

A perturbed version of an inexact generalized Newton method for solving nonsmooth equations

Marek J. Śmietański

Received: 30 January 2012 / Accepted: 14 June 2012 /

Published online: 1 July 2012

© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract In this paper, we present the combination of the inexact Newton method and the generalized Newton method for solving nonsmooth equations $F(x) = 0$, characterizing the local convergence in terms of the perturbations and residuals. We assume that both iteration matrices taken from the B-differential and vectors $F(x^{(k)})$ are perturbed at each step. Some results are motivated by the approach of Cătiuaș regarding to smooth equations. We study the conditions, which determine admissible magnitude of perturbations to preserve the convergence of method. Finally, the utility of these results is considered based on some variant of the perturbed inexact generalized Newton method for solving some general optimization problems.

Keywords Nonsmooth equations · Inexact Newton method · Inexact generalized Newton method · Perturbation · B-differential · Superlinear convergence

Mathematics Subject Classifications (2010) 65H10 · 49M15

1 Introduction

We consider the system of nonlinear equations

$$F(x) = 0, \tag{1}$$

M. J. Śmietański (✉)
Faculty of Mathematics and Computer Science, University of Łódź,
Banacha 22, 90-238 Łódź, Poland
e-mail: smietan@math.uni.lodz.pl

where F is a given Lipschitz continuous function from R^n into R^n . In the whole paper we assume that there exists $x^* \in R^n$ such that $F(x^*) = 0$.

A locally convergent iterative Newton method has the general form

$$x^{(k+1)} = x^{(k)} + s^{(k)},$$

where the step $s^{(k)}$ is a solution of some particular linear system. The classical Newton method (using the usual Jacobian of F at $x^{(k)}$) for solving smooth equations is locally and quadratically convergent to x^* under the standard assumptions: existence of a solution, continuously differentiability of F and nonsingularity of the Jacobian at the solution point (to guarantee that the increment $s^{(k)}$ exists in the neighborhood of the solution).

The superlinearly convergent generalized Newton method, which was introduced by Qi [26], has the following form

$$V_k s^{(k)} = -F(x^{(k)}), \quad V_k \in \partial_B F(x^{(k)}),$$

where $\partial_B F(x^{(k)})$ is the B-differential of F at $x^{(k)}$ and the matrix V_k is taken arbitrarily from $\partial_B F(x^{(k)})$. The iterative process is the classical Newton method if F is differentiable, because $\partial_B F(x) = \{F'(x)\}$, where $F'(x)$ denotes the Jacobian of F at x .

In turn, the inexact Newton method for solving smooth equations introduced by Dembo, Eisenstat and Steihaug in [8] is given by

$$F'(x^{(k)}) s^{(k)} = -F(x^{(k)}) + r_k$$

with

$$\|r_k\| \leq \eta_k \|F(x^{(k)})\|,$$

where $\{\eta_k\}$ is the forcing sequence such that $0 \leq \eta_k < 1$. From the computational point of view, we can treat the error term r_k as the perturbation of F at $x^{(k)}$.

The combination of the inexact Newton method and the quasi-Newton method for solving smooth equations has been described in several papers, e.g. [1, 2, 14]. Furthermore, the nonsmooth versions of the inexact generalized Newton method were considered among the others in [3, 15, 23, 34, 37]. Both the generalized Newton methods and the inexact Newton methods for solving nonsmooth equations are locally and superlinearly convergent to the solution under mild conditions. The main goal of this paper is to analyze the convergence of the perturbed version of the inexact generalized Newton method. Is the behavior of this method the same and what conditions on the perturbations do we have to guarantee for a good convergence?

Ypma [36] analyzed the rounding errors of the sequence of iterates generated by the undamped Newton-like method as the perturbations incorporated in linear system, however he did not give the explicit formulas regarding to the perturbations and residuals. The sufficient condition for local convergence of the perturbed Newton-like method for solving smooth equations was proved, but under rather strong assumptions imposed on F . Next, Cătinăș [4]

extended the local convergence theory of the inexact Newton method in the smooth case

$$[F'(x^{(k)}) + \Delta_k]s^{(k)} = [-F(x^{(k)}) + \delta_k] + \hat{r}_k, \quad (2)$$

characterizing the rate of convergence in terms of perturbations and residuals. In the iteration above, Δ_k and δ_k represent the perturbations, respectively to the Jacobians and the function evaluations, while \hat{r}_k are the residuals of the approximate solutions $s^{(k)}$ of the corresponding linear systems. The method (2) is linearly convergent for continuously differentiable equations defined by functions with nonsingular Jacobian at the solution point. The superlinear convergence is possible to attain with some strong condition imposed on perturbations. Cătinaş showed also that the inexact, inexact perturbed and quasi-Newton methods for solving smooth equations are equivalent models based on the characterization of convergence orders [5].

In this paper, we analyze the perturbed inexact generalized Newton method in the form

$$(V_k + \tilde{V}_k)s^{(k)} = [-F(x^{(k)}) + \tilde{F}(x^{(k)})] + \tilde{r}_k,$$

where \tilde{V}_k is some perturbation matrix, \tilde{F} is the perturbation of function F and \tilde{r}_k denotes the residual. Some results were obtained based on known theorems from [4, 15] and [7], however some extension of the relevant statements to the nonsmooth case was one of the main objectives of paper inspired by perturbed approach.

Really, application of the considered methods is much wider than solving nonsmooth equations and equations which arise from the reformulation of nonlinear complementarity problems and variational inequalities. There exist numerous schemes which allow generalization methods for solving univariate global optimization problems to the multidimensional case (including one-point based, diagonal, simplicial, space-filling curves and other popular approaches, see e.g. [11, 19]). Because in many problems functions may not possess a sufficient degree of smoothness, algorithms for solving nondifferentiable unconstrained problems can be important tools to solve them. For example, the univariate constrained minimization problem with inequality constraints can be simply reduced to the unconstrained one by using the penalty scheme $f_P(x) = f_0(x) + P \max\{g_1(x), g_2(x), \dots, g_m(x), 0\}$ with the penalty coefficient P (see [10]).

Chemical engineering, electronics and electrotechnics are among the fields where such methods can be used successfully (including [20, 30, 32]). The effective Newton-like methods for solving some multidimensional optimization problems with nonsmooth gradient were also introduced for LC^1 optimization problem (e.g. in [24, 27]) and for convex SC^1 minimization problem (e.g. in [18]). Some examples of source problems for the latter one are nonlinear minimax problem with a convex-concave saddle function and some nonlinear stochastic program. In turn, di Pillo (including in [21, 22]) presented some constructions of the unconstrained optimization problem with

the objective function, which may not be twice differentiable, by using penalty functions or Lagrange multiplier functions. Moreover, the promising results of solving unconstrained optimization problems by using the perturbed Newton-like method presented by Grapsa, Antonelou and Kostopulos [13] were an additional motivation for the study of the perturbed approach.

The paper is organized as follows. In Section 2, we recall some notions and we summarize some properties, which are needed for the rest of this paper. In Section 3, we analyze the convergence of the perturbed inexact generalized Newton method and we study some convergence conditions. In Section 4, we give other approach to the characterization of the convergence rate based on some general perturbation lemma. In Section 5, we apply the specific version of the perturbed method to solve some particular system of equations arising from nonlinear optimization problem and we present some numerical results.

2 Preliminaries

Throughout this paper, we regard vectors in R^n as a column vector. We use the Euclidean norm on R^n denoted by $\|\cdot\|$, together with its induced operator norm. However, it is easy to verify that results are independent of this choice. For a differentiable function $F : R^n \rightarrow R^n$, $F'(x)$ denotes the usual Jacobian matrix of the partial derivatives whenever x is a point, at which the necessary partial derivatives exist and D_F denotes the set where F is differentiable.

Let the function F be a Lipschitz continuous in the traditional sense, i.e. there exists $L \geq 0$ such that, for any $x, y \in D \subset R^n$, it holds that

$$\|F(x) - F(y)\| \leq L \|x - y\|.$$

According to Rademacher's Theorem [29], the Lipschitz continuity of F implies that F is differentiable almost everywhere. Then, the set

$$\partial_B F(x) = \left\{ \lim_{x_i \rightarrow x} F'(x_i), x_i \in D_F \right\}$$

is called the B-differential (the Bouligand subdifferential) of F at x (introduced by Qi [26]). The generalized Jacobian of F at x in the sense of Clarke [6] is

$$\partial F(x) = \text{conv } \partial_B F(x),$$

where conv denotes the convex hull.

We say that F is BD-regular at x , if all $V \in \partial_B F(x)$ are nonsingular.

Remark The above property is also called the strongly BD-regularity in some papers.

Proposition 1 (Martínez and Qi [15], Proposition 1) *If F is BD-regular at x , then $\mu = \max\{\|V^{-1}\| : V \in \partial_B F(x)\}$ exists and, for any given $\varepsilon > 0$, there exists a neighborhood $N(x)$ of x such that F is BD-regular at $N(x)$ and $\|V^{-1}\|$ is uniformly bounded by $\mu + \varepsilon$ for all $V_y \in \partial_B F(y)$ and all y in $N(x)$.*

In the whole work we assume that F satisfies the following conditions:

- (i) $F : D \rightarrow R^n$, where $D \subset R^n$ is an open set, is Lipschitz continuous;
- (ii) There exists an $x^* \in D$ such that $F(x^*) = 0$, i.e. x^* is a solution of (1);
- (iii) $N(x^*, r) = \{x \in R^n : \|x - x^*\| \leq r\} \subset D$ is some neighborhood of x^* and x^* is the only solution of (1) in $N(x^*, r)$.

3 Convergence of the perturbed inexact generalized Newton method

First, we show the local convergence of the Newton-like method $x^{(k+1)} = x^{(k)} + s^{(k)}$, where the step $s^{(k)}$ satisfies

$$V_k s^{(k)} = -F(x^{(k)}) + r_k, \quad V_k \in \partial_B F(x^{(k)}), \quad (3)$$

with

$$\|r_k\| \leq \eta_k \|F(x^{(k)})\|, \quad (4)$$

for some sequences $\{r_k\} \subset R^n$ and $\{\eta_k\} \subset R$ such that $\eta_k \leq \bar{\eta} < 1$. Naturally, this is the inexact generalized Newton method. The fundamental version of such method for solving nonsmooth equations was presented by Martínez and Qi [15]. The superlinear convergence of the method using the B-differential has been proved under semismoothness and BD-regularity assumptions. Pu and Tian [23] presented some practical version of method (with the generalized Jacobian), which is also superlinearly convergent. Some substantial extension of method with the B-differential was also given in [35], where additionally a globally convergent hybrid method with Armijo line search was presented. In turn, Bonettini and Tinti [3] proposed a nonmonotone variant of the inexact generalized Newton method with backtracking strategy for solving semismooth equations. Moreover, in [34] we introduce some parameterized version of the method described by (3) and (4) with the fixed forcing terms η_k for solving constrained nonsmooth equations.

Besides the usually considered assumptions we need some additional condition to guarantee the local convergence of the method. We can use one of the following:

Assumption A1 We say that Lipschitz continuous function F satisfies A1 at x , if, for any given $\gamma > 0$, there exists a constant $r > 0$, such that for any $y \in N(x, r)$ and any $V_y \in \partial_B F(y)$, it holds

$$\|F(y) - F(x) - V_y(y - x)\| \leq \gamma \|y - x\|. \quad (5)$$

Assumption A2 We say that Lipschitz continuous function F satisfies A2 at x , if there exists constant $r > 0$ such that, for any $y \in N(x, r)$ and any $V_y \in \partial_B F(y)$, it holds

$$F(y) - F(x) - V_y(y - x) = o(\|y - x\|). \quad (6)$$

Remark

- (i) Clearly, it is easy to notice that the condition (6) is really stronger than (5). On the other hand, the Assumption A2 is also weaker than strongly monotonicity and BD-regularity of F at x required together (see [23]). Moreover, if we use the condition stronger than Assumption A1 at x^* , then we obtain the higher convergence order of the method.
- (ii) Semismoothness (introduced by Mifflin [16]), C-differentiability (introduced by Qi [28]) and H-differentiability (introduced by Gowda and Ravindran [12]) are properties that imply A2. However, Gowda and Ravindran [12] showed that either the Clarke generalized Jacobian of a Lipschitz continuous function, the B-differential of a semismooth function (in particular, piecewise affine or piecewise smooth) and C-differential of a C-differentiable function are particular instances of the H-differential.

Now, we can state the following theorem, which gives the sufficient conditions for the local convergence of the inexact generalized Newton method for solving nonsmooth equations.

Theorem 2 *Assume that F is BD-regular at x^* and F satisfies Assumption A1 at x^* . Then, there exists a positive number $r > 0$ such that, if $x^{(0)} \in N(x^*, r)$ then the sequence generated by the method (3) satisfying (4) with $0 \leq \eta_k \leq \bar{\eta} < 1$ for $k = 0, 1, \dots$ is well-defined and converges linearly to x^* .*

Moreover, if F satisfies Assumption A2 at x^ and $\eta_k \rightarrow 0$ as $k \rightarrow \infty$, then $\{x^{(k)}\}$ converges superlinearly to x^* .*

Proof Since F is Lipschitz continuous, there exists constant $L \geq 0$ such that

$$\|F(x)\| \leq L \|x - x^*\| \quad (7)$$

for any $x \in D$.

Now, we claim that, for any given $\gamma > 0$, there exist positive numbers r and μ' such that, for any $y \in N(x^*, r)$ and any $V_y \in \partial_B F(y)$, we have

$$\|F(y) - F(x^*) - V_*(y - x^*)\| \leq \gamma \|y - x^*\| \quad (8)$$

and

$$\|V_y^{-1}\| \leq \mu'. \quad (9)$$

Inequality (8) is implied by Assumption A1. Inequality (9) is obtained by BD-regularity of F at x^* and Proposition 1.

Furthermore, since

$$\begin{aligned} x^{(k+1)} - x^* &= x^{(k)} - x^* + V_k^{-1} (-F(x^{(k)}) + r_k) \\ &= V_k^{-1} \cdot [r_k + V_k(x^{(k)} - x^*) - (F(x^{(k)}) - F(x^*))], \end{aligned}$$

taking norms, we obtain

$$\begin{aligned}\|x^{(k+1)} - x^*\| &\leq \|V_k^{-1}\| \cdot (\|r_k\| + \|F(x^{(k)}) - F(x^*) - V_k(x^{(k)} - x^*)\|) \\ &\leq \mu' (\eta_k \|F(x^{(k)})\| + \gamma \|x^{(k)} - x^*\|) \\ &\leq \mu' (\eta_k L + \gamma) \|x^{(k)} - x^*\|.\end{aligned}$$

using (9), (4), (8) and (7). If $\eta_k \leq \bar{\eta} < \frac{1}{L\mu'}$ and $\gamma < \frac{1}{\mu'} - \bar{\eta}L$, then the method (3) with condition (4) is linearly convergent to x^* . Clearly, the sequence generated by the considered method is well-defined in the neighborhood of x^* , i.e. for $x^{(k)} \in N(x^*, r)$.

Assumption A2 and convergence $\{\eta_k\}$ to 0 obviously imply the superlinear convergence based on the above considerations. \square

The above theorem is the generalized nonsmooth version of theorem presented by Dembo et al. in [8] for the smooth inexact Newton method. However, in [8] the convergence of the method has been obtained only in the special norm $\|y\|_* = \|F'(x^*)y\|$. In contrast, our results are norm-independent. On the other hand, the similar theorem was proved in [15], however only for the semismooth equations.

Now, we consider the perturbed inexact generalized Newton method in the following form

$$(V_k + \tilde{V}_k)s^{(k)} = [-F(x^{(k)}) + \tilde{F}(x^{(k)})] + \tilde{r}_k, \quad (10)$$

where \tilde{V}_k is an $n \times n$ perturbation matrix, $\tilde{F} : R^n \rightarrow R^n$ is a perturbation of function F and \tilde{r}_k denotes a residual of the approximate solution $s^{(k)}$ of the linear systems

$$(V_k + \tilde{V}_k)s^{(k)} = -F(x^{(k)}) + \tilde{F}(x^{(k)}). \quad (11)$$

If we assume that the systems (11) are solved exactly, i.e. $\tilde{r}_k = 0$ for all $k = 0, 1, \dots$, then the perturbed inexact generalized Newton method simplifies to the perturbed exact one. Some other approach to the exact version of the perturbed method will be consider in the next section.

Using the previous theorem, we can easy prove the following convergence property of the method (10) (similarly as in [5]).

Theorem 3 *Assume that F is BD-regular at x^* and F satisfies Assumption A1 at x^* . Moreover, we suppose that the perturbations \tilde{V}_k are such that the matrices $V_k + \tilde{V}_k$ are nonsingular for $k = 0, 1, \dots$. Then, there exists $\varepsilon > 0$ such that, if $x^{(0)} \in N(x^*, \varepsilon)$ and*

$$\begin{aligned}&\left\| \tilde{V}_k (V_k + \tilde{V}_k)^{-1} F(x^{(k)}) + \left[I - \tilde{V}_k (V_k + \tilde{V}_k)^{-1} \right] (\tilde{F}(x^{(k)}) + \tilde{r}_k) \right\| \\ &\leq \eta_k \|F(x^{(k)})\|\end{aligned}$$

for $k = 0, 1, \dots$, then the sequence $\{x^{(k)}\}$ generated by the method (10) is linearly convergent to x^* .

Moreover, if $\eta_k \rightarrow 0$ as $k \rightarrow \infty$, then $\{x^{(k)}\}$ converges superlinearly to x^* .

Proof The perturbed inexact generalized Newton method (10) can be viewed as the inexact generalized Newton method without perturbation, i.e. the method, in which

$$s^{(k)} = \left(V_k + \tilde{V}_k \right)^{-1} \left[-F(x^{(k)}) + \tilde{F}(x^{(k)}) + \tilde{r}_k \right].$$

So, we have

$$\begin{aligned} V_k s^{(k)} &= -\tilde{V}_k s^{(k)} - F(x^{(k)}) + \tilde{F}(x^{(k)}) + \tilde{r}_k \\ &= \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} F(x^{(k)}) - F(x^{(k)}) \\ &\quad - \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} \left[\tilde{F}(x^{(k)}) + \tilde{r}_k \right] + \tilde{F}(x^{(k)}) + \tilde{r}_k \\ &= \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} F(x^{(k)}) - F(x^{(k)}) \\ &\quad + \left[I - \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} \right] \left[\tilde{F}(x^{(k)}) + \tilde{r}_k \right]. \end{aligned}$$

Taking

$$r_k = \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} F(x^{(k)}) + \left[I - \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} \right] \left[\tilde{F}(x^{(k)}) + \tilde{r}_k \right], \quad (12)$$

we obtain the conclusion from Theorem 2. \square

Similarly as Cătiņaș [4] we can formulate the following corollary for our perturbed nonsmooth method.

Corollary 4 Suppose that the assumptions of Theorem 3 are fulfilled. Then, there exists $\varepsilon > 0$ such that, if $x^{(0)} \in N(x^*, \varepsilon)$ and

$$\left\| \tilde{V}_k \left(V_k + \tilde{V}_k \right)^{-1} \right\| \leq q_1 < 1 \text{ for } k = 0, 1, 2, \dots,$$

$$\left\| \tilde{F}(x^{(k)}) \right\| + \left\| \tilde{r}_k \right\| \leq \frac{\eta_k}{1 + q_1} \left\| F(x^{(k)}) \right\|,$$

$$\text{where } \eta_k \leq q_2 < 1 - q_1, \quad k = 0, 1, 2, \dots,$$

then the sequence generated by method (10) is linearly convergent to x^* with the asymptotic error constant $q_1 + q_2$, i.e.

$$\left\| x^{(k+1)} - x^* \right\| \leq (q_1 + q_2) \left\| x^{(k)} - x^* \right\|.$$

Moreover, if $\eta_k \rightarrow 0$ as $k \rightarrow \infty$, then $\{x^{(k)}\}$ converges superlinearly to x^* .

Proof The proof is the immediate consequence of the previous theorem. \square

Theorem 3 shows that the perturbed inexact generalized Newton method is locally convergent. The convergence rate of the method can be also characterized in terms of the rate of the relative residuals. Dembo et al. [8] proved that the usual inexact Newton method for solving smooth equations is superlinearly convergent to the solution of nonlinear equation if and only if $\|r_k\| = o(\|F(x^{(k)})\|)$ as $k \rightarrow \infty$. The same equivalence holds for our perturbed nonsmooth method, however under Assumption A2, because A1 is not strong enough. However, if we assume A2 at x^* , then we simultaneously obtain the higher convergence rate without the additional condition for the sequence of forcing terms $\{\eta_k\}$.

To prove the superlinear convergence of our method, the following lemma (which is similar to Lemma 2 in [34]) will be needed.

Lemma 5 *Let*

$$l = \max \left\{ 2\beta, \frac{1}{2\beta} + \|V_y\| \right\},$$

where

$$\beta = \max_{y \in N(x^*)} \|V_y^{-1}\|,$$

$V_y \in \partial_B F(y)$ and $N(x^*)$ is some neighborhood of x^* . If F is BD-regular at x^* and satisfies Assumption A2 at x^* then

$$\frac{1}{l} \|y - x^*\| \leq \|F(y)\| \leq l \|y - x^*\|$$

for all $y \in N(x^*)$.

Proof The proof is almost the same as the one for Lemma 2 in [34]. \square

First, we characterize the convergence order of the inexact generalized Newton method:

Theorem 6 *Assume that F is BD-regular at x^* and F satisfies Assumption A2 at x^* . If the sequence generated by method (3) satisfying (4) with $0 \leq \eta_k \leq \bar{\eta} < 1$ for $k = 0, 1, \dots$ converges to x^* , then the convergence is superlinear if and only if*

$$\|r_k\| = o(\|F(x^{(k)})\|) \text{ as } k \rightarrow \infty.$$

Proof Assume that the sequence $\{x^{(k)}\}$ converges superlinearly to x^* as $k \rightarrow \infty$. Since

$$r_k = F(x^{(k)}) - F(x^*) - V_k(x^{(k)} - x^*) + V_k(x^{(k+1)} - x^*),$$

taking norms, we obtain

$$\begin{aligned}\|r_k\| &\leq \|F(x^{(k)}) - F(x^*) - V_k(x^{(k)} - x^*)\| + \|V_k\| \|x^{(k+1)} - x^*\| \\ &= o(\|x^{(k)} - x^*\|)\end{aligned}$$

by A2 and the assumption that $x^{(k)} \rightarrow x^*$ superlinearly.

Hence, we have that

$$\|r_k\| = o(\|x^{(k)} - x^*\|) = o(\|F(x^{(k)})\|) \text{ as } k \rightarrow \infty,$$

by Lemma 5.

Conversely, assume that $\|r_k\| = o(\|F(x^{(k)})\|)$. As in the proof of Theorem 2,

$$\begin{aligned}\|x^{(k+1)} - x^*\| &\leq \|V_k^{-1}\| \cdot [\|r_k\| + \|F(x^{(k)}) - F(x^*) - V_k(x^{(k)} - x^*)\|] \\ &= \|V_k^{-1}\| \cdot [o(\|F(x^{(k)})\|) + o(\|x^{(k)} - x^*\|)]\end{aligned}$$

by the assumption that $\|r_k\| = o(\|F(x^{(k)})\|)$ and A2. Hence, it holds

$$\|x^{(k+1)} - x^*\| = o(\|x^{(k)} - x^*\|) \text{ as } k \rightarrow \infty$$

by BD-regularity of F at x^* , Proposition 1 and Lemma 5. \square

Clearly, the last theorem characterizes the convergence order of the inexact generalized Newton method in terms of the convergence rate of the relative residuals. In other words, in terms of the steps $s^{(k)}$, the sequence $\{x^{(k)}\}$ converges superlinearly to x^* if and only if

$$\|r_k\| = o(\|s^{(k)}\|) \text{ as } k \rightarrow \infty.$$

Now, we can establish the final result for the perturbed version of the inexact generalized Newton method. The next theorem gives both necessary and sufficient condition for superlinear convergence of the method.

Theorem 7 Assume that F is BD-regular at x^* and F satisfies Assumption A2 at x^* . Moreover, we suppose that the perturbations \tilde{V}_k are such that the matrices $V_k + \tilde{V}_k$ are nonsingular for $k = 0, 1, \dots$. Then, if the sequence $\{x^{(k)}\}$ generated by (10) converges to x^* , then the convergence is superlinear if and only if

$$\begin{aligned}&\left\| \tilde{V}_k (V_k + \tilde{V}_k)^{-1} F(x^{(k)}) + \left[I - \tilde{V}_k (V_k + \tilde{V}_k)^{-1} \right] (\tilde{F}(x^{(k)}) + \tilde{r}_k) \right\| \\ &= o(\|F(x^{(k)})\|) \text{ as } k \rightarrow \infty.\end{aligned}$$

Proof The proof is obvious based on the previous theorem, using (12). \square

Similarly as Cătinaș [4], we can formulate the corollary which characterizes the convergence order of the perturbed inexact generalized Newton method through the convergence of perturbations and residuals.

Corollary 8 *Suppose that the assumptions of Theorem 7 are fulfilled. Moreover, we assume that $\tilde{V}_k \rightarrow 0$, $\tilde{F}(x^{(k)}) \rightarrow 0$ and $\tilde{r}_k \rightarrow 0$ as $k \rightarrow \infty$. If the sequence $\{x^{(k)}\}$ generated by (10) converges to x^* and*

$$\|\tilde{F}(x^{(k)})\| = o(\|F(x^{(k)})\|) \quad \text{and} \quad \|\tilde{r}_k\| = o(\|F(x^{(k)})\|) \quad \text{as } k \rightarrow \infty,$$

then the convergence is superlinear.

At the end, it should be noticed that in method (10) invertibility of the perturbed matrices taken from the B-differential is not explicitly requested at each iteration step. Therefore, Theorem 7 may be restated by requirement that only the corresponding iterates were well-defined, i.e. the linear systems (11) have to be solvable. Following this idea introduced by Cătiuaș [5], Theorem 6 can be retrieved from the following extension of the second part of Theorem 7.

Theorem 9 *Assume that the method (10) is well-defined and convergent to x^* . Then, the convergence is superlinear if and only if*

$$\|-\tilde{V}_k s^{(k)} + \tilde{F}(x^{(k)}) + \tilde{r}_k\| = o(\|F(x^{(k)})\|).$$

Proof Writing the method (10) as the method (11) with residuals $r_k = -\tilde{V}_k s^{(k)} + \tilde{F}(x^{(k)}) + \tilde{r}_k$, we obtain needed result from Theorem 6. \square

4 Another approach to the convergence of the perturbed generalized Newton method

It should be emphasized that Corollary 8 gives the convergence characterization of the inexact method, when the residuals and perturbations converge to zero. Recall from the previous section: if we assume in advance that all residuals are equal to zero in (10), then we obtain the (exact) perturbed generalized Newton method (11). Clearly, suitable theorems and corollaries are also true for such version of method, if we really provide that residuals are zeros.

In this section, we present other approach to the convergence of the perturbed generalized Newton method, which is motivated by some known perturbation lemma. The result is the very general extension of Lemma 11.2.2. presented by Ortega and Rheinboldt [17]. On the other hand, our lemma is also the nonsmooth version of the perturbation lemma, which was proved by Cores and Tapia [7]. However, the conditions for the perturbations in the Newton method for solving smooth equations were given in [7], so we analyze convergence rate of the method (11). To obtain the widest possible generalization, we consider perturbations of both iteration matrix taken from the B-differential and the right-hand side of the Newton linear system.

Assumption A3 We say that F satisfies A3 at x with degree p , if there exists constant $r > 0$ such that, for any $y \in N(x, r)$ and any $V_y \in \partial_B F(y)$, it holds

$$F(y) - F(x) - V_y(y - x) = O(\|y - x\|^p). \quad (13)$$

Remark

- (i) Recall, that a function $g(x)$ is big O of $\|x - x^*\|$, i.e. $g(x) = O(\|x - x^*\|)$, if there exist a positive constant C and a neighborhood $N(x^*)$ of x^* such that

$$\|g(x)\| \leq C \|x - x^*\| \text{ for all } x \in N(x^*).$$

- (ii) Clearly, the Assumption A3 is stronger than A2.
 (iii) p -order semismoothness (introduced by Qi and Sun [25]) is property, which implies A3. In turn, strongly C -differentiability (introduced by Qi [27]) is property, which implies A3 with degree 2. Additionally, note that if the B-derivative of F is Lipschitzian then F is strongly semismooth (see [25], Proposition 3.5), which also implies A3 with order 2.

Lemma 10 Assume that the function F satisfies A3 with degree p in an open convex set D containing x^* , F is BD-regular at x^* and in a neighborhood $D_* \subset D$ of the solution x^* we have

$$\begin{aligned} \|\tilde{F}(x)\| &= O(\|x - x^*\|^q) \text{ for some } q > 0, \\ \|\tilde{V}_x\| &= O(\|x - x^*\|^r) \text{ for some } r > 0 \end{aligned}$$

for all $\tilde{V}_x \in \partial_B F(x)$. Then, there exist a neighborhood $N(x^*)$ of x^* contained in D_* and positive constants C_1, C_2, C_3, C_4 such that, for all $x \in N(x^*)$ and $x^{(k+1)}$ given by (11), we have

$$\begin{aligned} \|x^{(k+1)} - x^*\| &\leq C_1 \|x^{(k)} - x^*\|^{p+1} + C_2 \|x^{(k)} - x^*\|^q \\ &\quad + C_3 \|x^{(k)} - x^*\|^{r+1} + C_4 \|x^{(k)} - x^*\|^{q+r}. \end{aligned} \quad (14)$$

Proof First, choose $N(x^*, r)$ (i.e. $r > 0$) such that, for all $x \in N(x^*, r)$, we have

$$\|V_x^{-1}\| \|\tilde{V}_x\| \leq \delta < 1 \text{ for some } \delta > 0.$$

Now, from Theorem 3.6 by Stewart [33] we have that, for $x \in N(x^*, r)$, the matrix $V_x + \tilde{V}_x$ is invertible and

$$[V_x + \tilde{V}_x]^{-1} = [I + W_x] V_x^{-1},$$

for some W_x satisfying

$$\|W_x\| \leq \frac{\|V_x^{-1}\| \|\tilde{V}_x\|}{1 - \|V_x^{-1}\| \|\tilde{V}_x\|}. \quad (15)$$

If we denote $W_k = W_{x^{(k)}}$, we obtain

$$x^{(k+1)} - x^* = x^{(k)} - x^* - (I + W_k) V_k^{-1} \left[F(x^{(k)}) + \tilde{F}(x^{(k)}) \right].$$

Hence, it follows that

$$\begin{aligned} & \|x^{(k+1)} - x^*\| \\ &= \|x^{(k)} - x^* - (I + W_k) V_k^{-1} [F(x^{(k)}) + \tilde{F}(x^{(k)})]\| \\ &\leq \|V_k^{-1}\| \|V_k(x^{(k)} - x^*) - (I + W_k) [F(x^{(k)}) + \tilde{F}(x^{(k)})]\| \\ &= \|V_k^{-1}\| \|V_k(x^{(k)} - x^*) - F(x^{(k)}) - \tilde{F}(x^{(k)}) - W_k F(x^{(k)}) - W_k \tilde{F}(x^{(k)})\| \\ &\leq \|V_k^{-1}\| \left[\|F(x^*) - F(x^{(k)}) - V_k(x^{(k)} - x^*)\| + \|\tilde{F}(x^{(k)})\| \right. \\ &\quad \left. + \|W_k\| \|F(x^*) - F(x^{(k)})\| + \|W_k\| \|\tilde{F}(x^{(k)})\| \right]. \end{aligned}$$

Inequality (14) follows from the last inequality, assumption A3 with degree p , Lipschitz continuity of F and (15). \square

Remark Hereby, we will obtain at least superlinear convergence of the perturbed method (11) with $p > 0$, $q > 1$ and $r > 0$ and exactly quadratic convergence, if $p = 1$, $q = 2$ and $r = 1$.

5 Application in optimization and numerical results

Consider the general unconstrained optimization problem

$$\min_{x \in R^n} g(x), \quad (16)$$

where $g : R^n \rightarrow R$ is assumed to be a differentiable function and $x = (x_1, \dots, x_n)^T \in R^n$. As it is known, all the local minimizers of the objective function g are stationary points. At these points the gradient $\nabla g(x) = F(x) = (f_1(x), \dots, f_n(x))^T$ vanishes, i.e.

$$\nabla g(x) = F(x) = 0. \quad (17)$$

If the Hessian matrix $H(x)$ of g is symmetric and positive definite then solving the problem (16) is equivalent to solving the problem (17). In the smooth cases, the Newton method given by the iterative formula

$$x^{(k+1)} = x^{(k)} - H(x^{(k)})^{-1} F(x^{(k)})$$

is the successful algorithm for solving problem (16) (see e.g. [17] and [9]). However, if the gradient $\nabla g(x)$ is only Lipschitz continuous function then the Hessian matrix does not exist and the classical Newton method does not work.

Inspired by the promising results of Grapsa et al. [13], we deal the non-smooth version of the perturbed Newton method, which is the particular case of the method described in this paper.

Consider the mapping $W = (w_1, \dots, w_n)^T : D \rightarrow R^n$ in the form

$$w_i(x) = f_i(x) + \sum_{j=1}^n t_j x_j, \quad i = 1, \dots, n,$$

where $T = (t_1, \dots, t_n)^T$ is the vector of the perturbed parameters t_j , $j = 1, \dots, n$, which satisfies the equality with the inner product

$$\langle x, T \rangle = 0. \quad (18)$$

If $(V_k)_j$ denotes the j th column-vector of the matrix V_k taken from the B-differential of F at $x^{(k)}$, i.e. $V_k \in \partial_B F(x^{(k)})$, then the perturbed parameters can be given by

$$t_j = \begin{cases} \frac{f_j(x^{(k)})}{\|(V_k)_j\|_2^2} & \text{for } j = 1, \dots, n-1 \\ -\frac{1}{x_n^{(k)}} \sum_{i=1}^{n-1} t_i x_i^{(k)} & \text{for } j = n \text{ and } x_n^{(k)} \neq 0 \end{cases}. \quad (19)$$

Then, it is easy to notice that the condition (18) holds.

Now, we present the adoption of the Newton-like method for solving nonsmooth equation $W(x) = 0$. The fundamental version of the new iterative scheme is given by

$$x^{(k+1)} = x^{(k)} - (V_k + \tilde{V}_k)^{-1} W(x^{(k)}), \quad (20)$$

where $V_k \in \partial_B F(x^{(k)})$, \tilde{V}_k is the perturbation matrix with $(\tilde{V}_k)_{ij} = t_j$, $j = 1, \dots, n$ for all $i = 1, \dots, n$. Note that $\|\tilde{V}_k\| \rightarrow 0$ as $k \rightarrow \infty$ by the definition of t_j .

Directly by observing that \tilde{V}_k are given based on (19), we obtain that $\|\tilde{V}_k\| = O(\|x - x^*\|)$. Hence, the Lemma 10 implies quadratic convergence of the method (20), if the gradient ∇g satisfies Assumption A3 with degree 1.

Clearly, if we assume that the solution of the Newton equation is not exact, then we may obtain the inexact version of perturbed generalized Newton method with residuals

$$x^{(k+1)} = x^{(k)} - (V_k + \tilde{V}_k)^{-1} W(x^{(k)}) + \tilde{r}_k,$$

which is some particular case of the method (10). Hence, the superlinear convergence is achieved, when

$$\left\| \tilde{V}_k (V_k + \tilde{V}_k)^{-1} F(x^{(k)}) + \left[I - \tilde{V}_k (V_k + \tilde{V}_k)^{-1} \right] \tilde{r}_k \right\| = o(\|F(x^{(k)})\|).$$

Moreover, the perturbation matrices can also be regarded as tool to overcome singularity of the iteration matrices V_k , when the starting point $x^{(0)}$ is too far

from a solution. The BD-regularity assumption guarantees the nonsingularity of V_k only in some neighborhood of x^* (see Proposition 1).

Finally, we consider the constrained optimization problem

$$\begin{cases} \min_{x \in R^n} g(x) \\ g_i(x) \leq 0, \quad i = 1, \dots, m \end{cases}, \quad (21)$$

where $g : R^n \rightarrow R$ and $g_i : R^n \rightarrow R, i = 1, \dots, m$, are twice differentiable. To define the equivalent unconstrained problem (16), we can use the penalty scheme as follows

$$\min_{x \in R^n} g_P(x) := g(x) + P \sum_{i=1}^m (\max\{0, g_i(x)\})^2 + \|\nabla g(x)\|^2, \quad (22)$$

which is some version of the di Pillo-Grippo type Lagrange multiplier function $g_P : R^{n+m} \rightarrow R$ (as e.g. in [22]). Problems (21) and (22) are equivalent under some conditions. The new objective function may not be twice differentiable, but has LC^1 gradient and satisfies A2 at x , where $g_i(x) = 0$.

In order to study the behavior of the nonsmooth perturbed method, we solve several problems. The new algorithm was implemented in C++. All calculations are carried out in double precision and on a computer with Intel Core i7 3.20GHz processor and 8GB RAM using Dev-C++. The termination criterion is always $\|x^{(k+1)} - x^{(k)}\|_2 \leq 10^{-8}$ or $\|F(x^{(k)})\|_2 \leq 10^{-10}$ is not reached after 1000 iterations, where $\|\cdot\|_2$ denotes the Euclidean norm. The forcing terms are chosen as follows: $\eta_k = 0.5$ for all k or $\eta_k = (10k)^{-1}$. Moreover, all tests were conducted with various initial points.

Solving the problem (21) with the following functions (Examples 1–3, [31]), we use (22) with various penalty coefficients P .

Example 1

$$\begin{aligned} g(x) &= (x_1 - 2)^2 + (x_2 - 1)^2, \\ g_1(x) &= x_1^2 - x_2, \\ g_2(x) &= x_2^2 - x_1. \end{aligned}$$

Example 2

$$\begin{aligned} g(x) &= \frac{-(9 - (x_1 - 3)^2)x_2^3}{27\sqrt{3}}, \\ g_1(x) &= -\frac{x_1}{\sqrt{3}} + x_2, \\ g_2(x) &= -x_1 - \sqrt{3}x_2, \\ g_3(x) &= x_1 + \sqrt{3}x_2 - 6. \end{aligned}$$

Table 1 Numerical results for all examples

Example	P	η_k	$x^{(0)}$	N	$x^{(0)}$	N
1	10	<i>const</i>	(0.5, 0.5)	11	(1, 1)	12
1	10	<i>decr</i>	(0.5, 0.5)	9	(1, 1)	9
1	10	<i>const</i>	(10, 10)	×	(−10, −10)	×
1	10	<i>decr</i>	(10, 10)	×	(−10, −10)	×
1	100	<i>const</i>	(0.5, 0.5)	14	(1, 1)	17
1	100	<i>decr</i>	(0.5, 0.5)	12	(1, 1)	13
2	10	<i>const</i>	(2, 0.5)	7	(4, 1)	9
2	10	<i>decr</i>	(2, 0.5)	5	(4, 1)	6
2	10	<i>const</i>	(6, 1.5)	11	(10, 5)	×
2	10	<i>decr</i>	(6, 1.5)	8	(10, 5)	×
2	100	<i>const</i>	(2, 0.5)	6	(4, 1)	9
2	100	<i>decr</i>	(2, 0.5)	5	(4, 1)	7
3	10	<i>const</i>	(0, 0, 0, 0)	19	(0, 0.5, 1.5, −0.5)	23
3	10	<i>decr</i>	(0, 0, 0, 0)	14	(0, 0.5, 1.5, −0.5)	19
3	10	<i>const</i>	(1, 1, 1, 1)	×	(0, 0.8, 1.8, −0.8)	34
3	10	<i>decr</i>	(1, 1, 1, 1)	×	(0, 0.8, 1.8, −0.8)	21
3	100	<i>const</i>	(0, 0, 0, 0)	×	(0, 0.5, 1.5, −0.5)	×
3	100	<i>decr</i>	(0, 0, 0, 0)	17	(0, 0.5, 1.5, −0.5)	×

Example 3

$$g(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5(x_1 + x_2) - 21x_3 + 7x_4,$$

$$g_1(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 - x_3 - x_4 - 8,$$

$$g_2(x) = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 9,$$

$$g_3(x) = 2x_1^2 + x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_2 - x_4 - 5.$$

Table 1 presents results that we obtained solving all examples. Initial points $x^{(0)}$, forcing terms η_k (*const* denotes the constant sequence and *decr* denotes the decreasing sequence), penalty coefficient P and the number of iterations N are shown in this table. × denotes the failure.

6 Conclusions

In this paper, we study the perturbed version of the inexact generalized Newton method for solving nonsmooth equations. Importance of the perturbed inexact Newton approach lies in how easy convergence analysis can be applied to study specific methods. Many of the existing versions of the Newton-like methods can be treated as the particular case of (10), e.g. when the matrices from the B-differential are perturbed or the Newton linear equation is solved insufficiently exactly. Clearly, the results of Cătinaș allowed us to extend the local convergence analysis of perturbed method to the nonsmooth case and we were able to characterize the convergence order for the method in the relatively general form. Additionally, we showed that the inexact generalized Newton and the perturbed inexact generalized Newton methods are closely

related in some natural way, because one can be used to characterize the convergence order of the other.

In the general nonsmooth case, the BD-regularity and semismoothness are sufficient to obtain superlinear convergence of various special versions of the generalized Newton method, while the CD-regularity or strongly semismoothness gives even quadratic convergence. Results of our studies are consistent with these well-known facts, which were confirmed by Theorems 3, 7 and Lemma 10.

Results of the numerical tests show that the new perturbed version of the inexact generalized Newton method can be effectively used to solve not only nonsmooth equations but also non-twice differentiable optimization problems. However, a degenerate behavior can be observed for some combinations of parameters and initial points (Example 3). In turn, use of the decreasing forcing sequence can improve convergence behavior and it is easy to observe that the choice of penalty coefficient has relevant effect on speed of convergence in terms of the number of iterations.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Birgin, E.G., Krejić, N., Martínez, J.M.: Globally convergent inexact quasi-Newton methods for solving nonlinear systems. *Numer. Algorithms* **32**, 249–260 (2003)
2. Bonettini, S.: A nonmonotone inexact Newton method. *Optim. Methods Softw.* **20**, 475–491 (2005)
3. Bonettini, S., Tinti, F.: A nonmonotone semismooth inexact Newton method. *Optim. Methods Softw.* **22**, 637–657 (2007)
4. Cătiņaș, E.: Inexact perturbed Newton methods and application to a class of Krylov solvers. *J. Optim. Theory Appl.* **108**, 543–571 (2001)
5. Cătiņaș, E.: The inexact, inexact perturbed and quasi-Newton methods are equivalent models. *Math. Comput.* **74**, 291–301 (2004)
6. Clarke, F.H.: *Nonsmooth Analysis*. Wiley, New York (1983)
7. Cores, D., Tapia, R.A.: Perturbation lemma for the Newton method with application to the SQP Newton method. *J. Optim. Theory Appl.* **97**, 271–280 (1998)
8. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. *SIAM J. Numer. Anal.* **14**, 400–408 (1982)
9. Dennis, J.E., Schnabel, R.B.: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs (1983)
10. Famularo, D., Sergeev, Y.D., Pugliese, P.: Test problems for Lipschitz univariate global optimization with multiextremal constraints. In: Dzemyda, G., Saltenis, V., Žilinskas, A. (eds.) *Stochastic and Global Optimization*, pp. 93–109. Kluwer, Dordrecht (2002)
11. Floudas, C.A., Pardalos, P.M.: *State of the Art in Global Optimization*. Kluwer, Dordrecht (1996)
12. Gowda, M.S., Ravindran, G.: Algebraic univalence theorems for nonsmooth functions. *J. Math. Anal. Appl.* **252**, 917–935 (2000)
13. Grapsa, T.N., Antonelou, G.E., Kostopoulos, A.E.: Perturbed Newton method for unconstrained optimization. In: *Conference in Numerical Analysis NumAn 2007*, Kalamata, Greece, 2007, pp. 77–80. <http://www.math.upatras.gr/numan2007/NumAn2007.pdf> (2007). Accessed 6 June 2012

14. Martínez, J.M.: Quasi-inexact-Newton methods with global convergence for solving constrained nonlinear systems. *Nonlinear Anal. Theor. Meth. Appl.* **30**, 1–8 (1997)
15. Martínez, J.M., Qi, L.: Inexact Newton method for solving nonsmooth equations. *J. Comput. Appl. Math.* **60**, 127–145 (1995)
16. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. *SIAM J. Control Optim.* **15**, 959–972 (1977)
17. Ortega, J.M., Rheinboldt, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York (1970)
18. Pang, J.S., Qi, L.: A globally convergent Newton method for convex SC^1 minimization problem. *J. Optim. Theory Appl.* **85**, 633–648 (1995)
19. Pardalos, P.M., Rosen, J.B. (eds.): *Computational Methods in Global Optimization*. In: *Annals of Operation Research*, Vol. 25. J.C. Baltzer AG, Basel (1990)
20. Patwardhan, A.A., Karim, M.N., Shah, R.: Controller tuning by a least-squares method. *AIChE J.* **33**, 1735–1737 (1987)
21. di Pillo G.: Exact penalty method. In: Spedicato, E. (eds.) *Algorithms for Continuous Optimization: The State of Art*, pp. 209–253. Kluwer, Dordrecht (1994)
22. di Pillo G., Lucidi S.: On exact augmented Lagrangian functions in nonlinear programming problems. In: di Pillo, G., Gianessi, F. (eds.) *Nonlinear Optimization and Applications*, pp. 85–103. Plenum Press, New York (1996)
23. Pu, D., Tian, W.: Globally convergent inexact generalized Newton's methods for nonsmooth equations. *J. Comput. Appl. Math.* **138**, 37–49 (2002)
24. Qi, L., Chen, X.: A globally convergent successive approximation method for severely nonsmooth equations. *SIAM J. Control Optim.* **33**, 402–418 (1995)
25. Qi, L., Sun, D.: A nonsmooth version of Newton's method. *Math. Program.* **58**, 353–367 (1993)
26. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. *Math. Oper. Res.* **18**, 227–244 (1993)
27. Qi, L.: Superlinearly convergent approximate Newton methods for LC^1 optimization problems. *Math. Program.* **64**, 277–295 (1994)
28. Qi, L.: C-differential operators, C-differentiability and generalized Newton methods. *Applied Mathematics Report AMR 96/5*, University of New South Wales, Australia (1996)
29. Rademacher, H.: Ueber partielle und totale Differenzierbarkeit I. *Math. Ann.* **79**, 340–359 (1919)
30. Ralston, P.A.S., Watson, K.R., Patwardhan, A.A., Deshpande, P.B.: A computer algorithm for optimized control. *Ind. Eng. Chem. Prod. Res. Dev.* **24**, 1132–1136 (1985)
31. Schittkowski K.: *More Test Examples for Nonlinear Programming Codes*. Springer, Berlin (1987)
32. Sergeyev, Y.D., Daponte, P., Grimaldi, D., Molinaro, A.: Two methods for solving optimization problems arising in electronic measurements and electrical engineering. *SIAM J. Optim.* **10**, 1–21 (1999)
33. Stewart, G.W.: *Introduction to Matrix Computations*. Academic Press, London (1973)
34. Śmietański, M.J.: Inexact quasi-Newton global convergent method for solving constrained nonsmooth equations. *Int. J. Comput. Math.* **84**, 1157–1170 (2007)
35. Śmietański, M.J.: Some superlinearly convergent inexact quasi-Newton method for solving nonsmooth equations. *Optim. Methods Softw.* **27**, 405–417 (2012)
36. Ypma, T.J.: The effect of rounding errors on Newton-like methods. *IMA J. Numer. Anal.* **3**, 109–118 (1983)
37. Zhu, D.: Affine scaling inexact generalized Newton algorithm with interior backtracking technique for solving bound-constrained semismooth equations. *J. Comput. Appl. Math.* **187**, 227–252 (2006)