \frac{\|u_{k+1} - u_k\|_\infty}{\|u_{k+1}\|_\infty}, \frac{\|v_{k+1} - v_k\|_\infty}{\|v_{k+1}\|_\infty} \right\} \leq \frac{\sqrt{n}}{2} \cdot \epsilon,$$

where $n$ the size of matrix $D$ given by (42) and $\epsilon = 2^{-52} \approx 2.2204 \times 10^{-16}$. The CPU time (in seconds) is computed by using the MATLAB function `cputime`. In each test, we run the same program 10 times and choose the average time as the time spent by the algorithm. Moreover, we use “iter” to stand for the number of the iterations needed.

Figure 1 depicts the iteration history for the problem size $n = 1024, 2048$ and 4096 with six different pairs of $(\alpha, c)$, respectively, namely, $(0.5, 1/3)$, $(0.5, 2/9)$, $(0.5, 1/9)$, $(0.25, 2/5)$, $(0.25, 1/3)$ and $(0.25, 1/10)$. As one can see in the figure,
Fig. 1  The iteration history for various $(\alpha, c)$ when the problem size $n = 1024, 2048, 4096$, respectively

for each case, the speedup is obtained as the value $L\beta = c(1 + \alpha)$ decreases. More convergence results including the number of iterations, the relative residual and the CPU time for various problem size $n$ are listed in Tables 1, 2 and 3, respectively. Obviously, we see from these tables that it requires less time when the value $L\beta = c(1 + \alpha)$ is taken smaller.
In conclusion, the above numerical experiments confirm our convergence results stated in Corollary 1 for the two-step Newton method (3).

4.2 Convection-diffusion equation

This subsection concerns a nontrivial example of functions \((F, L)\) satisfying the \(L\)-average Lipschitz condition given in Definition 1.

Consider the following two-dimensional nonlinear convection-diffusion equation

\[
\begin{aligned}
-(u_{xx} + u_{yy}) + q_1 u_x + q_2 u_y &= u^c \\
\quad u(x, y) &= 0 
\end{aligned} \quad \text{for } (x, y) \in \Omega, \\
\quad u(x, y) &= 0 \quad \text{for } (x, y) \in \partial \Omega,
\]

where \(\Omega = (0, 1) \times (0, 1)\), with \(\partial \Omega\) its boundary, \(q_1\) and \(q_2\) are positive constants used to measure the magnitudes of the convective terms, and \(c \in [1, 2]\). Applying the five-point finite-difference scheme to the diffusive terms and the central difference scheme to the convective terms, respectively, a system of nonlinear equations is

Table 1  The results for the problem size \(n = 1024\)

| \((\alpha, c)\)     | \(L^\beta\) | iter | Res         | CPU time (s) |
|---------------------|-------------|------|-------------|--------------|
| (0.5, 1/3)          | 1/2         | 5    | 3.8969e-16  | 7.6050       |
| (0.5, 2/9)          | 1/3         | 5    | 4.0858e-16  | 7.4600       |
| (0.5, 1/9)          | 1/6         | 4    | 1.9201e-15  | 5.9390       |
| (0.25, 2/5)         | 1/2         | 5    | 3.7210e-16  | 7.4178       |
| (0.25, 1/3)         | 5/12        | 5    | 3.8521e-16  | 7.3523       |
| (0.25, 1/10)        | 1/8         | 4    | 2.1370e-15  | 5.9062       |

Table 2  The results for the problem size \(n = 2048\)

| \((\alpha, c)\)     | \(L^\beta\) | iter | Res         | CPU time (s) |
|---------------------|-------------|------|-------------|--------------|
| (0.5, 1/3)          | 1/2         | 5    | 3.8969e-16  | 50.6707      |
| (0.5, 2/9)          | 1/3         | 5    | 4.0858e-16  | 49.3416      |
| (0.5, 1/9)          | 1/6         | 4    | 1.9201e-15  | 40.7350      |
| (0.25, 2/5)         | 1/2         | 5    | 5.815e-16   | 52.0825      |
| (0.25, 1/3)         | 5/12        | 5    | 3.9646e-16  | 51.6239      |
| (0.25, 1/10)        | 1/8         | 4    | 2.7781e-15  | 40.7989      |
The results for the problem size $n = 4096$

| $(\alpha, c)$ | $L\beta$ | iter | Res        | CPU time (s) |
|----------------|----------|------|------------|--------------|
| $(0.5, 1/3)$   | $1/2$    | 5    | $5.8453e-16$ | 351.6434     |
| $(0.5, 2/9)$   | $1/3$    | 5    | $4.0858e-16$ | 350.8556     |
| $(0.5, 1/9)$   | $1/6$    | 4    | $1.9200e-15$ | 290.0932     |
| $(0.25, 2/5)$  | $5/12$   | 5    | $7.4419e-16$ | 363.0767     |
| $(0.25, 1/3)$  | $5/12$   | 5    | $5.7781e-16$ | 353.1909     |
| $(0.25, 1/10)$ | $1/8$    | 4    | $2.5644e-15$ | 275.7286     |

obtained as the following form

$$
f(u) := M u + h^2 \phi(u) = 0,
$$

where

$$
u = (u_1, u_2, \ldots, u_N)^T, \quad u_i = (u_{i1}, u_{i2}, \ldots, u_{iN})^T, \quad i = 1, 2, \ldots, N,
$$
h = $\frac{1}{N+1}$ is the equidistant step-size with $N$ as a prescribed positive integer,

$$
\phi(u) = (u_{c1}, u_{c2}, \ldots, u_{cn})^T, \quad n = N \times N,
$$

and

$$
M = T_x \otimes I + I \otimes T_y.
$$

Here, $\otimes$ means the Kronecker product, $T_x$ and $T_y$ are tridiagonal matrices given by

$$
T_x = \begin{bmatrix}
2 & -1 + Re_1 & & \\
-1 - Re_1 & 2 & -1 + Re_1 & \\
& \ddots & \ddots & \ddots & \\
& & -1 - Re_1 & -1 + Re_1 & 2
\end{bmatrix}
$$

and

$$
T_y = \begin{bmatrix}
2 & -1 + Re_2 & & \\
-1 - Re_2 & 2 & -1 + Re_2 & \\
& \ddots & \ddots & \ddots & \\
& & -1 - Re_2 & -1 + Re_2 & 2
\end{bmatrix}
$$

with

$$
Re_i = \frac{1}{2} q_i h, \quad i = 1, 2,
$$

and $Re = \max\{Re_1, Re_2\}$ is the mesh Reynolds number.

The Jacobian matrix of $f$ defined by (49) at $u$ has the following form

$$
f'(u) = M + ch^2 D
$$

with $D = \text{diag}(u_{c1}^{-1}, u_{c2}^{-1}, \ldots, u_{cn}^{-1})$. Choose initial guess

$$
u_0 = e := (1, 1, \ldots, 1)^T.$$
Then, we have
\[ f'(u_0) = M + ch^2 I. \]
Thus, for any \( u, v \in \mathbb{R}^n \), it follows that
\[
\|f'(u_0)^{-1}(f'(v) - f'(u))\| \leq \int_{\|u-u_0\| + \|v-u\|} L(t) \, dt \leq ch^2 \alpha \|v - u\|^{c-1},
\]
with \( L(t) = c(c - 1)h^2 \alpha t^{c-2} \), where \( \alpha := \|f'(u_0)^{-1}\| \). Therefore, Theorem 1 with \( c \in [1, 2) \) is applicable, but not Corollary 1, to concluding that the sequence \( \{u_k\} \) generated by the two-step Newton method with initial guess \( u_0 = e \) converges to a local solution of (48).

Remark 6 If \( c = 2 \), then \( L(t) \) reduces a positive constant. That is, \( f' \) satisfies Lipschitz condition. Thus, Corollary 1 is applicable in this case.

5 Conclusions

In this paper, we have presented the semilocal convergence analysis for two-step Newton method (3) under the assumption that the first derivative \( F' \) satisfies the \( L \)-average Lipschitz conditions (14) in Banach spaces. The main results are contained in Theorem 1. When the unified convergence criteria \( 0 < \beta \leq b \) given by Wang in [63] is satisfied, the existence and uniqueness of a solution \( x^* \in B(x_0, r) \) of (1) are shown, and the superquadratic convergence of the sequence \( \{x_k\} \) is also proved. Moreover, we proved that the sequence \( \{x_k\} \) is Q-cubically convergent if the condition (24) is satisfied additionally. Three special cases which include the Kantorovich-type conditions, \( \gamma \)-conditions and Nesterov-Nemirovskii conditions have been provided. We also have applied our convergence result to solve the approximation of minimal positive solution for NSARE (41) arising from transport theory. Moreover, we consider a two-dimensional nonlinear convection-diffusion equation to show the advantage of considering the \( L \)-average Lipschitz condition rather than the classical Lipschitz condition. One goal of our future research is to exploit this general theory to develop more practical and efficient two-step inexact Newton method for solving IEP that can easily capture practical applications in large-scale settings.

Funding This work was supported in part by the Fujian Province National Science Foundation of China (Grant No. 2016J05015), Distinguished Young Scientific Research Talents Plan in Universities of Fujian Province, and the Education Research Projects for Young Teachers of Fujian Provincial Education Department (Grant No. JT180300).

References

1. Alvarez, F., Bolte, J., Munier, J.: A unifying local convergence result for Newton’s method in Riemannian manifolds. Found. Comput. Math. 8(2), 197–226 (2008)
2. Amat, S., Busquier, S., Gutiérrez, J.M.: Third-order iterative methods with applications to Hammerstein equations: a unified approach. J. Comput. Appl. Math. 235(9), 2936–2943 (2011)
3. Appell, J., De Pascale, E., Evkhuta, N.A., Zabrejko, P.P.: On the two-step Newton method for the solution of nonlinear operator equations. Math. Nachr. 172(1), 5–14 (1995)
4. Argyros, I.K.: A new semilocal convergence theorem for Newton’s method in Banach space using hypotheses on the second Fréchet-derivative. J. Comput. Appl. Math. 130(1-2), 369–373 (2001)
5. Argyros, I.K., Hilout, S.: Extending the applicability of the Gauss–Newton method under average Lipschitz–type conditions. Numer. Algorithms 58(1), 23–52 (2011)
6. Argyros, I.K., Khatriri, S.K.: Weak convergence conditions for the Newton’s method in Banach space using general majorizing sequences. Appl. Math. Comput. 263, 59–72 (2015)
7. Bai, Z., Gao, Y., Lu, L.: Fast iterative schemes for nonsymmetric algebraic Riccati equations arising from transport theory. SIAM J. Sci. Comput. 30(2), 804–818 (2008)
8. Bertaccini, D., Golub, G.H., Serra-Capizzano, S.: Spectral analysis of a preconditioned iterative method for the convection-diffusion equation. SIAM J. Matrix Anal. Appl. 29(1), 260–278 (2007)
9. Bittencourt, T., Ferreira, O.P.: Kantorovich’s theorem on Newton’s method under majorant condition in Riemannian manifolds. J. Global Optim. 68(2), 387–411 (2017)
10. Candela, V., Marquina, A.: Recurrence relations for rational cubic methods i: The Halley method. Computing 44(2), 169–184 (1990)
11. Candela, V., Marquina, A.: Recurrence relations for rational cubic methods II: The Chebyshev method. Computing 45(2), 355–367 (1990)
12. Chen, P.: Approximate zeros of quadratically convergent algorithms. Math. Comput. 63(207), 247–270 (1994)
13. Chen, X., Wen, C., Sun, H.W.: Two-step Newton-type methods for solving inverse eigenvalue problems. Numer. Linear Algebra Appl. 25(5), e2185 (2018)
14. Dedieu, J.P.: Approximate solutions of analytic inequality systems. SIAM J. Optim. 11(2), 411–425 (2000)
15. Dedieu, J.P., Kim, M.H.: Newton’s method for analytic systems of equations with constant rank derivatives. J. Complexity 18(1), 187–209 (2002)
16. Dedieu, J.P., Priouret, P., Malajovich, G.: Newton’s method on Riemannian manifolds: Covariant alpha theory.IMA J. Numer. Anal. 23(3), 395–419 (2003)
17. Dedieu, J.P., Shub, M.: Multihomogeneous Newton methods. Math. Comput. 69(231), 1071–1098 (2000)
18. Dedieu, J.P., Shub, M.: Newton’s method for overdetermined systems of equations. Math. Comput. 69(231), 1099–1115 (2000)
19. Deuflhard, P., Heindl, G.: Affine invariant convergence theorems for Newton’s method and extensions to related methods. SIAM J. Numer. Anal. 16(1), 1–10 (1979)
20. Ezquerro, J.A., Hernández, M.A.: Generalized differentiability conditions for Newton’s method. IMA J. Numer. Anal. 22(4), 519–530 (2002)
21. Ezquerro, J.A., Hernández, M.A.: On the R-order of the Halley method. J. Math. Anal. Appl. 303(2), 591–601 (2005)
22. Ezquerro, J.A., Hernández, M.A.: An improvement of the region of accessibility of Chebyshev’s method from Newton’s method. Math. Comput. 78(267), 1613–1627 (2009)
23. Ezquerro, J.A., Hernández, M.A.: An optimization of Chebyshev’s method. J. Complexity 25(4), 343–361 (2009)
24. Ezquerro, J.A., Hernández, M.A., Magreñán, A.A.: Starting points for Newton’s method under a center Lipschitz condition for the second derivative. J. Comput. Appl. Math. 330, 721–731 (2018)
25. Ferreira, O.P.: Local convergence of Newton’s method in Banach space from the viewpoint of the majorant principle.IMA J. Numer. Anal. 29(3), 746–759 (2009)
26. Ferreira, O.P.: A robust semi-local convergence analysis of Newton’s method for cone inclusion problems in Banach spaces under affine invariant majorant condition. J. Comput. Appl. Math. 279, 318–335 (2015)
27. Ferreira, O.P., Svaiter, B.F.: Kantorovich’s majorants principle for Newton’s method. Comput. Optim. Appl. 42(2), 213–229 (2009)
28. Ferreira, O.P., Svaiter, B.F.: A robust Kantorovich’s theorem on the inexact Newton method with relative residual error tolerance. J. Complex. 28(3), 346–363 (2012)
29. Gragg, W.B., Tapia, R.A.: Optimal error bounds for the newton-Kantorovich theorem. SIAM J. Numer. Anal. 11(1), 10–13 (1974)
30. Guo, C.H., Laub, A.J.: On the iterative solution of a class of nonsymmetric algebraic Riccati equations. SIAM J. Matrix Anal. Appl. 22(2), 376–391 (2000)
31. Guo, X.P., Duff, I.S.: Semilocal and global convergence of the newton-HSS method for systems of nonlinear equations. Numer Linear Algebra Appl. 18(3), 299–315 (2011)
32. Gutiérrez, J.M., Hernández, M.A.: Newton’s method under weak Kantorovich conditions. IMA J. Numer. Anal. 20(4), 521–532 (2000)
33. Han, D., Wang, X.: The error estimates of Halley’s method. Numer. Math. JCU (Engl. Ser.) 6, 231–240 (1997)
34. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms i: Fundamentals, vol. 305 Springer science & business media (1993)
35. Huang, Z.: A note on the Kantorovich theorem for Newton iteration. J. Comput. Appl. Math. 47(2), 211–217 (1993)
36. Iannazzo, B.: A family of rational iterations and its application to the computation of the matrix $p$th root. SIAM J. Matrix Anal. Appl. 30, 1445–1462 (2008)
37. Jay, L.O.: A note on Q-order of convergence. BIT Numer. Math. 41(2), 422–429 (2001)
38. Juang, J.: Existence of algebraic matrix Riccati equations arising in transport theory. Linear Algebra Appl. 230, 89–100 (1995)
39. Kantorovich, L.V., Akilov, G.P. Functional Analysis, 2nd edn. Pergamon Press, Oxford (1982)
40. Kelley, C.T.: Numerical methods for nonlinear equations. Acta Numer. 27, 207–287 (2018)
41. Li, C., Hu, N., Wang, J.: Convergence behavior of Gauss-Newton’s method and extensions of the Smale point estimate theory. J. Complexity 26(3), 268–295 (2010)
42. Li, C., Ng, K.: Majorizing functions and convergence of the Gauss-Newton for convex composite optimization. SIAM J. Optim. 18(2), 613–642 (2007)
43. Li, C., Ng, K.: Approximate solutions for abstract inequality systems. SIAM J. Optim. 23(2), 1237–1256 (2013)
44. Li, C., Ng, K.: Extended Newton methods for conic inequalities: Approximate solutions and the extended Smale $\alpha$-theory. J. Math. Anal. Appl. 440(2), 636–660 (2016)
45. Li, C., Ng, K.: Quantitative analysis for perturbed abstract inequality systems in Banach spaces. SIAM J. Optim. 28(4), 2872–2901 (2018)
46. Li, C., Wang, J.: Newton’s method on Riemannian manifolds: Smale’s point estimate theory under the $\gamma$–condition. IMA J. Numer. Anal. 26(2), 228–251 (2006)
47. Lin, Y., Bao, L.: Convergence analysis of the Newton–Shamanskii method for a nonsymmetric algebraic Riccati equation. Numer. Linear Algebra Appl. 15(6), 535–546 (2008)
48. Ling, Y., Huang, Z.: An analysis on the efficiency of Euler’s method for computing the matrix $p$th root. Numer. Linear Algebra Appl. 24(6), e2104 (2017)
49. Ling, Y., Xu, X.: On the semilocal convergence behavior for Halley’s method. Comput. Optim. Appl. 58(3), 597–618 (2014)
50. Ling, Y., Xu, X.: On one-parameter family of Newton-like iterations for solving nonsymmetric algebraic Riccati equation from transport theory. J. Nonlinear Convex Anal. 18(10), 1833–1848 (2017)
51. Lu, L.Z.: Solution form and simple iteration of a nonsymmetric algebraic Riccati equation arising in transport theory. SIAM J. Matrix Anal. Appl. 26(3), 679–685 (2005)
52. Ma, W., Chen, X.S.: Two-step inexact Newton-type method for inverse singular value problems. Numer. Algorithms. https://doi.org/10.1007/s11075-019-00783-x. (to appear) (2019)
53. Magreñán Ruiz, A.A., Argyros, I.K.: Two-step Newton methods. J. Complexity 30(4), 533–553 (2014)
54. Nakatsukasa, Y., Bai, Z., Gygi, F.: Optimizing Halley’s iteration for computing the matrix polar decomposition. SIAM J. Matrix Anal. Appl. 31(5), 2700–2720 (2010)
55. Nakatsukasa, Y., Freund, R.W.: Computing fundamental matrix decompositions accurately via the matrix sign function in two iterations: The power of Zolotarev’s functions. SIAM Rev. 58(3), 461–493 (2016)
56. Nesterov, Y., Nemirovskii, A.: Interior-point Polynomial Algorithms in Convex Programming. SIAM, Philadelphia (1994)
57. Potra, F.A.: On Q-order and R-order of convergence. J. Optim. Theory Appl. 63(3), 415–431 (1989)
58. Shub, M., Smale, S.: Complexity of b´ezout’s theorem IV: Probability of success, extensions. SIAM J. Numer. Anal. 33(1), 128–148 (1996)
59. Smale, S.: Newton’s method estimates from data at one point. In: Ewing, R., Gross, K., Martin, C. (eds.) The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, pp. 185–196. Springer, New York (1986)
60. Smale, S.: The fundamental theory for solving equations. In: Proceeding of the International Congress of Mathematicians, p. 185. AMS, Providence (1987)
61. Smale, S.: Complexity theory and numerical analysis. Acta Numer. 6, 523–552 (1997)
62. Wang, J., Hu, Y., Yu, C.K.W., Li, C., Yang, X.: Extended Newton methods for multiobjective optimization: majorizing function technique and convergence analysis. SIAM J. Optim. 29(3), 2388–2421 (2019)
63. Wang, X.: Convergence of Newton’s method and inverse functions theorem in Banach space. Math. Comput. 68, 169–186 (1999)
64. Wang, X., Han, D.: On dominating sequence method in the point estimate and Smale theorem. Sci. China Ser. A 33(2), 135–144 (1990)
65. Wang, X., Han, D.: Criterion $\alpha$ and Newton’s method under weak conditions. Chinese J. Numer. Appl. Math. 19(1), 96–105 (1997)
66. Wang, X., Li, C.: Local and global behavior for algorithms of solving equations. Chinese Sc. Bull. 46(6), 444–451 (2001)
67. Xie, Z.: Concerning Kantorovich-type theorem of Newton’s method and its majorants (in Chinese). Adv. Math. 41(6), 641–654 (2012)
68. Xu, X., Li, C.: Convergence criterion of Newton’s method for singular systems with constant rank derivatives. J. Math. Anal. Appl. 345(2), 689–701 (2008)
69. Zabrejko, P.P., Nguyen, D.F.: The majorant method in the theory of Newton-Kantorovich approximations and the Ptak error estimates. Numer. Funct. Anal. Optim. 9(5-6), 671–684 (1987)
70. Zhong, H.X., Chen, G.L., Guo, X.P.: Convergence of the newton-HSS method under the Lipschitz condition with the L-average. J. Comput. Anal. Appl 25(5), 952–964 (2018)

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.