

# Local convergence analysis of inexact Newton-like methods under majorant condition

O.P. Ferreira · M.L.N. Gonçalves

Received: 23 July 2008 / Published online: 22 April 2009  
© Springer Science+Business Media, LLC 2009

**Abstract** We present a local convergence analysis of inexact Newton-like methods for solving nonlinear equations under majorant conditions. This analysis provides an estimate of the convergence radius and a clear relationship between the majorant function, which relaxes the Lipschitz continuity of the derivative, and the nonlinear operator under consideration. It also allow us to obtain some important special cases.

**Keywords** Inexact Newton method · Majorant condition · Local convergence

## 1 Introduction

Newton's method and its variations (see [6]), including the inexact Newton methods, are the most efficient methods known for solving nonlinear equations

$$F(x) = 0, \quad (1)$$

where  $F : \Omega \rightarrow \mathbb{R}^n$  is a continuously differentiable function and  $\Omega \subseteq \mathbb{R}^n$  is an open set. The inexact Newton method was introduced by Dembo, Eisenstat and Steihaug in [5] denoting any method which, given an initial point  $x_0$ , generates the sequence  $\{x_k\}$  as follows:

$$x_{k+1} = x_k + S_k, \quad k = 0, 1, \dots,$$

---

O.P. Ferreira was supported in part by CNPq Grants 480101/2008-6, 302024/2008-5, PRONEX–Optimization (FAPERJ/CNPq) and FUNAPE/UFG.

M.L.N. Gonçalves was supported in part by CAPES.

O.P. Ferreira (✉) · M.L.N. Gonçalves  
IME/UFG, Campus II, Caixa Postal 131, 74001-970 Goiânia, GO, Brazil  
e-mail: [orizon@mat.ufg.br](mailto:orizon@mat.ufg.br)

M.L.N. Gonçalves  
e-mail: [maxlng@hotmail.com](mailto:maxlng@hotmail.com)

where  $S_k$  is the solution of the linear system

$$F'(x_k)S_k = -F(x_k) + r_k,$$

for a suitable residual  $r_k \in \mathbb{R}^n$ . Let  $x_*$  be a solution of (1) such that  $F'(x_*)$  is invertible. As shown in [5], if  $\|r_k\| \leq \theta_k \|F(x_k)\|$  for  $k = 0, 1, \dots$  and  $\{\theta_k\}$  is a sequence of forcing terms such that  $0 \leq \theta_k < 1$  then there exists  $\epsilon > 0$  such that the sequence  $\{x_k\}$ , for any initial point  $x_0 \in B(x_*, \epsilon) = \{x \in \mathbb{R}^n : \|x_* - x\| < \epsilon\}$ , is well defined and converges linearly to  $x_*$  in the norm  $\|y\|_* = \|F'(x_*)y\|$ , where  $\|\cdot\|$  is any norm in  $\mathbb{R}^n$ . It is worth noting that, in [5], no Lipschitz condition is assumed on the derivative  $F'$  to prove that  $\{x_k\}$  is well defined and linearly converging; however, no estimate of the convergence radius  $\epsilon$  is provided. As pointed out by [11] (see also [13]) the result of [5] is difficult to apply due to dependence of the norm  $\|\cdot\|_*$ , which is not computable.

Formally, the inexact Newton-like methods for solving the non-linear equation (1), which we will consider, are described as follows: Given an initial point  $x_0 \in \Omega$ , define

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F(x_k) + r_k, \quad k = 0, 1, \dots,$$

where  $B(x_k)$  is a suitable invertible approximation of the derivative  $F'(x_k)$  and the residual  $r_k$  satisfies

$$\|P_k r_k\| \leq \theta_k \|P_k F(x_k)\|,$$

for suitable forcing sequence  $\{\theta_k\}$  and some invertible matrix sequence  $\{P_k\}$  of preconditioners for the above linear equation defining the steep  $S_k$ . This method was considered for the first time in [13], and was also considered in [4]. In particular, letting  $P_k \equiv I$  be the identity matrix and  $B_k = F'(x_k)$  for each  $k$ , we obtain the inexact Newton method considered in [5, 12] and [20].

Inexact Newton-like methods may fail to converge and may even fail to be well defined. To ensure that the method is well defined and converges to a solution of a given non-linear equation, some conditions must be imposed. For instance, the classical convergence analysis (see [5]) requires the initial iterate to be “close enough” to a solution and the first derivative of the non-linear function to be invertible in this solution. Moreover, for estimating the convergence radius, the Lipschitz continuity or something like Lipschitz continuity, of the first derivative is also assumed (see [4, 13] and [20]).

In the last twenty-five years, there have been papers dealing with the issue of convergence of the Newton methods, including the inexact Newton methods and the Gauss-Newton methods, by relaxing the assumption of Lipschitz continuity of the derivative (see [1, 3, 4, 8, 9, 12, 17, 18] and [20]). In addition to improving the convergence theory (this allows us to estimate the convergence radius and to enlarge the range of application) some modifications of the Lipschitz condition also permit us to unify several results. Works dealing with this subject include [1, 8, 9, 17] and [18].

Our aim in this paper is to present a new local convergence analysis for inexact Newton-like methods under majorant condition. In our analysis, the classical Lipschitz condition is relaxed using a majorant function. It is worth pointing out that this condition is equivalent to Wang’s condition introduced in [17] and used by Chen and

Li in [4] to study the inexact Newton-like methods. The convergence analysis presented is linear in an arbitrary norm. It provides a new estimate for the convergence radius and a clear relationship between the majorant function and the nonlinear operator under consideration. Moreover, this analysis merges in the analysis on Newton method, see [17] and [9], and Newton-like methods, see [13] (see also [4]), besides affine invariant results are obtained in some particular cases (see [19]). The results also allows us to obtain some special cases that can be evaluated as an application (which has not appeared in the literature).

The organization of the paper is as follows. In Sect. 1.1, we list some notations and basic results used in our presentation. In Sect. 2 the main result is stated, and in Sect. 2.1 some properties involving the majorant function are established. In Sect. 2.2 we presented the relationships between the majorant function and the non-linear operator. In Sect. 2.3 the main result is proved and some applications of this result are given in Sect. 3. Some final remarks are made in Sect. 4.

### 1.1 Notations and auxiliary results

The following notations and results are used throughout our presentation. Let  $\mathbb{R}^n$  be with a norm  $\|\cdot\|$ . The open and closed ball at  $a \in \mathbb{R}^n$  and radius  $\delta > 0$  are denoted, respectively by

$$B(a, \delta) = \{x \in \mathbb{R}^n; \|x - a\| < \delta\}, \quad B[a, \delta] = \{x \in \mathbb{R}^n; \|x - a\| \leq \delta\}.$$

Let  $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$  be the space of linear operators of  $\mathbb{R}^n$ . Define the operator norm associated to the norm  $\|\cdot\|$  as

$$\|T\| := \sup\{\|Tx\|, \|x\| \leq 1\}, \quad T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n).$$

The condition number of an invertible operator  $T$  is denoted by  $\text{cond}(T) := \|T\| \|T^{-1}\|$ .

**Lemma 1** (Banach's lemma) *Let  $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$  and  $I \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ , the identity operator. If  $\|B - I\| < 1$ , then  $B$  is invertible and  $\|B^{-1}\| \leq 1/(1 - \|B - I\|)$ .*

*Proof* See the proof of Lemma 1, p. 189 of Smale [15] with  $A = I$  and  $c = \|B - I\|$ .  $\square$

**Proposition 2** *If  $0 \leq t < 1$ , then  $\sum_{i=0}^{\infty} (i+2)(i+1)t^i = 2/(1-t)^3$ .*

*Proof* Take  $k = 2$  in Lemma 3, p. 161 of Blum, Cucker, Shub and Smale [2].  $\square$

Also, the following auxiliary result of elementary convex analysis will be needed:

**Proposition 3** *Let  $\epsilon > 0$  and  $\tau \in [0, 1]$ . If  $\varphi : [0, \epsilon) \rightarrow \mathbb{R}$  is convex, then  $l : (0, \epsilon) \rightarrow \mathbb{R}$  defined by*

$$l(t) = \frac{\varphi(t) - \varphi(\tau t)}{t},$$

*is non-increasing.*

*Proof* See Theorem 4.1.1 and Remark 4.1.2 on p. 21 of Hiriart-Urruty and Lemaréchal [10]. □

## 2 Local convergence of inexact Newton-like method

Our goal is to state and prove a local theorem for *inexact Newton-like methods*. Assuming that the non-linear equation  $F(x) = 0$  has a solution  $x_*$ , we will, under mild conditions, prove that the inexact Newton-like method is well defined and that the generated sequence converges linearly to this solution. The statement of the theorem is as follows:

**Theorem 4** *Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow \mathbb{R}^n$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and*

$$\kappa := \sup \{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

*Suppose that  $F(x_*) = 0$ ,  $F'(x_*)$  is invertible and there exists a  $f : [0, R) \rightarrow \mathbb{R}$  continuously differentiable such that*

$$\left\| F'(x_*)^{-1} [F'(x) - F'(x_* + \tau(x - x_*))] \right\| \leq f'(\|x - x_*\|) - f'(\tau\|x - x_*\|), \quad (2)$$

*for  $\tau \in [0, 1]$  and  $x \in B(x_*, \kappa)$ , where*

- (h1)  $f(0) = 0$  and  $f'(0) = -1$ ;
- (h2)  $f'$  is convex and strictly increasing.

*Take  $0 \leq \vartheta < 1$ ,  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1\vartheta + \omega_2 < 1$ . Let  $\nu := \sup\{t \in [0, R) : f'(t) < 0\}$ ,*

$$\begin{aligned} \rho &:= \sup\{t \in (0, \nu) : \omega_1(1 + \vartheta)[f(t)/(tf'(t)) - 1] + \omega_1\vartheta + \omega_2 < 1\}, \\ \sigma &:= \min\{\kappa, \rho\}. \end{aligned}$$

*Then, the inexact Newton-like methods for solving  $F(x) = 0$ , with initial point  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$*

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F(x_k) + r_k, \quad k = 0, 1, \dots, \quad (3)$$

*with the following conditions for the forcing term  $\theta_k$  and the residual  $r_k$ ,*

$$0 \leq \theta_k \text{cond}(P_k F'(x_k)) \leq \vartheta, \quad \|P_k r_k\| \leq \theta_k \|P_k F(x_k)\|, \quad k = 0, 1, \dots,$$

*where  $\{P_k\}$  is a sequence of invertible matrix (preconditioners for the linear system in (3)) and  $B(x_k)$  is an invertible approximation of  $F'(x_k)$  satisfying*

$$\|B(x_k)^{-1} F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1} F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, the generated sequence  $\{x_k\}$  is contained in  $B(x_*, \sigma)$  and converges to  $x_*$ . Moreover,  $\{x_k\}$  converges linearly as follows

$$\|x_{k+1} - x_*\| \leq \left[ \omega_1(1 + \vartheta) \left( \frac{f(\|x_0 - x_*\|)}{\|x_0 - x_*\| f'(\|x_0 - x_*\|)} - 1 \right) + \omega_1\vartheta + \omega_2 \right] \times \|x_k - x_*\|, \quad k = 0, 1, \dots$$

Note that letting the majorant function  $f : [0, \kappa) \rightarrow \mathbb{R}$  be given by  $f(t) = Kt^2/2 - t$ ,  $B(x_k) = F'(x_k)$ ,  $P_k = I$ ,  $\omega_1 = 1$  and  $\omega_2 = 0$  in Theorem 4, we obtain the following local convergence result for the inexact Newton method:

**Theorem 5** *Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow \mathbb{R}^n$  be continuously differentiable in  $\Omega$ . Take  $x_* \in \Omega$  and let  $\kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}$ . Assume that  $F'(x_*)$  is invertible,  $F(x_*) = 0$ , there exists a  $K > 0$  such that*

$$\left\| F'(x_*)^{-1} [F'(x) - F'(y)] \right\| \leq K \|x - y\|, \quad \forall x, y \in B(x_*, \kappa).$$

Take  $0 \leq \vartheta < 1$ . Let

$$\sigma := \min \{ \kappa, 2(1 - \vartheta)/(K(3 - \vartheta)) \}.$$

Then, the inexact Newton-like method for solving  $F(x) = 0$ , with the initial point  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F(x_k) + r_k, \quad k = 0, 1, \dots,$$

with the following conditions for the forcing term  $\theta_k$  and the residual  $r_k$ ,

$$0 \leq \theta_k \text{ cond}(F'(x_k)) \leq \vartheta, \quad \|r_k\| \leq \theta_k \|F(x_k)\|, \quad k = 0, 1, \dots,$$

is well defined, the generated sequence  $\{x_k\}$  is contained in  $B(x_*, \sigma)$  and converges to  $x_*$ . Moreover,  $\{x_k\}$  converges linearly as follow

$$\|x_{k+1} - x_*\| \leq \left[ (1 + \vartheta) \frac{K \|x_0 - x_*\|}{2(1 - K \|x_0 - x_*\|)} + \vartheta \right] \|x_k - x_*\|, \quad k = 0, 1, \dots$$

*Remark 1* Note that letting  $\vartheta = 0$  (in this case  $\theta_k \equiv 0$  and  $r_k \equiv 0$ ) in Theorem 5 we obtain  $r = \min\{\kappa, 2/(3K)\}$ . As was shown in [16] (see also, [19]) this is the best possible convergence radius for Newton’s Method. Therefore, for vanishing residuals, Theorem 5 merges into the theory of Newton’s Method and, as a consequence, Theorem 4 does too.

In order to prove Theorem 4 we need some results. From here on, we assume that all assumptions of Theorem 4 hold.

## 2.1 The majorant function

Our first goal is to show that the constant  $\kappa$  associated with  $\Omega$  and the constants  $\nu$ ,  $\rho$  and  $\sigma$  associated with the majorant function  $f$  are positive. Also, we will prove some results related to the function  $f$ .

We begin by noting that  $\kappa > 0$ , because  $\Omega$  is an open set and  $x_* \in \Omega$ .

**Proposition 6** *The following statements hold:*

- (i)  $\nu > 0$ ;
- (ii)  $f'(t) < 0, \forall t \in [0, \nu)$ ;
- (iii) *The map  $[0, \nu) \ni t \mapsto 1/|f'(t)|$  is strictly increasing and*
- (iv)  $t - f(t)/f'(t) < 0, \forall t \in [0, \nu)$ .

*Proof* As  $f'$  is continuous in  $(0, R)$  and  $f'(0) = -1$ , there exists a  $\delta > 0$  such that  $f'(t) < 0$  for all  $t \in (0, \delta)$ . So,  $\nu \geq \delta$  and item (i) is proved.

For proving item (ii), use (h2) and the definition of  $\nu$ . Now, for establishing the validity of item (iii), combine (h2) and item (ii).

Since  $f'$  is strictly increasing we have  $f$  is strictly convex. So,

$$f(0) > f(t) - tf'(t), \quad \forall t \in (0, R).$$

Because  $f(0) = 0$  and  $f'(t) < 0$  for all  $t \in (0, \nu)$ , the inequality in item (iv) follows from above equation.  $\square$

Let  $n_f$  be the Newton iteration associated with the majorant function,

$$\begin{aligned} n_f : [0, \nu) &\rightarrow (-\infty, 0] \\ t &\mapsto t - f(t)/f'(t). \end{aligned} \quad (4)$$

From Proposition 6 we have  $f' < 0$  in  $[0, \nu)$ . Hence, the Newton iteration associated with the majorant function is well defined in  $[0, \nu)$ .

**Proposition 7** *The map  $(0, \nu) \ni t \mapsto |n_f(t)|/t^2$  is strictly increasing.*

*Proof* Using item (iv) of Proposition 6 and (h1) we obtain, after simple algebraic manipulation, that

$$\frac{|n_f(t)|}{t^2} = \frac{1}{|f'(t)|} \int_0^1 \frac{f'(t) - f'(\tau t)}{t} d\tau, \quad \forall t \in (0, \nu). \quad (5)$$

On the other hand, since  $f'$  is strictly increasing, we obtain that the map

$$(0, \nu) \ni t \mapsto \frac{f'(t) - f'(\tau t)}{t},$$

is positive for all  $\tau \in (0, 1)$ . Also, from (h2) we know that  $f'$  is convex. So, applying Proposition 3 with  $f' = \varphi$  and  $\epsilon = \nu$ , we conclude that the last map is increasing.

Hence the second term in the right hand side of (5) is positive and increasing. Therefore, since Proposition 6 implies that the first term in the right hand side of (5) is positive and strictly increasing, we conclude the statement.  $\square$

**Corollary 8** *The map  $(0, \nu) \ni t \mapsto |n_f(t)|/t$  is strictly increasing.*

*Proof* It is immediate, by noting that  $|n_f(t)|/t = (|n_f(t)|/t^2)t$  is a product of two strictly increasing functions.  $\square$

**Proposition 9** *The constant  $\rho$  is positive and there holds*

$$\omega_1(1 + \vartheta) \frac{|n_f(t)|}{t} + \omega_1\vartheta + \omega_2 < 1, \quad \forall t \in (0, \rho).$$

*Proof* Using Proposition 6 and the definition (4), we have

$$0 < f(t)/(tf'(t)) - 1 = (f(t)/f'(t) - t)/t = |n_f(t)|/t, \quad \forall t \in (0, \nu). \quad (6)$$

Now, because Proposition 7 implies that  $|n_f(t)|/t^2$  is bounded near zero, we obtain

$$\lim_{t \rightarrow 0} |n_f(t)|/t = \lim_{t \rightarrow 0} (|n_f(t)|/t^2)t = 0. \quad (7)$$

Thus, since  $1 - (\omega_1\vartheta + \omega_2)/\omega_1(1 + \vartheta) > 0$ , using (6) and (7) we conclude that there exists a  $\delta > 0$  such that

$$0 < (f(t)/(tf'(t)) - 1) < 1 - (\omega_1\vartheta + \omega_2)/\omega_1(1 + \vartheta), \quad \forall t \in (0, \delta),$$

or, equivalently,

$$0 < \omega_1(1 + \vartheta)[f(t)/(tf'(t)) - 1] + \omega_1\vartheta + \omega_2 < 1, \quad \forall t \in (0, \delta). \quad (8)$$

Hence, combining the last inequality and the definition of  $\rho$ , we have  $\delta \leq \rho$ , which is a proof of the first statement.

For concluding the proof, we use the definition of  $\rho$ , inequality (8), (6) and Corollary 8.  $\square$

## 2.2 Relationship of the majorant function with the non-linear operator

In this section we will present the main relationships between the majorant function  $f$  and the non-linear operator  $F$ .

**Lemma 10** *Let  $x \in \Omega$ . If  $\|x - x_*\| < \min\{\nu, \kappa\}$ , then  $F'(x)$  is invertible and*

$$\|F'(x)^{-1}F'(x_*)\| \leq 1/|f'(\|x - x_*\|)|.$$

*In particular,  $F'$  is invertible in  $B(x_*, \sigma)$ .*

*Proof* Let  $x \in \Omega$  such that  $\|x - x_*\| < \min\{\nu, \kappa\}$ . So  $f'(\|x - x_*\|) < 0$  which, together with (2), implies

$$\begin{aligned} \|F'(x_*)^{-1}F'(x) - I\| &= \|F'(x_*)^{-1}[F'(x) - F'(x_*)]\| \\ &\leq f'(\|x - x_*\|) - f'(0) < -f'(0) = 1. \end{aligned}$$

Thus, Lemma 1 and the last equation imply that  $F'(x_*)^{-1}F'(x)$  is invertible, as well as  $F'(x)$ , and

$$\begin{aligned} \|F'(x)^{-1}F'(x_*)\| &\leq \frac{1}{1 - \|F'(x_*)^{-1}F'(x) - I\|} \leq \frac{1}{1 - (f'(\|x - x_*\|) - f'(0))} \\ &= \frac{1}{|f'(\|x - x_*\|)|}, \end{aligned}$$

where we assume that  $f'(0) = -1$  and  $f' < 0$  in  $[0, \nu)$  in the last equality. As  $\sigma \leq \nu$  the last part is proved. □

The Newton iteration at a point happens to be a zero of the linearization of  $F$  at such point, which is also the first-order Taylor expansion of  $F$ . So, we study the linearization error at point in  $\Omega$

$$E_F(x, y) := F(y) - [F(x) + F'(x)(y - x)], \quad y, x \in \Omega. \tag{9}$$

We will bound this error by the error of the linearization of the majorant function  $f$

$$e_f(t, u) := f(u) - [f(t) + f'(t)(u - t)], \quad t, u \in [0, R). \tag{10}$$

**Lemma 11** *If  $\|x_* - x\| < \kappa$ , then there holds  $\|F'(x_*)^{-1}E_F(x, x_*)\| \leq e_f(\|x - x_*\|, 0)$ .*

*Proof* Since  $B(x_*, \kappa)$  is convex, we obtain that  $x_* + \tau(x - x_*) \in B(x_*, \kappa)$ , for  $0 \leq \tau \leq 1$ . Thus, as  $F$  is continuously differentiable in  $\Omega$ , the definition of  $E_F$  and some simple manipulations yield

$$\|F'(x_*)^{-1}E_F(x, x_*)\| \leq \int_0^1 \left\| F'(x_*)^{-1}[F'(x) - F'(x_* + \tau(x - x_*))]\right\| \|x_* - x\| \, d\tau.$$

From the last inequality and the assumption (2), we obtain

$$\|F'(x_*)^{-1}E_F(x, x_*)\| \leq \int_0^1 [f'(\|x - x_*\|) - f'(\tau\|x - x_*\|)] \|x - x_*\| \, d\tau.$$

Evaluating the above integral and using the definition of  $e_f$ , the statement follows. □

Define the Newton step to the functions  $F$  and  $f$  by the following equalities:

$$S_F(x) := -F'(x)^{-1}F(x), \quad s_f(t) := -f(t)/f'(t). \tag{11}$$



**Lemma 12** *If  $\|x - x_*\| < \min\{\nu, \kappa\}$ , then  $\|S_F(x)\| \leq s_f(\|x - x_*\|)$ .*

*Proof* Using (11),  $F(x_*) = 0$  and some algebraic manipulation, it follows from (9) that

$$\begin{aligned} \|S_F(x)\| &= \left\| -F'(x)^{-1} \left( F(x_*) - [F(x) + F'(x)(x_* - x)] \right) + (x_* - x) \right\| \\ &\leq \|F'(x)^{-1} F'(x_*)\| \|F'(x_*)^{-1} (F(x_*) - [F(x) + F'(x)(x_* - x)])\| + \|x_* - x\| \\ &= \|F'(x)^{-1} F'(x_*)\| \|F'(x_*)^{-1} E_F(x, x_*)\| + \|x - x_*\|. \end{aligned}$$

Combining the last equation with Lemmas 10 and 11 we have

$$\|S_F(x)\| \leq \frac{e_f(\|x - x_*\|, 0)}{|f'(\|x - x_*\|)|} + \|x - x_*\|.$$

Since  $f' < 0$  in  $[0, \nu]$  and  $\|x - x_*\| < \nu$ , we obtain from last inequality (10) and (h1), that

$$\begin{aligned} \|S_F(x)\| &\leq \frac{f(0) - f(\|x - x_*\|) + f'(\|x - x_*\|)\|x - x_*\|}{-f'(\|x - x_*\|)} + \|x - x_*\| \\ &= \frac{f(\|x - x_*\|)}{f'(\|x - x_*\|)}. \end{aligned}$$

So, the last inequality together with the second equality in (11) implies the desired inequality. □

**Lemma 13** *Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow \mathbb{R}^n$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and  $\kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}$ . Suppose that  $F(x_*) = 0$ ,  $F'(x_*)$  is invertible and there exists a  $f : [0, R) \rightarrow \mathbb{R}$  continuously differentiable satisfying (2), (h1) and (h2). Let  $\vartheta, \omega_2, \omega_1, \nu, \rho$  and  $\sigma$  as in Theorem 4. Assume that  $x \in B(x_*, \sigma) \setminus \{x_*\}$ , i.e.,  $0 < \|x - x_*\| < \sigma$ . Define*

$$x_+ = x + S, \quad B(x)S = -F(x) + r, \tag{12}$$

where  $B(x)$  is a invertible approximation of  $F'(x)$  satisfying

$$\|B(x)^{-1}F'(x)\| \leq \omega_1, \quad \|B(x)^{-1}F'(x) - I\| \leq \omega_2, \tag{13}$$

and the forcing term  $\theta$  and the residual  $r$  satisfy

$$0 \leq \theta \operatorname{cond}(PF'(x)) \leq \vartheta, \quad \|Pr\| \leq \theta \|PF(x)\|, \tag{14}$$

with  $P$  a invertible matrix (preconditioner for the liner system in (12)). Then  $x_+$  is well defined and there holds

$$\|x_+ - x_*\| \leq \left[ \omega_1(1 + \vartheta) \frac{|nf(\|x - x_*\|)|}{\|x - x_*\|} + \omega_1\vartheta + \omega_2 \right] \|x - x_*\|.$$

In particular,

$$\|x_+ - x_*\| < \|x - x_*\|.$$

*Proof* First note that, as  $\|x - x_*\| < \sigma$ , it follows from Lemma 10 that  $F'(x)$  is invertible. Now, let  $B(x)$  a invertible approximation of it satisfying (13). Thus,  $x_+$  is well defined. Now, as  $F(x_*) = 0$ , some simple algebraic manipulation and (12) yield

$$\begin{aligned} x_+ - x_* &= B(x)^{-1} (F(x_*) - [F(x) + F'(x)(x_* - x)]) \\ &\quad + (B(x)^{-1} F'(x) - I)(x_* - x) + B(x)^{-1} r. \end{aligned}$$

So, the above equation and (9) give

$$x_+ - x_* = B(x)^{-1} E_F(x, x_*) + (B(x)^{-1} F'(x) - I)(x_* - x) + B^{-1}(x)r.$$

Again, some algebraic manipulation in the above equation, together with the properties of the norm, imply

$$\begin{aligned} \|x_+ - x_*\| &\leq \|B(x)^{-1} F'(x)\| \|F'(x)^{-1} F'(x_*)\| \|F'(x_*)^{-1} E_F(x, x_*)\| \\ &\quad + \|B(x)^{-1} F'(x) - I\| \|x - x_*\| \\ &\quad + \|B(x)^{-1} F'(x)\| \|F'(x)^{-1} P^{-1}\| \|Pr\|. \end{aligned}$$

Taking into account the assumptions (13) and (14) we obtain from the last inequality that

$$\begin{aligned} \|x_+ - x_*\| &\leq \omega_1 \|F'(x)^{-1} F'(x_*)\| \|F'(x_*)^{-1} E_F(x, x_*)\| + \omega_2 \|x - x_*\| \\ &\quad + \omega_1 \theta \|F'(x)^{-1} P^{-1}\| \|PF(x)\|. \end{aligned}$$

On the other hand, (14) implies  $\theta \|(PF'(x))^{-1}\| \|PF'(x)\| \leq \vartheta$ . So, it follows from (11) and (14) that

$$\begin{aligned} \omega_1 \theta \|F'(x)^{-1} P^{-1}\| \|PF(x)\| &\leq \omega_1 \theta \|(PF'(x))^{-1}\| \|PF'(x)\| \|S_F(x)\| \\ &\leq \omega_1 \vartheta \|S_F(x)\|. \end{aligned}$$

Hence, it follows from the two latter equations that

$$\begin{aligned} \|x_+ - x_*\| &\leq \omega_1 \|F'(x)^{-1} F'(x_*)\| \|F'(x_*)^{-1} E_F(x, x_*)\| \\ &\quad + \omega_2 \|x - x_*\| + \omega_1 \vartheta \|S_F(x)\|. \end{aligned}$$

Combining the last equation with Lemma 10, Lemmas 11 and 12 we conclude that

$$\|x_+ - x_*\| \leq \omega_1 \frac{e_f(\|x - x_*\|, 0)}{|f'(\|x - x_*\|)|} + \omega_2 \|x - x_*\| + \omega_1 \vartheta s_f(\|x - x_*\|).$$

Now, using (10), (4), (h1) and (11) we have, by direct calculus,

$$\frac{e_f(\|x - x_*\|, 0)}{|f'(\|x - x_*\|)|} = |n_f(\|x - x_*\|)|, \quad s_f(\|x - x_*\|) = |n_f(\|x - x_*\|)| + \|x - x_*\|.$$

Therefore, it follows from above inequality and the two latter equalities that

$$\|x_+ - x_*\| \leq \omega_1 |n_f(\|x - x_*\|)| + \omega_2 \|x - x_*\| + \omega_1 \vartheta \left( |n_f(\|x - x_*\|)| + \|x - x_*\| \right),$$

which is equivalent to the first inequality of the lemma.

Because  $x \in B(x_*, \sigma) \setminus \{x_*\}$ , i.e.,  $0 < \|x - x_*\| < \sigma$  we obtain the last inequality of the lemma by combining the first one and Proposition 9 with  $t = \|x - x_*\|$ .  $\square$

### 2.3 Proof of Theorem 4

We are now in position to prove Theorem 4.

*Proof* Since  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$ , i.e.,  $0 < \|x_0 - x_*\| < \sigma$ , a straightforward induction argument and the last inequality in Lemma 13 implies that the sequence  $\{x_k\}$  generated by inexact Newton-like methods is well defined and contained in  $B(x_*, \sigma)$ .

Our task is now to show that  $\{x_k\}$  converges to  $x_*$ . Because,  $\{x_k\}$  is well defined and contained in  $B(x_*, \sigma)$ , applying Lemma 13 with  $x_+ = x_{k+1}$ ,  $x = x_k$ ,  $r = r_k$ ,  $B(x) = B(x_k)$ ,  $P = P_k$  and  $\theta = \theta_k$  we obtain

$$\|x_{k+1} - x_*\| \leq \left[ \omega_1 (1 + \vartheta) \frac{|n_f(\|x_k - x_*\|)|}{\|x_k - x_*\|} + \omega_1 \vartheta + \omega_2 \right] \|x_k - x_*\|, \quad k = 0, 1, \dots \tag{15}$$

$$\|x_{k+1} - x_*\| < \|x_k - x_*\|, \quad k = 0, 1, \dots$$

In particular, the last inequality implies that  $\|x_{k+1} - x_*\| < \|x_0 - x_*\|$ , for  $k = 0, 1, \dots$ , which, together with (15) and Corollary 8, gives

$$\|x_{k+1} - x_*\| \leq \left[ \omega_1 (1 + \vartheta) \frac{|n_f(\|x_0 - x_*\|)|}{\|x_0 - x_*\|} + \omega_1 \vartheta + \omega_2 \right] \|x_k - x_*\|, \quad k = 0, 1, \dots \tag{16}$$

As Proposition 9 with  $t = \|x_0 - x_*\|$  gives  $\omega_1 (1 + \vartheta) |n_f(\|x_0 - x_*\|)| / \|x_0 - x_*\| + \omega_1 \vartheta + \omega_2 < 1$ , we conclude from above equation that  $\{\|x_k - x_*\|\}$  converges to zero. So,  $\{x_k\}$  converges to  $x_*$ .

It remains to prove the last inequality of the theorem. For this, use (16) and the definition in (4).  $\square$

*Remark 2* If a continuously differentiable function  $f : [0, \kappa) \rightarrow \mathbb{R}$  is a majorant function satisfying the conditions (h1) and (h2), then the function  $h : (-\kappa, \kappa) \rightarrow \mathbb{R}$  is defined by

$$h(t) = \begin{cases} -f(-t), & t \in (-\kappa, 0], \\ f(t), & t \in [0, \kappa) \end{cases} \tag{17}$$

satisfies all hypotheses of Theorem 4. Indeed, it is straightforward to show that  $h(0) = 0$ ,  $h'(0) = -1$ ,  $h'(t) = f'(|t|)$  and that

$$\left| h'(0)^{-1} [h'(t) - h'(\tau t)] \right| = f'(|t|) - f'(\tau|t|), \quad \tau \in [0, 1], t \in (-\kappa, \kappa).$$

So,  $F = h$ ,  $n = 1$ ,  $\Omega = (-\kappa, \kappa)$  and  $x_* = 0$  satisfy all hypotheses of Theorem 4. Therefore, we can apply Theorem 4 to solve  $h(t) = 0$ .

Note that if  $f'$  is not Lipschitz, then  $h'$  is also not Lipschitz. Therefore, we conclude that Theorem 4 enlarges the range of application of theorems on inexact Newton-like Methods having the Lipschitz condition as a hypothesis on the first derivative of the non-linear operator under consideration.

Now, we will give some examples of majorant functions satisfying the conditions (h1) and (h2) with the first derivative that is not Lipschitz.

*Example 1* The following functions satisfy the conditions (h1) and (h2):

- i)  $f : [0, +\infty) \rightarrow \mathbb{R}$  such that  $f(t) = e^t - 2t - 1$ ;
- ii)  $f : [0, \kappa) \rightarrow \mathbb{R}$  such that  $f(t) = -(1/\kappa) \ln(1 - \kappa t) - 2t$ ;
- iii)  $f : [0, \kappa) \rightarrow \mathbb{R}$  such that  $f(t) = [(1 - \kappa t)/\kappa] \ln(1 - \kappa t)$ .

Note that the first derivatives of each of the functions above are not Lipschitz.

*Remark 3* The assumption (2) was crucial for our analysis. It is worth pointing out that, under appropriate regularity conditions on the nonlinear operator  $F$ , the assumption (2) always holds in a suitable neighborhood of  $x_*$ . For instance, if  $F$  is twice continuously differentiable, then the majorant function  $f : [0, \kappa) \rightarrow \mathbb{R}$  defined by  $f(t) = Kt^2/2 - t$ , where  $K = \sup\{\|F'(x_*)^{-1}F''(x)\| : x \in B[x_*, \kappa)\}$  satisfies the assumption (2). Estimating the constant  $K$  is a very difficult problem. Therefore, the goal is to identify classes of nonlinear operators for which it is possible to obtain a majorant function. We will give some examples of such classes in the next section.

### 3 Special cases

In this section we present three special cases of Theorem 4. Namely, convergence results under an affine invariant Lipschitz condition, Smale's condition for analytical functions and Nesterov-Nemirovskii's condition for self-concordant functions.

#### 3.1 Convergence result for affine invariant Lipschitz condition

In this section we show a correspondent theorem to Theorem 4 under an affine invariant Lipschitz condition (see [3, 7] and [13]) instead of the general assumption (2).

**Theorem 14** *Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow \mathbb{R}^n$  be continuously differentiable in  $\Omega$ . Take  $x_* \in \Omega$  and let*

$$\kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\}.$$

*Assume that  $F'(x_*)$  is invertible,  $F(x_*) = 0$ , and there exists a  $K > 0$  such that*

$$\left\| F'(x_*)^{-1} [F'(x) - F'(y)] \right\| \leq K \|x - y\|, \quad \forall x, y \in B(x_*, \kappa). \quad (18)$$

Take  $0 \leq \vartheta < 1, 0 \leq \omega_2 < \omega_1$  such that  $\omega_1\vartheta + \omega_2 < 1$ . Let

$$\sigma := \min \{ \kappa, 2(1 - \vartheta\omega_1 - \omega_2) / (K(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)) \}.$$

Then, the inexact Newton-like methods for solving  $F(x) = 0$ , with initial point  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F(x_k) + r_k, \quad k = 0, 1, \dots, \tag{19}$$

with the following conditions for the forcing term  $\theta_k$  and the residual  $r_k$ ,

$$0 \leq \theta_k \text{cond}(P_k F'(x_k)) \leq \vartheta, \quad \|P_k r_k\| \leq \theta_k \|P_k F(x_k)\|, \quad k = 0, 1, \dots, \tag{20}$$

where  $\{P_k\}$  is a sequence of invertible matrix (preconditioners for the linear system in (19)) and  $B(x_k)$  is an invertible approximation of  $F'(x_k)$  satisfying

$$\|B(x_k)^{-1}F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, the generated sequence  $\{x_k\}$  is contained in  $B(x_*, \sigma)$  and converges to  $x_*$ . Moreover,  $\{x_k\}$  converges linearly as follows

$$\|x_{k+1} - x_*\| \leq \left[ \omega_1(1 + \vartheta) \frac{K\|x_0 - x_*\|}{2(1 - K\|x_0 - x_*\|)} + \omega_1\vartheta + \omega_2 \right] \|x_k - x_*\|,$$

$$k = 0, 1, \dots$$

*Proof* It is immediately possible to prove that  $F, x_*$  and  $f : [0, \kappa] \rightarrow \mathbb{R}$  defined by  $f(t) = Kt^2/2 - t$ , satisfy the inequality (2) and the conditions (h1) and (h2) in Theorem 4. In this case, it is easy to see that the constants  $\rho$  and  $\nu$ , as defined in Theorem 4, satisfy

$$\rho = 2(1 - \vartheta\omega_1 - \omega_2) / (K(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)) \leq \nu = 1/K,$$

as a consequence

$$\sigma := \min \{ \kappa, 2(1 - \vartheta\omega_1 - \omega_2) / (K(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)) \}.$$

Therefore, as  $F, \sigma, f$  and  $x_*$  satisfy all of the hypotheses of Theorem 4, taking  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$  the statements of the theorem follow from Theorem 4.  $\square$

Although the condition (18) is affine invariant (it is insensitive with respect to transformation of the map  $F$  of the form  $F \mapsto AF$ ), iteration (19) and the condition for the residual (20) is not affine invariant. So, Theorem 14 is not affine invariant. Now, taking  $B(x_k) = F'(x_k)$  in iteration (19) and  $P_k = F'(x_k)^{-1}$  in the condition for the residual (20), the Theorem 14 becomes affine invariant. It is easy to see that, for the theorem that uses the Lipschitz condition

$$\|F'(x) - F'(y)\| \leq L\|x - y\|, \quad \forall x, y \in B(x_*, \kappa),$$

instead of the affine invariant Lipschitz condition (18), the convergence radius is given by

$$\sigma := \min \left\{ \kappa, 2(1 - \vartheta \omega_1 - \omega_2) / (L \|F'(x_*)\| (2 + \omega_1 - \vartheta \omega_1 - 2\omega_2)) \right\}.$$

We point out that the convergence radius of affine invariant theorems are insensitive to invertible transformation of the map  $F$ , but that theorems with the Lipschitz condition (see next example) are sensitive. For more details about affine invariant theorem see [7].

*Example 2* Assume that  $B(x_k) = F'(x_k)$ ,  $P_k = F'(x_k)^{-1}$ ,  $\omega_1 = 1$  and  $\omega_2 = 0$  in Theorem 14. Let  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be given by  $F(x_1, x_2) = (x_1^2/2 - x_1, x_2^2/2 - x_2)$ . Note that  $F(0, 0) = (0, 0)$ . Using the Euclidean vector norm and the associated operators norm, it is easy to see that  $\|F''(x_1, x_2)\| = 1$  and the Lipschitz constant for  $F'$  is 1. In this case, the Lipschitz condition for  $F'$  and the affine invariant Lipschitz condition (18) are equal. Therefore, by applying Theorem 14 we conclude that the convergence radius for solving  $F(x_1, x_2) = 0$  is  $2(1 - \vartheta)/(3 - \vartheta)$ . Let the invertible matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1/\epsilon \end{bmatrix}, \quad 0 < \epsilon < 1,$$

and the map  $G : \Omega \rightarrow \mathbb{R}^n$  given by  $G(x) = AF(x)$ . Hence  $G'(x_1, x_2) = AF'(x_1, x_2)$ . Moreover,

$$\|G'(0, 0)^{-1}\| = 1, \quad \|G''(x_1, x_2)\| = 1/\epsilon,$$

and the Lipschitz constant for  $G'$  is  $1/\epsilon$ . Applying Theorem 14 with  $B(x_k) = G'(x_k)$  and  $P_k = G'(x_k)^{-1}$ , we conclude that, due to its insensitivity to invertible transformation, the convergence radius for solving  $G(x_1, x_2) = 0$  is also  $2(1 - \vartheta)/(3 - \vartheta)$ . However, if in Theorem 14 the Lipschitz condition for  $G'$  is assumed instead of the affine invariant Lipschitz condition (18), then the convergence radius is  $(2\epsilon)(1 - \vartheta)/(3 - \vartheta)$ .

### 3.2 Convergence result under Smale’s condition

In this section we show a correspondent theorem to Theorem 4 under Smale’s condition. For more details about Smale’s condition see [15].

**Theorem 15** *Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow \mathbb{R}^n$  an analytic function. Take  $x_* \in \Omega$  such that  $F'(x_*)$  is invertible and  $F(x_*) = 0$  and let*

$$\kappa := \sup\{t > 0 : B(x_*, t) \subset \Omega\} \quad \text{and} \quad \gamma := \sup_{n>1} \left\| \frac{F'(x_*)^{-1} F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty. \tag{21}$$

Take  $0 \leq \vartheta < 1$ ,  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1\vartheta + \omega_2 < 1$ . Let  $a = \omega_1(1 + \vartheta)$ ,  $b = (1 - \omega_1\vartheta - \omega_2)$  and

$$\sigma := \min \left\{ \kappa, \frac{a + 4b - (\sqrt{(a + 4b)^2 - 8b^2})}{4b\gamma} \right\}.$$

Then, the inexact Newton-like methods for solving  $F(x) = 0$ , with initial point  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F(x_k) + r_k, \quad k = 0, 1, \dots, \tag{22}$$

with the following conditions for the forcing term  $\theta_k$  and the residual  $r_k$ ,

$$0 \leq \theta_k \text{cond}(P_k F'(x_k)) \leq \vartheta, \quad \|P_k r_k\| \leq \theta_k \|P_k F(x_k)\|, \quad k = 0, 1, \dots,$$

where  $\{P_k\}$  is a sequence of invertible matrix (preconditioners for the linear system in (22)) and  $B(x_k)$  is an invertible approximation of  $F'(x_k)$  satisfying

$$\|B(x_k)^{-1}F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, the generated sequence  $\{x_k\}$  is contained in  $B(x_*, \sigma)$  and converges to  $x_*$ . Moreover,  $\{x_k\}$  converges linearly as follows

$$\|x_{k+1} - x_*\| \leq \left[ \omega_1(1 + \vartheta) \frac{\gamma \|x_0 - x_*\|}{2(1 - \gamma \|x_0 - x_*\|)^2 - 1} + \omega_1\vartheta + \omega_2 \right] \|x_k - x_*\|, \\ k = 0, 1, \dots$$

We need the following result to prove the above theorem.

**Lemma 16** Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow \mathbb{R}^n$  an analytic function. Suppose that  $x_* \in \Omega$ ,  $F'(x_*)$  is invertible and that  $B(x_*, 1/\gamma) \subset \Omega$ , where  $\gamma$  is defined in (21). Then, for all  $x \in B(x_*, 1/\gamma)$  there holds

$$\|F'(x_*)^{-1}F''(x)\| \leq (2\gamma)/(1 - \gamma\|x - x_*\|)^3.$$

*Proof* Let  $x \in \Omega$ . Since  $F$  is an analytic function, we have

$$F'(x_*)^{-1}F''(x) = \sum_{n=0}^{\infty} \frac{1}{n!} F'(x_*)^{-1}F^{(n+2)}(x_*)(x - x_*)^n.$$

Combining (21) and the above equation we obtain, after some simple calculus, that

$$\|F'(x_*)^{-1}F''(x)\| \leq \gamma \sum_{n=0}^{\infty} (n + 2)(n + 1)(\gamma\|x - x_*\|)^n.$$

On the other hand, as  $B(x_*, 1/\gamma) \subset \Omega$  we have  $\gamma \|x - x_*\| < 1$ . So, from Proposition 2 we conclude

$$\frac{2}{(1 - \gamma \|x - x_*\|)^3} = \sum_{n=0}^{\infty} (n + 2)(n + 1)(\gamma \|x - x_*\|)^n.$$

Combining the two above equations, we obtain the desired result. □

The next result gives a condition that is easier to check than condition (2), whenever the functions under consideration are twice continuously differentiable.

**Lemma 17** *Let  $\Omega \subseteq \mathbb{R}^n$  be an open set and  $F : \Omega \rightarrow Y$  be twice continuously on  $\Omega$ . Let  $x_* \in \Omega$  with  $F'(x_*)$  be invertible. If there exists a  $f : [0, R) \rightarrow \mathbb{R}$  twice continuously differentiable such that*

$$\|F'(x_*)^{-1} F''(x)\| \leq f''(\|x - x_*\|), \tag{23}$$

for all  $x \in \Omega$  such that  $\|x - x_*\| < R$ . Then  $F$  and  $f$  satisfy (2).

*Proof* Taking  $\tau \in [0, 1]$  and  $x \in \Omega$ , such that  $x_* + \tau(x - x_*) \in \Omega$  and  $\|x - x_*\| < R$ , we obtain that

$$\begin{aligned} & \|F'(x_*)^{-1} [F'(x) - F'(x_* + \tau(x - x_*))]\| \\ & \leq \int_{\tau}^1 \|F'(x_*)^{-1} F''(x_* + t(x - x_*))\| \|x - x_*\| dt. \end{aligned}$$

Now, as  $\|x - x_*\| < R$  and  $f$  satisfies (23), we obtain from the last inequality that

$$\|F'(x_*)^{-1} [F'(x) - F'(x_* + \tau(x - x_*))]\| \leq \int_{\tau}^1 f''(t\|x - x_*\|) \|x - x_*\| dt.$$

Evaluating the latter integral, the statement follows. □

*Proof of Theorem 15* Assume that all hypotheses of Theorem 15 hold. Consider the real function  $f : [0, 1/\gamma) \rightarrow \mathbb{R}$  defined by

$$f(t) = \frac{t}{1 - \gamma t} - 2t.$$

It is straightforward to show that  $f$  is analytic and that

$$\begin{aligned} f(0) &= 0, & f'(t) &= 1/(1 - \gamma t)^2 - 2, & f'(0) &= -1, \\ f''(t) &= (2\gamma)/(1 - \gamma t)^3, & f^n(0) &= n!\gamma^{n-1}, \end{aligned}$$

for  $n \geq 2$ . From the last four equalities it is easy to see that  $f$  satisfies (h1) and (h2). Now, since  $f''(t) = (2\gamma)/(1 - \gamma t)^3$  combining Lemmas 17, 16 we conclude that  $F$  and  $f$  satisfy (2) with  $R = 1/\gamma$ . Define

$$v := \sup\{t \in [0, 1/\gamma) : f'(t) < 0\},$$



$$\rho := \sup\{t \in (0, \nu) : \omega_1(1 + \vartheta)[f(t)/(tf'(t)) - 1] + \omega_1\vartheta + \omega_2 < 1\}.$$

In this case, it is easy to see that the constants  $\nu$  and  $\rho$  satisfy

$$\rho = \frac{a + 4b - (\sqrt{(a + 4b)^2 - 8b^2})}{4b}, \quad \nu = \frac{\sqrt{2} - 1}{\sqrt{2}\gamma}, \quad \rho < \nu < \frac{1}{\gamma},$$

where  $a = \omega_1(1 + \vartheta)$ ,  $b = (1 - \omega_1\vartheta - \omega_2)$ . Finally, let  $\sigma := \min\{\kappa, \rho\}$ . Therefore, as  $F$ ,  $\sigma$ ,  $f$  and  $x_*$  satisfy all hypothesis of Theorem 4, taking  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$ , the statements of the theorem follow from Theorem 4. □

### 3.3 Convergence result under The Nesterov-Nemirovskii condition

In this section we show a correspondent theorem to Theorem 4 under the Nesterov-Nemirovskii condition (see [14]).

Let  $\Omega \subset \mathbb{R}^n$  be a convex set. A function  $g : \Omega \rightarrow \mathbb{R}$  is called  $a$ -self-concordant with the parameter  $a > 0$ , if  $g \in C^3(\Omega)$ , i.e., three times continuously differentiable in  $\Omega$ , is a convex function on  $\Omega$  and satisfies the following inequality

$$|g'''(x)[h, h, h]| \leq 2a^{-1/2}(g''(x)[h, h])^{3/2}, \quad \forall x \in \Omega, \forall h \in \mathbb{R}^n. \tag{24}$$

Take  $x_* \in \Omega$  such that  $g''(x_*)$  is invertible. Define  $X := (\mathbb{R}^n, \langle \cdot, \cdot \rangle_{x_*})$  as the Euclidean space  $\mathbb{R}^n$  with the inner product and the associated norm defined, respectively, by

$$\langle u, v \rangle_{x_*} := a^{-1} \langle g''(x_*)u, v \rangle, \quad \|u\|_{x_*} := \sqrt{\langle u, u \rangle_{x_*}},$$

where  $\langle \cdot, \cdot \rangle$  is the Euclidean inner product. So, the open and closed ball of radius  $r > 0$  centered at  $x_*$  (Dikin’s ellipsoid of radius  $r > 0$  centered at  $x_*$ ) in  $X$  are defined, respectively, as

$$W_r(x_*) := \{x \in \mathbb{R}^n : \|x - x_*\|_{x_*} < r\}, \quad W_r[x_*] := \{x \in \mathbb{R}^n : \|x - x_*\|_{x_*} \leq r\}.$$

**Theorem 18** *Let  $\Omega \subseteq X$  be a convex set and  $g : \Omega \rightarrow \mathbb{R}$  an  $a$ -self-concordant function. Take  $x_* \in \Omega$  with  $g''(x_*)$  invertible and let  $\kappa := \sup\{u > 0 : W_u(x_*) \subset \Omega\}$ . Suppose that  $g'(x_*) = 0$ . Take  $0 \leq \vartheta < 1$ ,  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1\vartheta + \omega_2 < 1$ . Let  $a = \omega_1(1 + \vartheta)$ ,  $b = (1 - \omega_1\vartheta - \omega_2)$  and*

$$\sigma := \min \left\{ \kappa, \frac{a + 4b - (\sqrt{(a + 4b)^2 - 8b^2})}{4b} \right\}.$$

*Then, the inexact Newton-like methods for solving  $F(x) = 0$ , with initial point  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$*

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F(x_k) + r_k, \quad k = 0, 1, \dots, \tag{25}$$

*with the following conditions for the forcing term  $\theta_k$  and the residual  $r_k$ ,*

$$0 \leq \theta_k \text{ cond}(P_k F'(x_k)) \leq \vartheta, \quad \|P_k r_k\| \leq \theta_k \|P_k F(x_k)\|,$$

where  $P_k$  is a invertible matrix (preconditioner for the linear system in (25)) and  $B(x_k)$  is an invertible approximation of  $F'(x_k)$  satisfying

$$\|B(x_k)^{-1} F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1} F'(x_k) - I\| \leq \omega_2,$$

is well defined, the generated sequence  $\{x_k\}$  is contained in  $B(x_*, \sigma)$  and converges to  $x_*$ . Moreover,  $\{x_k\}$  converges linearly as follows

$$\|x_{k+1} - x_*\| \leq \left[ \omega_1(1 + \vartheta) \frac{\|x_0 - x_*\|}{2(1 - \|x_0 - x_*\|)^2 - 1} + \omega_1\vartheta + \omega_2 \right] \|x_k - x_*\|,$$

$$\forall k = 0, 1, \dots$$

We need some auxiliary results about self-concordant functions to prove the above theorem. We begin with two well known propositions in the theory of self-concordant functions, from Nesterov and Nemirovskii [14].

**Proposition 19** *Let  $\Omega \subset X$  be an open convex set and let  $g : \Omega \rightarrow \mathbb{R}$  be an  $a$ -self-concordant function. Then,*

$$|g'''(x)[h_1, h_2, h_3]| \leq 2a^{-1/2} \prod_{i=1}^3 (g''(x)[h_i, h_i])^{1/2}, \quad \forall x \in \Omega, \forall h_1, h_2, h_3 \in X.$$

*Proof* See Proposition 9.1.1, Appendix 1, p. 361 of [14]. □

**Proposition 20** *Let  $\Omega \subset X$  be an open convex set and let  $g : \Omega \rightarrow \mathbb{R}$  be an  $a$ -self-concordant function. Assume that  $W_1(x_*) \subset \Omega$ . Then there holds*

$$g''(x)[h, h] \leq \frac{1}{(1 - \|x - x_*\|_{x_*})^2} g''(x_*)[h, h], \quad \forall x \in W_1(x_*), \forall h \in X.$$

*Proof* See Theorem 2.1.1 p. 13 of [14]. □

The next result is a combination of the two last propositions, which has appeared in [1, Lemma 5.1]. We include the proof here.

**Lemma 21** *Let  $\Omega \subset X$  be an open convex set and let  $g : \Omega \rightarrow \mathbb{R}$  be an  $a$ -self-concordant function. Assume that  $W_1(x_*) \subset \Omega$ . Then*

$$\|g''(x_*)^{-1} g'''(x)\|_{x_*} \leq \frac{2}{(1 - \|x - x_*\|_{x_*})^3}, \quad \forall x \in W_1(x_*).$$

*Proof* Letting  $x \in W_1(x_*)$  and  $h_1, h_2, h_3 \in \mathbb{R}^n$  we have from (24) and Proposition 19

$$\begin{aligned} |\langle g''(x_*)^{-1} g'''(x)h_1h_2, h_3 \rangle_{x_*}| &= a^{-1} |\langle g''(x_*)(g''(x_*)^{-1} g'''(x))h_1h_2, h_3 \rangle| \\ &= a^{-1} |g'''(x)[h_1, h_2, h_3]| \\ &\leq 2a^{-3/2} \prod_{i=1}^3 (g''(x)[h_i, h_i])^{1/2}. \end{aligned}$$

Since  $\|g''(x_*)^{-1}g'''(x)\|_{x_*} := \sup\{|\langle g''(x_*)^{-1}g'''(x)h_1h_2, h_3 \rangle_{x_*}| : \|h_i\|_{x_*} \leq 1, i = 1, 2, 3\}$ , we have from the last inequality

$$\|g''(x_*)^{-1}g'''(x)\|_{x_*} \leq 2a^{-3/2} \sup \left\{ \prod_{i=1}^3 (g''(x)[h_i, h_i])^{1/2} : \|h_i\|_{x_*} \leq 1, i = 1, 2, 3 \right\}. \tag{26}$$

Therefore, it follows from (26) and Proposition 20 that

$$\begin{aligned} & \|g''(x_*)^{-1}g'''(x)\|_{x_*} \\ & \leq \frac{2a^{-3/2}}{(1 - \|x - x_*\|_{x_*})^3} \sup \left\{ \prod_{i=1}^3 g''(x_*)[h_i, h_i]^{1/2} : \|h_i\|_{x_*} \leq 1, i = 1, 2, 3 \right\} \\ & = \frac{2}{(1 - \|x - x_*\|_{x_*})^3} \sup \left\{ \prod_{i=1}^3 a^{-1/2} g''(x_*)[h_i, h_i]^{1/2} : \|h_i\|_{x_*} \leq 1, i = 1, 2, 3 \right\} \\ & = \frac{2}{(1 - \|x - x_*\|_{x_*})^3} \sup \left\{ \prod_{i=1}^3 \|h_i\|_{x_*} : \|h_i\|_{x_*} \leq 1, i = 1, 2, 3 \right\} \\ & \leq \frac{2}{(1 - \|x - x_*\|_{x_*})^3}, \end{aligned}$$

which is a proof of the lemma. □

*Proof of Theorem 18* Assume that all hypotheses of Theorem 15 hold. Consider the real function  $f : [0, 1) \rightarrow \mathbb{R}$  defined by

$$f(t) = \frac{t}{1-t} - 2t.$$

It is straightforward to show that  $f$  is analytic and that

$$\begin{aligned} f(0) &= 0, & f'(t) &= 1/(1-t)^2 - 2, & f'(0) &= -1, \\ f''(t) &= 2/(1-t)^3, & f^n(0) &= n!, \end{aligned}$$

for  $n \geq 2$ . From the last four equalities, it is easy to conclude that  $f$  satisfies (h1) and (h2). Now, combining Lemmas 17, 21 and the latter equality we obtain that  $g'$  and  $f$  satisfy (2) with  $R = 1$ . Define

$$\begin{aligned} \nu &:= \sup\{t \in [0, 1) : f'(t) < 0\}, \\ \rho &:= \sup\{t \in (0, \nu) : \omega_1(1 + \vartheta)[f(t)/(tf'(t)) - 1] + \omega_1\vartheta + \omega_2 < 1\}. \end{aligned}$$

In this case, it is easy to see that the constants  $\nu$  and  $\rho$  satisfy

$$\rho = \frac{a + 4b - (\sqrt{(a + 4b)^2 - 8b^2})}{4b}, \quad \nu = \frac{\sqrt{2} - 1}{\sqrt{2}}, \quad \rho < \nu < 1,$$

where  $a = \omega_1(1 + \vartheta)$ ,  $b = (1 - \omega_1\vartheta - \omega_2)$ . Finally, let  $\sigma := \min\{\kappa, \rho\}$ . Therefore, as  $F = g'$ ,  $\sigma$ ,  $f$  and  $x_*$  since the above satisfy all hypotheses of Theorem 4, taking  $x_0 \in B(x_*, \sigma) \setminus \{x_*\}$ , the statements of the theorem follow from Theorem 4.  $\square$

#### 4 Final remarks

As pointed out by Morini in [13] if preconditioning  $P_k$ , satisfying

$$\|P_k r_k\| \leq \theta_k \|P_k F(x_k)\|, \quad (27)$$

for some forcing sequence  $\{\theta_k\}$ , is applied in finding the inexact Newton steep, then the inverse proportionality between each forcing term  $\theta_k$  and  $\text{cond}(P_k F'(x_k))$  stated in the following assumption:

$$0 < \theta_k \text{cond}(P_k F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots, \quad (28)$$

is sufficient to guarantee convergence, and may be overly restrictive to bound the sequence  $\{\theta_k\}$ , always such that the matrices  $P_k F'(x_k)$ , for  $k = 0, 1, \dots$ , are badly conditioned. Moreover,  $\theta_k$  does not depend on  $\text{cond}(P_k)$  but only on the  $\text{cond}(P_k F'(x_k))$  and a suitable choice of scaling matrix  $P_k$  leads to a relaxation of the forcing terms.

Using the assumptions (27) and (28), we presented a new local convergence analysis for inexact Newton-like methods under majorant condition. In our analysis, the affine invariant Lipschitz condition (see [3, 7] and [13]) is relaxed by using the majorant condition (see (2) in Theorem 4). Although the condition (2) is equivalent to the Chen and Li condition (see (1.4) in [4]), our analysis elucidates the relationship of the majorant function with the non-linear operator under consideration (see Lemma 10). In addition, Lemmas 11, 12 and 13 allow us to obtain the special cases (Theorems 14, 15 and 18) of Theorem 4 as an application.

Finally, we point out that the Kantorovich analysis produced a semilocal convergence result, in that it ensures convergence of Newton's method under very mild assumptions and proves the existence of a solution. On the other hand, local analysis gives us the optimal convergence radius.

**Acknowledgements** The author would like to thank the two anonymous referees for their constructive comments that have helped to substantially improve the presentation of the paper.

#### 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. Blum, L., Cucker, F., Shub, M., Smale, S.: *Complexity and Real Computation*. Springer, New York (1997)
3. Chen, J.: The convergence analysis of inexact Gauss-Newton methods for nonlinear problems. *Comput. Optim. Appl.* **40**(1), 97–118 (2008)
4. Chen, J., Li, W.: Convergence behaviour of inexact Newton methods under weak Lipschitz condition. *J. Comput. Appl. Math.* **191**(1), 143–164 (2006)
5. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. *SIAM J. Numer. Anal.* **19**(2), 400–408 (1982)

6. Dennis, J.E., Schnabel, R.B.: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs (1983)
7. Deuffhard, P., Heindl, G.: Affine invariant convergence for Newton's method and extensions to related methods. *SIAM J. Numer. Anal.* **16**(1), 1–10 (1979)
8. Ferreira, O.P.: Local convergence of Newton's method in Banach space from the viewpoint of the majorants principle. *IMA J. Numer. Anal.* (2008, to appear). doi:[10.1093/imanum/drn036](https://doi.org/10.1093/imanum/drn036)
9. Ferreira, O.P., Svaiter, B.F.: Kantorovich's majorants principle for Newton's method. *Comput. Optim. Appl.* **42**, 213–229 (2009)
10. Hiriart-Urruty, J.-B., Lemaréchal, C.: *Convex Analysis and Minimization Algorithms I*. Springer, Berlin (1993)
11. Martinez, J.M., Qi, L.: Inexact Newton methods for solving nonsmooth equations. *J. Comput. Appl. Math.* **60**(1–2), 127–145 (1995)
12. Moret, I.: A Kantorovich-type theorem for inexact Newton methods. *Numer. Funct. Anal. Optim.* **10**(3–4), 351–365 (1989)
13. Morini, B.: Convergence behaviour of inexact Newton methods. *Math. Comput.* **68**(228), 1605–1613 (1999)
14. Nesterov, Y., Nemirovskii, A.: *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics, vol. 13. SIAM, Philadelphia (1994)
15. Smale, S.: Newton 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)
16. Traub, J.F., Wozniakowski, H.: Convergence and complexity of Newton iteration for operator equation. *J. Assoc. Comput. Mach.* **26**(2), 250–258 (1979)
17. Wang, X.: Convergence of Newton methods and uniqueness of the solution of equations in Banach space. *IMA J. Numer. Anal.* **20**(1), 123–134 (2000)
18. Wu, M.: A convergence theorem for the Newton-like methods under some kind of weak Lipschitz conditions. *J. Math. Anal. Appl.* **339**(2), 1425–1431 (2008)
19. Ypma, T.J.: Affine invariant convergence results for Newton's methods. *BIT* **22**(1), 108–118 (1982)
20. Ypma, T.J.: Local convergence of inexact Newton methods. *SIAM J. Numer. Anal.* **21**(3), 583–590 (1984)