Stability of metric regularity with set-valued perturbations and application to Newton's method for solving generalized equations

Samir Adly, Huynh Van Ngai, Nguyen Van Vu

Stability of metric regularity with set-valued perturbations and application to Newton's method for solving generalized equations

Samir Adly[1] · **Huynh Van Ngai**[2] · **Nguyen Van Vu**[1]

Abstract In this paper, we deal firstly with the question of the stability of the metric regularity under set-valued perturbation. By adopting the measure of closeness between two multifunctions, we establish some stability results on the semi-local/local metric regularity. These results are applied to study the convergence of iterative schemes of Newton-type methods for solving generalized equations in which the set-valued part is approximated. Some examples illustrating the applicability of the proposed method are discussed.

Keywords Generalized equation · Metric regularity · Newton's method · Linear/superlinear convergence

1 Introduction

The metric regularity is a key ingredient in variational analysis and optimization. In recent decades, many papers are devoted to study this property and its applications (see, e.g. [3,

The research of Huynh Van Ngai is supported by NAFOSTED, under Grant: 101.01–2016.27.

✉ Samir Adly
samir.adly@unilim.fr

Huynh Van Ngai
ngaivn@yahoo.com

Nguyen Van Vu
van-vu.nguyen@etu.unilim.fr

[1] Laboratoire XLIM UMR-CNRS 6172, Université de Limoges, 123 Avenue Albert Thomas, 87060 Limoges Cedex, France

[2] Department of Mathematics, University of Quy Nhon, 170 An Duong Vuong, Quy Nhon, Binh Dinh, Vietnam

4, 11, 12, 17–20] and references given therein). Recently, the metric regularity was used to investigate the convergence of the Newton type methods for solving generalized equations [1–3, 7, 8, 16]. Historically, this concept of metric regularity goes back to the openness of linear continuous mapping in the Banach Open Mapping Theorem and to the local openness of smooth nonlinear operators in the Lyusternik-Graves Theorem. For a given set-valued mapping $\Phi : X \rightrightarrows Y$ between two metric spaces X, Y, which are endowed with metrics both denoted by $d(\cdot, \cdot)$, and a given subset $V \in X \times Y$, Φ is said to be *metrically regular* on V with modulus $\tau > 0$ if

$$d\big(x, \Phi^{-1}(y)\big) \leqslant \tau d\big(y, \Phi(x)\big), \quad \text{for all} \quad (x, y) \in V.$$

If for a given $(\bar{x}, \bar{y})$ belonging to the graph of Φ, the inequality above holds for some neighborhood V of $(\bar{x}, \bar{y})$, then we say Φ is (locally) metrically regular at $(\bar{x}, \bar{y})$.

In this paper, we are firstly concerned with the stability of the metric regularity when the set-valued map under consideration is perturbed. When Y is a normed space, it was shown by Ioffe [11], and by Dontchev-Lewis-Rockafellar [6] that, if Φ is metrically regular at $(\bar{x}, \bar{y})$ with modulus τ, then $\Phi + g$ is metrically regular with modulus $\tau/(1 - \tau L)$ at $(\bar{x}, \bar{y} + g(\bar{x}))$ for any locally Lipschitz mapping $g : X \to Y$ at $\bar{x}$ with Lipschitz modulus $L \in (0, \tau^{-1})$. Unfortunately, without additional assumption, this stability property of the local metric regularity does not hold in general when the function $g : X \to Y$ is replaced, in the sum $\Phi + g$, by some set-valued map $\Psi : X \rightrightarrows Y$.

In [12], Ioffe considered the stability of metric regularity with respect to set-valued perturbations, and has established a stability result of the global metric regularity. In that paper, instead of using perturbations of usual addition type, the author has made use of a quantity to measure the closeness between two set-valued mappings. Then some related results have been investigated in [1, 19]. In the present paper, by using the quantity to measuring the closeness between two set-valued mappings, introduced by Ioffe, we firstly establish a stability property of the semi-local metric regularity under set-valued perturbations. For the local metric regularity, a *weaker* stability is presented. Secondly, we deal with generalized equations in Banach space of the form

$$0 \in f(x) + F(x), \tag{1}$$

where $f : X \longrightarrow Y$ is a function between two Banach spaces X and Y, and $F : X \rightrightarrows Y$ is a set-valued mapping. Throughout the paper, we assume that f is at least of class C^1 and that F has a closed graph.

Such a model (1) appears in a very wide range of problems in applied mathematics, engineering and sciences, such as variational inequality, complementarity problem, optimization [7, 16]. Because of its importance, this kind of problem have been studied extensively.

In some suitable situation, one can transform (1) into a nonlinear equation. If $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$, and $F = N_K$ is the normal cone mapping associated to a closed convex set $K \subset \mathbb{R}^n$, (1) can be rewritten into an equation using for example the normal maps introduced by Robinson [21]. Unfortunately, this strategy might not be possible in general. Some Newton-type algorithms (exact and inexact) have been developed for solving inclusions (1) (see, e.g. [1, 3, 5, 8, 13, 23, 24]). Much more surveys about Newton-type method for inclusion can be found in [14, 15]. Up to our knowledge, in almost all existing algorithms in the literature, the set-valued part of inclusion (1) is not considered to be approximated. In the recent paper [9], the authors introduced an iterative scheme for which both f and F are approximated. Then, local analysis of convergence results were proved under certain regularity criterion at a solution as well as the differentiability (in a generalized sense) of F at that point.

However, in that paper, throughout the algorithm, the multifunction F is approximated by a fixed positively homogeneous multifunction H at the solution which is generally unknown.

In this work, we propose an iterative scheme for solving (1), in which the multifunction F is approximated suitably at each iteration. Precisely, by $\mathcal{PH}(X, Y)$, we mean the collection of all positively homogeneous multifunction (cf. [9, 11, 20]) between two Banach spaces X and Y. Given a map $H : X \longrightarrow \mathcal{PH}(X, Y)$, suppose that for some open convex subset $\Omega \subset X$, the set-valued map F is $H(x)$-differentiable at $x \in \Omega$ in the sense of [20] (see Section 2 for the definition). Choosing a starting point x_0 nearby a solution, the proposed algorithm generates a sequence (x_k) by solving the following auxiliary problem

$$0 \in f(x_k) + Df(x_k)(x_{k+1} - x_k) + H(x_k)(x_{k+1} - x_k) + F(x_k). \tag{2}$$

Here, the notation Df indicates the first-order derivative of f. It is clear that if $H(x)(\cdot) = H(\cdot)$ for every x, then (2) recovers the one studied in [9].

The rest of the paper is organized as follows. Section 2 presents the local/semi-local stability results for metric regularity property under set-valued perturbations. These stability results are the key ingredients in the analysis of the convergence properties of the iterative sequences generated by (2). In Section 3, we establish the local and semi-local convergence results for iterative schemes of Newton-type methods of the form (2). Some practical examples, which illustrate methodologically the proposed method, are reported in the last section.

2 Stability of the Metric Regularity

Throughout the paper, all objects are always related to Banach spaces, which are usually denoted by upper character X, Y, etc. Open and closed balls in X with center x and radius r will be written as $\mathbb{B}_X(x, r)$ and $\bar{\mathbb{B}}_X(x, r)$. When dealing with the open (closed) unit ball of X, we use the notation $\mathbb{B}_X$ (resp. $\bar{\mathbb{B}}_X$). We will use the common notation $\|\cdot\|$ for the norm of any Banach space.

Let C and D be two subsets of a space X, we denote their sum $C + D = \{u + v : u \in C, v \in D\}$. If λ is a scalar, then the product λC is the set $\lambda C = \{\lambda u : u \in C\}$. A cone in X is any subset C such that $\lambda C \subset C$ whenever $\lambda \geqslant 0$.

As usual, the distance from a point z to a set K is defined by

$$d(z, K) = \inf_{z' \in K} \|z - z'\|.$$

For $K, K' \subset X$, the quantity $e(K', K) = \sup_{z \in K'} d(z, K)$ is called the excess from K' to K. Finally, we define $d^{\mathcal{H}}(K', K) = \max\{e(K', K), e(K, K')\}$ as the Hausdorff distance between K' and K.

Any set-valued map $T : X \rightrightarrows Y$ is totally determined by its graph $\mathrm{Gr}(T) := \{(x, y) \in X \times Y : y \in T(x)\}$. Corresponding to a set-valued map T, one defines its inverse through the relation

$$(x, y) \in \mathrm{Gr}(T) \Leftrightarrow (y, x) \in \mathrm{Gr}(T^{-1}).$$

We says that T is closed provided that $\mathrm{Gr}(T)$ is a closed set in $X \times Y$. The mapping T is called positively homogeneous, or shortly $T \in \mathcal{PH}(X, Y)$, iff $\mathrm{Gr}(T)$ is a cone. For a mapping $T \in \mathcal{PH}(X, Y)$, its outer norm is the following quantity [9, 20]

$$|T|^+ = \sup_{\|v\| \leqslant 1} \sup_{w \in T(v)} \|w\| = \inf\{\kappa > 0 : T(\mathbb{B}_X) \subset \kappa \mathbb{B}_Y\}, \tag{3}$$

where the conventions $\inf(\emptyset) = +\infty$ and $\sup(\emptyset) = -\infty$ are used. Note that in the case $|T|^+$ is finite, one has $T(0) = \{0\}$ (see e.g. [9]).

Given a mapping $S : X \rightrightarrows Y$ and $T \in \mathcal{PH}(X, Y)$. One says that S is strictly $T-$differentiable at some point $\bar{x} \in X$ if for each $\delta > 0$, there exists a neighborhood V of $\bar{x}$ such that

$$S(x') \subset S(x) + T(x' - x) + \delta \left\| x' - x \right\| \mathbb{B}_Y, \forall x, x' \in V. \tag{4}$$

If x is replaced by $\bar{x}$ in (4) then S is said to be outer T-differentiable at $\bar{x}$. For more details about the differentiation of set-valued mapping, we refer to C.H.J. Pang [20].

The study of convergence analysis developed in this paper needs the notion of regularity [4, 11]. Recall that a mapping $\Phi : X \rightrightarrows Y$ is said to be metrically regular on a set $V \subset X \times Y$ with a modulus $\tau > 0$, iff

$$d\big(x, \Phi^{-1}(y)\big) \leqslant \tau d\big(y, \Phi(x)\big), \quad \text{for all} \quad (x, y) \in V. \tag{5}$$

In this case, we write $\tau \in \mathrm{REG}(\Phi, V)$.

Metric regularity property plays an important role in the analysis of convergence of algorithm (2). The rest of current section is left to prove some stability results which will be important for the study later. These developments are based on a concept of measuring the closeness between two set-valued mappings. Specifically, given $G_1, G_2 : X \rightrightarrows Y$, one defines

$$\sigma_{G_1,G_2}(x, r) := \sup_{\eta_2 \in G_2(x)} \inf_{\eta_1 \in G_1(x)} \sup_{\substack{d(x,u) \leqslant r, \\ \zeta_1 \in G_1(u)}} \inf_{\zeta_2 \in G_2(u)} \|\eta_2 - \eta_1 + \zeta_1 - \zeta_2\|, \tag{6}$$

for each $x \in X$ and $r > 0$. Such a quantity was introduced in [12] (see also [19]).

Let us now present the main results of this section.

Theorem 1 (stability of semi-local metric regularity) *Given two Banach spaces X, Y, and let $\Phi : X \rightrightarrows Y$ and $\Psi : X \rightrightarrows Y$ are two mappings with closed graph. Let r, s, r', s', κ and μ be positively real numbers such that*

$$\kappa\mu < 1, r' + \frac{\kappa}{1 - \kappa\mu} s' < r, s' < s. \tag{7}$$

Suppose

(i) *Φ is metrically regular on the set*

$$V_{r,s}(\Phi, \bar{x}) := \big\{(z, w) \in X \times Y : \|z - \bar{x}\| \leqslant r, d\big(w, \Phi(z)\big) \leqslant s\big\}$$

with a modulus κ,

(ii) *$\sigma_{\Phi,\Psi}(x, \rho) \leqslant \mu\rho$, for all $x \in \bar{\mathbb{B}}_X(\bar{x}, r)$ and $\rho \leqslant \kappa s$.*

If we set $\tau = \frac{1}{1-\kappa\mu}\kappa$, then $\tau \in \mathrm{REG}\big(\Psi, V_{r',s'}(\Psi, \bar{x})\big)$, where

$$V_{r',s'}(\Psi, \bar{x}) := \big\{(z, w) \in X \times Y : \|z - \bar{x}\| \leqslant r', d\big(w, \Psi(z)\big) \leqslant s'\big\}.$$

Proof Pick $(x, y) \in V_{r',s'}(\Psi, \bar{x})$ with $y \notin \Psi(x)$ (the other case is trivial). We need to establish the following inequality

$$d\big(x, \Psi^{-1}(y)\big) \leqslant \tau d\big(y, \Psi(x)\big). \tag{8}$$

Let $\delta > 0$ and $\varepsilon > 0$ such that

$$\kappa(\mu + \delta) < 1, r' + \frac{\kappa}{1 - \kappa(\mu + \delta)}(s' + \varepsilon) < r, s' + \varepsilon < s,$$

we shall claim

$$d\big(x, \Psi^{-1}(y)\big) \leqslant \frac{\kappa}{1-\kappa(\mu+\delta)}(\nu+\varepsilon), \tag{9}$$

where $\nu = d\big(y, \Psi(x)\big)$. To do this, let us first take some $z_0 \in \Psi(x)$ with $\|y - z_0\| < \nu + \varepsilon$. Define $\rho_0 = \kappa(\nu+\varepsilon) \leqslant \kappa(s'+\varepsilon) < \kappa s$, then assumption (1) implies $\sigma_{\Phi,\Psi}(x, \rho_0) \leqslant \mu\rho_0$. Thus, there exists $w_0 \in \Phi(x)$ satisfying

$$\sup_{d(x,u)\leqslant\rho_0, \zeta\in\Phi(u)} \inf_{\zeta'\in\Psi(u)} \big\|z_0 - w_0 + \zeta - \zeta'\big\| < (\mu+\delta)\rho_0. \tag{10}$$

If we set $x_0 = x$ and $y_0 = y + w_0 - z_0$, then from $\|x_0 - \bar{x}\| \leqslant r' < r$ and

$$d\big(y_0, \Phi(x_0)\big) = d\big(y_0, \Phi(x)\big) \leqslant \|y_0 - w_0\| = \|y - z_0\| < \nu + \varepsilon < s,$$

the pair (x_0, y_0) belongs to $V_{r,s}(\Phi, \bar{x})$. By applying condition (1) we get

$$d\big(x_0, \Phi^{-1}(y_0)\big) \leqslant \kappa d\big(y_0, \Phi(x_0)\big) < \kappa(\nu+\varepsilon) = \rho_0,$$

which ensures $\mathbb{B}_X(x_0, \rho_0) \cap \Phi^{-1}(y_0) \neq \emptyset$. Taking $x_1 \in \mathbb{B}_X(x_0, \rho_0) \cap \Phi^{-1}(y_0)$, (10) give us

$$\sup_{\zeta\in\Phi(x_1)} \inf_{\zeta'\in\Psi(x_1)} \big\|z_0 - w_0 + \zeta - \zeta'\big\| < (\mu+\delta)\rho_0.$$

This allows us to select a point $z_1 \in \Psi(x_1)$ possessing property below

$$\|z_0 - w_0 + y_0 - z_1\| < (\mu+\delta)\rho_0.$$

Due to the choice of ε and δ, we have

$$\|\bar{x} - x_1\| \leqslant \|\bar{x} - x_0\| + \|x_0 - x_1\| < r' + \kappa(\nu+\varepsilon) < r' + \frac{\kappa}{1-\kappa(\mu+\delta)}(s'+\varepsilon),$$

so x_1 is inside $\mathbb{B}_X(\bar{x}, r)$. By virtue of hypothesis (1), we obtain

$$\inf_{w\in\Phi(x_1)} \sup_{\substack{d(x_1,u)\leqslant\rho_1,\\ \eta\in\Phi(u)}} \inf_{\eta'\in\Psi(u)} \big\|z_1 - w + \eta - \eta'\big\| \leqslant \sigma_{\Phi,\Psi}(x_1, \rho_1) \leqslant \mu\rho_1,$$

where $\rho_1 := \kappa(\mu+\delta)\rho_0 < \rho_0$. Therefore, $\Phi(x_1)$ contains some element w_1 for which the following assertion is fulfilled

$$\sup_{d(x_1,u)\leqslant\rho_1, \eta\in\Phi(u)} \inf_{\eta'\in\Psi(u)} \big\|z_1 - w_1 + \eta - \eta'\big\| < (\mu+\delta)\rho_1. \tag{11}$$

Let us define $y_1 = y + w_1 - z_1$, then it holds that

$$\begin{aligned} d\big(y_1, \Phi(x_1)\big) &\leqslant \|y_1 - w_1\| = \|y - z_1\| = \|z_0 - w_0 + y_0 - z_1\| \\ &< (\mu+\delta)\rho_0 = [\kappa(\mu+\delta)](\nu+\varepsilon) < s. \end{aligned}$$

Thus, the inclusion $(x_1, y_1) \in V_{r,s}(\Phi, \bar{x})$ is now clear. By making use of condition (1) in the statement of Theorem 1, we arrive

$$d\big(x_1, \Phi^{-1}(y_1)\big) \leqslant \kappa d\big(y_1, \Phi(x_1)\big) < \kappa(\mu+\delta)\rho_0 = \rho_1.$$

The latter shows that there is some $x_2 \in \Phi^{-1}(y_1)$ with $\|x_1 - x_2\| < \rho_1$. Under assignment $u = x_2$, (11) ensures the existence of a point, written as z_2, belonging to $\Psi(x_2)$ and

$$\|z_1 - w_1 + y_1 - z_2\| < (\mu+\delta)\rho_1.$$

Passing to the inductive step, suppose the iterations $x_1, \ldots, x_n$ in X are known for some $n \geqslant 2$. As suggested by the arguments above, let us involve $w_0 \in \Phi(x_0), \ldots, w_{n-1} \in \Phi(x_{n-1})$ and $z_0 \in \Psi(x_0), \ldots, z_n \in \Psi(x_n)$ for which the following conditions are satisfied

- $x_j \in \Phi^{-1}(y_{j-1})$, where $y_{j-1} = y + w_{j-1} - z_{j-1}$, $j = 1, \ldots, n$;
- $\|x_j - x_{j+1}\| < \rho_j$, $\|z_j - w_j + y_j - z_{j+1}\| < (\mu + \delta)\rho_j$, $j = 0, \ldots, n-1$;
- $\rho_j = [\kappa(\mu + \delta)]^j \rho_0$, $j = 0, \ldots, n-1$.

Due to the triangle inequality, one has

$$\begin{aligned}\|x_n - \bar{x}\| &\leqslant \sum_{j=0}^{n-1} \|x_{j+1} - x_j\| + \|x_0 - \bar{x}\| < \sum_{j=0}^{n-1} [\kappa(\mu + \delta)]^j \rho_0 + r' \\ &< \frac{1}{1 - \kappa(\mu + \delta)} \kappa(\nu + \varepsilon) + r' < r.\end{aligned}$$

Define a new parameter $\rho_n := [\kappa(\mu+\delta)]^n \rho_0$. By invoking the hypothesis (1) once more, we can write

$$\inf_{\xi \in \Phi(x_n)} \sup_{\substack{d(x_n, u) \leqslant \rho_n, \\ \upsilon \in \Phi(u)}} \inf_{\upsilon' \in \Psi(u)} \|z_n - \xi + \upsilon - \upsilon'\| \leqslant \sigma_{\Phi,\Psi}(x_n, \rho_n) \leqslant \mu\rho_n.$$

Let us take $w_n \in \Phi(x_n)$ such that

$$\sup_{d(x_n, u) \leqslant \rho_n, \upsilon \in \Phi(u)} \inf_{\upsilon' \in \Psi(u)} \|z_n - w_n + \upsilon - \upsilon'\| < (\mu + \delta)\rho_n. \tag{12}$$

In order to proceed to the construction, we put $y_n = y + w_n - z_n$. Then, it is possible to estimate the distance $d\big(y_n, \Phi(x_n)\big)$ as follows

$$\begin{aligned}d\big(y_n, \Phi(x_n)\big) &\leqslant \|y_n - w_n\| = \|y - z_n\| = \|z_{n-1} - w_{n-1} + y_{n-1} - z_n\| \\ &< (\mu + \delta)\rho_{n-1} = (\mu + \delta)[\kappa(\mu + \delta)]^{n-1}\rho_0 \\ &= [\kappa(\mu + \delta)]^n(\nu + \varepsilon) < s.\end{aligned}$$

Combining with $\|x_n - \bar{x}\| < r$, we conclude $(x_n, y_n) \in V_{r,s}(\Phi, \bar{x})$. As a result,

$$d\big(x_n, \Phi^{-1}(y_n)\big) \leqslant \kappa d\big(y_n, \Phi(x_n)\big) < \kappa(\mu + \delta)\rho_{n-1} = [\kappa(\mu + \delta)]^n \rho_0 = \rho_n.$$

Hence, the set $\Phi^{-1}(y_n)$ must contain at least one element, saying x_{n+1}, which satisfies $\|x_n - x_{n+1}\| < \rho_n$. After substituting $u = x_{n+1}$, (12) tells us

$$\inf_{\upsilon \in \Psi(x_{n+1})} \|z_n - w_n + y_n - \upsilon\| < (\mu + \delta)\rho_n.$$

From this, we can find $z_{n+1} \in \Psi(x_{n+1})$ such that

$$\|z_n - w_n + y_n - z_{n+1}\| < (\mu + \delta)\rho_n.$$

Repeating the current process, all sequences (x_n), (y_n), (z_n) and (w_n) are well-defined by induction.

Recall that $\kappa(\mu+\delta) < 1$, the series $\sum_{k \geqslant 0} \rho_k = \sum_{k \geqslant 0} [\kappa(\mu+\delta)]^k \rho_0$ converges. Because of $\|x_n - x_{n+1}\| < \rho_n$, (x_n) is Cauchy sequence. Consequently, x_k converges in X to a limit x^*. In order to attain the necessary conclusion, we just verify $y \in \Psi(x^*)$ and $\|x - x^*\| \leqslant \frac{\kappa}{1-\kappa(\mu+\delta)}(\nu + \varepsilon)$. Indeed, according to the construction $z_n \in \Psi(x_n)$. But since

$$\|y - z_n\| = \|z_{n-1} - w_{n-1} + y_{n-1} - z_n\| < (\mu + \delta)\rho_{n-1},$$

z_n converges to y as $n \to \infty$. By passing to the limit in the inclusion $z_n \in \Psi(x_n)$, and note that Ψ has closed graph, we obtain $y \in \Psi(x^*)$. For the error bound, one has

$$\begin{aligned}\left\|x - x^*\right\| = \left\|x_0 - x^*\right\| &= \left\| \sum_{k\geqslant 0} (x_k - x_{k+1}) \right\| \leqslant \sum_{k\geqslant 0} \|x_k - x_{k+1}\| \leqslant \sum_{k\geqslant 0} \rho_k \\ &= \sum_{k\geqslant 0} [\kappa(\mu+\delta)]^k \rho_0 = \frac{1}{1-\kappa(\mu+\delta)} \kappa(\nu+\varepsilon).\end{aligned}$$

Taking into account $y \in \Psi(x^*)$, (9) is valid as well. Finally, let $\varepsilon \to 0$ and $\delta \to 0$ in (9) we obtain (8). The proof is done. □

When studying the local convergence of the scheme in (2), we may need a stability result related to the local metric regularity property for set-valued map. The following statement is in this sense.

Theorem 2 *Let $\Phi, \Psi : X \rightrightarrows Y$ be two closed set-valued maps, and $(\bar{x}, \bar{y}) \in \mathrm{Gr}\Phi$. Consider some positive parameters κ, r and s such that $\kappa \in \mathrm{REG}\big(\Phi, V\big)$ for $V = \mathbb{B}_X(\bar{x}, r) \times \mathbb{B}_Y(\bar{y}, s)$. Let $\mu > 0$, $\delta > 0$ and $\nu > 0$ satisfy*

$$\kappa\mu < 1, \frac{\kappa}{1-\kappa\mu}\nu < r, \delta + (1+\kappa\mu)\,\nu < s.$$

In addition, assume that there is $\bar{z} \in \Psi(\bar{x})$ with

$$\inf_{v\in\Phi(x)} \sup_{w\in\Psi(x)} \|(w-\bar{z}) - (v-\bar{y})\| \leqslant \delta, \quad \textit{for} \quad x \in \mathbb{B}_X(\bar{x}, r). \tag{13}$$

If

$$\sigma_{\Phi,\Psi}(x, \varepsilon) \leqslant \mu\varepsilon, \quad \textit{when} \quad x \in \mathbb{B}_X(\bar{x}, r) \quad \textit{and} \quad \varepsilon < r, \tag{14}$$

then one has

$$d\big(\bar{x}, \Psi^{-1}(z)\big) \leqslant \tau d\big(z, \Psi(\bar{x})\big), \quad \textit{whenever} \quad z \in \mathbb{B}_Y(\bar{z}, \nu), \tag{15}$$

where $\tau = \frac{1}{1-\kappa\mu}\kappa$.

Proof Let's pick $\kappa' > \kappa$ and $\mu' > \mu$ which fulfill

$$\kappa'\mu' < 1, \frac{\kappa}{1-\kappa'\mu'}\nu < r, \delta + \big(1+\kappa'\mu'\big)\,\nu < s.$$

Fix $z \in \mathbb{B}_Y(\bar{z}, \nu)$ and let $x_0 = \bar{x}$. The case $z \in \Psi(\bar{x})$ is trivial. Otherwise, denoting $C = d\big(z, \Psi(\bar{x})\big) > 0$. Observe that $C \leqslant \|z - \bar{z}\| < \nu$, so for some $\alpha > 0$ small enough, we can select $w_0 \in \Psi(\bar{x})$ with $\|z - w_0\| < C + \alpha < \nu$. Put $r_0 = \kappa(C+\alpha)$, the inequality $r_0 < \kappa\nu < r$ is evident. Thus, assumption (14) implies

$$\sup_{\xi\in\Psi(x_0)} \inf_{\xi'\in\Phi(x_0)} \sup_{\substack{\|x-x_0\|\leqslant r_0,\\ \zeta\in\Phi(x)}} \inf_{\zeta'\in\Psi(x)} \left\|\xi - \xi' + \zeta - \zeta'\right\| = \sigma_{\Phi,\Psi}(x_0, r_0) \leqslant \mu r_0.$$

Since $w_0 \in \Psi(x_0)$, we have

$$\inf_{\xi'\in\Phi(x_0)} \sup_{\substack{\|x-x_0\|\leqslant r_0,\\ \zeta\in\Phi(x)}} \inf_{\zeta'\in\Psi(x)} \left\|w_0 - \xi' + \zeta - \zeta'\right\| \leqslant \mu r_0 < \mu' r_0,$$

which ensures the existence of $v_0 \in \Phi(x_0)$ such that

$$\sup_{\|x-x_0\|\leqslant r_0, \zeta\in\Phi(x)} \inf_{\zeta'\in\Psi(x)} \left\|w_0 - v_0 + \zeta - \zeta'\right\| < \mu' r_0. \tag{16}$$

By setting $y_0 = z - w_0 + v_0$, we claim $\|y_0 - \bar{y}\| < s$. Indeed, the triangle inequality in Y tells us

$$\begin{aligned}\|y_0 - \bar{y}\| &\leqslant \|z - \bar{z}\| + \|w_0 - v_0 + \bar{y} - \bar{z}\| \\ &\leqslant \|z - \bar{z}\| + \left\|w_0 - v_0 + \zeta - \zeta'\right\| + \left\|(\zeta' - \bar{z}) - (\zeta - \bar{y})\right\|,\end{aligned}$$

where ζ varies in $\Phi(x_0)$ and ζ' in $\Psi(x_0)$. Thus,

$$\begin{aligned}\|y_0 - \bar{y}\| &\leqslant \|z - \bar{z}\| + \sup_{\zeta \in \Phi(x_0)} \inf_{\zeta' \in \Psi(x_0)} \left\|w_0 - v_0 + \zeta - \zeta'\right\| \\ &\quad + \inf_{\zeta \in \Phi(x_0)} \sup_{\zeta' \in \Psi(x_0)} \left\|(\zeta' - \bar{z}) - (\zeta - \bar{y})\right\| \\ &< \nu + \mu' r_0 + \delta < (1 + \kappa\mu')\nu + \delta < s.\end{aligned}$$

Using metric regularity property of Φ, we find

$$d\big(x_0, \Phi^{-1}(y_0)\big) \leqslant \kappa d\big(y_0, \Phi(x_0)\big) \leqslant \kappa \|y_0 - v_0\| = \kappa \|z - w_0\| < \kappa(C + \alpha).$$

Therefore, there exists $x_1 \in \Phi^{-1}(y_0)$ such that $\|x_0 - x_1\| < \kappa(C + \alpha) = r_0$.
Next, invoking (16)

$$d\big(w_0 - v_0 + y_0, \Psi(x_1)\big) \leqslant \sup_{\|x - x_0\| \leqslant r_0, \zeta \in \Phi(x)} \inf_{\zeta' \in \Psi(x)} \left\|w_0 - v_0 + \zeta - \zeta'\right\| < \mu' r_0,$$

which permit us to determine an element w_1 in $\Psi(x_1)$ obeying the inequality $\|w_0 - v_0 + y_0 - w_1\| < \mu' r_0$. Let $r_1 = \kappa'\mu' r_0$. After substituting $x = x_1$ and $\varepsilon = r_1$ into the left-hand side of (14), we deduce

$$\inf_{v \in \Phi(x_1)} \sup_{\substack{\|x - x_1\| \leqslant r_1 \\ \xi \in \Phi(x)}} \inf_{\xi' \in \Psi(x)} \left\|w_1 - v + \xi - \xi'\right\| \leqslant \sigma_{\Phi,\Psi}(x_1, r_1) \leqslant \mu r_1.$$

This shows that, there is v_1 in the set $\Phi(x_1)$ such that

$$\sup_{\|x - x_1\| \leqslant r_1, \xi \in \Phi(x)} \inf_{\xi' \in \Psi(x)} \left\|w_1 - v_1 + \xi - \xi'\right\| < \mu' r_1.$$

Setting now $y_1 = z - w_1 + v_1$, we require (x_1, y_1) lying into the set V. In fact, the choice of x_1 above have affirmed $x_1 \in \mathbb{B}_X(\bar{x}, r)$. Moreover, as similar as the case of y_0, the following estimation holds

$$\begin{aligned}\|y_1 - \bar{y}\| &\leqslant \|z - \bar{z}\| + \sup_{\xi \in \Phi(x_1)} \inf_{\xi' \in \Psi(x_1)} \left\|w_1 - v_1 + \xi - \xi'\right\| \\ &\quad + \inf_{\xi \in \Phi(x_1)} \sup_{\xi' \in \Psi(x_1)} \left\|(\xi' - \bar{z}) - (\xi - \bar{y})\right\| \\ &< \nu + \mu' r_1 + \delta < \nu + \mu'(\kappa'\mu')\kappa\nu + \delta < \left(1 + \kappa'\mu'\right)\nu + \delta < s.\end{aligned}$$

This implies the inclusion $(x_1, y_1) \in V$. As a result,

$$d\big(x_1, \Phi^{-1}(y_1)\big) \leqslant \kappa d\big(y_1, \Phi(x_1)\big) \leqslant \kappa \|y_1 - v_1\| = \kappa \|z - w_1\|.$$

Recalling $z = y_0 + w_0 - v_0$, which gives us $\|z - w_1\| = \|w_0 - v_0 + y_0 - w_1\| < \mu' r_0$. Hence,

$$d\big(x_1, \Phi^{-1}(y_1)\big) < \kappa\mu' r_0 < (\kappa'\mu') r_0.$$

Consequently, the set $\Phi^{-1}(y_1)$ contains a point x_2 such that $\|x_1 - x_2\| < (\kappa'\mu') r_0$.

The construction continues with the inductive process. Suppose that the iterations $x_0 = \bar{x}, x_1, \ldots, x_n$ are given for $n \geqslant 2$. Alternatively, we include some points $v_0 \in$

$\Phi(x_0), \ldots, v_{n-1} \in \Phi(x_{n-1})$ along with $w_0 \in \Psi(x_0), \ldots, w_{n-1} \in \Psi(x_{n-1})$ satisfying simultaneously the conditions below:

- $x_{k+1} \in \Phi^{-1}(y_k)$ for $y_k = z - w_k + v_k$ and $k \leqslant n$;
- $\|x_k - x_{k+1}\| < r_k$, with $r_k = \left(\kappa'\mu'\right)^k r_0$;
- $\|w_{k-1} - v_{k-1} + y_{k-1} - w_k\| < \mu' r_{k-1}$ holds for $k = 1, \ldots, n-1$;
- $\sup_{\|x-x_k\| \leqslant r_k, v \in \Phi(x)} \inf_{v' \in \Psi(x)} \left\|w_k - v_k + v - v'\right\| < \mu' r_k$.

With the goal of generating x_{n+1}, we exploit the hypothesis $\kappa \in \mathrm{REG}\big(\Phi, V\big)$. To do this, let's first notice that x_n is not out of the ball $\mathbb{B}_X(\bar{x}, r)$, which is a consequence of the following estimations

$$\|x_n - \bar{x}\| \leqslant \sum_{k=0}^{n-1} \|x_k - x_{k+1}\| < \sum_{k=0}^{n-1} \left(\kappa'\mu'\right)^k r_0 < \frac{1}{1-\kappa'\mu'} \kappa v < r.$$

Recall that $\|x_{n-1} - x_n\| < r_{n-1}$ and $y_{n-1} \in \Phi(x_n)$, we deduce

$$d\big(w_{n-1} - v_{n-1} + y_{n-1}, \Psi(x_n)\big) \leqslant \sup_{\substack{\|x-x_{n-1}\| \leqslant r_{n-1},\\ v \in \Phi(x)}} \inf_{v' \in \Psi(x)} \left\|w_{n-1} - v_{n-1} + v - v'\right\|.$$

In other words, $d\big(w_{n-1} - v_{n-1} + y_{n-1}, \Psi(x_n)\big) < \mu' r_{n-1}$. We take a point w_n in $\Psi(x_n)$ fulfilling inequality $\|w_{n-1} - v_{n-1} + y_{n-1} - w_n\| < \mu' r_{n-1}$. Put $r_n := \left(\kappa'\mu'\right)^n r_0$. By involving the hypothesis (14) once again, we derive

$$\inf_{v \in \Phi(x_n)} \sup_{\substack{\|x-x_n\| \leqslant r_n,\\ \zeta \in \Phi(x)}} \inf_{\zeta' \in \Psi(x)} \left\|w_n - v + \zeta - \zeta'\right\| \leqslant \sigma_{\Phi,\Psi}(x_n, r_n) \leqslant \mu r_n.$$

Thanks to this, it is possible to find an element $v_n \in \Phi(x_n)$ such that

$$\sup_{\|x-x_n\| \leqslant r_n, \zeta \in \Phi(x)} \inf_{\zeta' \in \Psi(x)} \left\|w_n - v_n + \zeta - \zeta'\right\| < \mu' r_n.$$

Define now $y_n = z - w_n + v_n$. After repeating analogous arguments as $n = 0$, one gets

$$\begin{aligned} \|y_n - \bar{y}\| &\leqslant \|z - \bar{z}\| + \sup_{\omega \in \Phi(x_n)} \inf_{\omega' \in \Psi(x_n)} \left\|w_n - v_n + \omega - \omega'\right\| \\ &\quad + \inf_{\omega \in \Phi(x_n)} \sup_{\omega' \in \Psi(x_n)} \left\|(\omega' - \bar{z}) - (\omega - \bar{y})\right\| \\ &< \nu + \mu' r_n + \delta < \nu + \mu'(\kappa'\mu')^n \kappa \nu + \delta < \left(1 + \kappa'\mu'\right)\nu + \delta < s. \end{aligned}$$

In summary, (x_n, y_n) belongs to V. Including the fact $\kappa \in \mathrm{REG}\big(\Phi, V\big)$, we find

$$d\big(x_n, \Phi^{-1}(y_n)\big) \leqslant \kappa d\big(y_n, \Phi(x_n)\big) \leqslant \kappa \|y_n - v_n\| = \kappa \|z - w_n\|.$$

Observing $y_{n-1} = z - w_{n-1} + v_{n-1}$, the latter implies

$$d\big(x_n, \Phi^{-1}(y_n)\big) \leqslant \kappa \|z - w_n\| = \kappa \|w_{n-1} - v_{n-1} + y_{n-1} - w_n\| < \kappa'\mu' r_{n-1}.$$

Hence, the set $\Phi^{-1}(y_n)$ contain some element x_{n+1} so that $\|x_n - x_{n+1}\| < r_n$ is fulfilled. Induction step is done.

According to the construction, it holds that

$$\begin{aligned} \|x_n - x_{n+k}\| &\leqslant \sum_{j=0}^{k-1} \left\|x_{n+j} - x_{n+j+1}\right\| < \sum_{j=0}^{k-1} \left(\kappa'\mu'\right)^{n+j} r_0 \\ &< \left(\kappa'\mu'\right)^n \frac{\kappa}{1-\kappa'\mu'}(C + \alpha). \end{aligned} \tag{17}$$

Since $\kappa'\mu' < 1$, (17) yields $\limsup_{m,n\to\infty} \|x_m - x_n\| = 0$. Thus, the limit $x^* = \lim_{n\to\infty} x_n$ exists. By fixing $n = 0$ and letting $k \to \infty$, we derive $\|\bar{x} - x^*\| \leqslant \frac{\kappa}{1-\kappa'\mu'}(C + \alpha)$. On the other hand, recalling $\|z - w_n\| = \|w_{n-1} - v_{n-1} + y_{n-1} - w_n\|$, and taking into account $\|w_{n-1} - v_{n-1} + y_{n-1} - w_n\| < \mu' r_{n-1}$, we conclude $w_n \to z$ as $n \to \infty$. But w_n is in $\Psi(x_n)$, so $z \in \Psi(x^*)$ due to the closedness of $\mathrm{Gr}(\Psi)$. Consequently,

$$d\big(\bar{x}, \Psi^{-1}(z)\big) \leqslant \big\|\bar{x} - x^*\big\| \leqslant \frac{\kappa}{1-\kappa'\mu'}(C + \alpha).$$

Since α is arbitrarily small, and both parameters κ', μ' are chosen independently of z, we obtain $d\big(\bar{x}, \Psi^{-1}(z)\big) \leqslant \frac{\kappa}{1-\kappa\mu}C$. This completes the proof. □

3 Convergence of the Newton-Type Algorithm Using Set-Valued Differentiation

We start with the following definition, that will be useful later.

Definition 1 Let $H : X \longrightarrow \mathcal{PH}(X, Y)$ be a given map, and let $\Omega \subset X$ be a nonempty open subset.

(i) A mapping $\Phi : X \rightrightarrows Y$ is said to be pointwise strictly differentiable with respect to H on Ω if and only if for any $x \in \Omega$ and $\varepsilon > 0$, there exists $\delta = \delta(x, \varepsilon) > 0$ such that

$$\Phi(z') \subset \Phi(z) + H(x)(z' - z) + \varepsilon \big\|z' - z\big\| \mathbb{B}_Y;\ \forall z, z' \in \mathbb{B}_X(x, \delta) \cap \Omega. \tag{18}$$

(ii) $\Phi : X \rightrightarrows Y$ is differentiable with respect to H uniformly on Ω providing for all $\varepsilon > 0$ there is $\delta > 0$ satisfying

$$\Phi(x') \subset \Phi(x) + H(x)(x' - x) + \varepsilon \big\|x' - x\big\| \mathbb{B}_Y, \tag{19}$$

whenever $x, x' \in \Omega$ and $\big\|x' - x\big\| \leqslant \delta$.

On the other hand, with the aim of studying the convergence of the scheme (2), we assign to each map $H : X \longrightarrow \mathcal{PH}(X, Y)$ a function defined as follows

$$\Lambda_H(x, x', t) := \sup_{\|u\| \leqslant t} e\big(H(x)(u) + H(x')(-u), 0\big);\ x, x' \in X, t \geqslant 0. \tag{20}$$

Remark 1 Suppose that both $|H(x)|^+$ and $\big|H(x')\big|^+$ are finite. From the definition of outer norm one has $H(x)(u) \subset |H(x)|^+ \|u\| \bar{\mathbb{B}}_Y$ as well as $H(x')(-u) \subset \big|H(x')\big|^+ \|u\| \bar{\mathbb{B}}_Y$. Consequently,

$$e\big(H(x)(u) + H(x')(-u), 0\big) \leqslant \big(|H(x)|^+ + \big|H(x')\big|^+\big) \|u\|.$$

That is, when $|H(\cdot)|^+$ is bounded, all real-valued functions $\Lambda_H(x, x', \cdot)$ can be majorized by one linear map.

Returning back to the main results, a semi-local convergence analysis for (2) is presented in the next theorem.

Theorem 3 (semi-local analysis) *Let Ω be an open convex subset of X on which Df is Lipschitz continuous with a modulus $L > 0$. Given a map $H : X \longrightarrow \mathcal{PH}(X, Y)$ so that F is differentiable with respect to H uniformly on Ω. Let's fix some $x \in \Omega$ and $r > 0$ with $\Omega_x = \bar{\mathbb{B}}_X(x, r) \subset \Omega$. Suppose that ρ_0 and ε_0 are some positive numbers satisfying*

$$F(z') \subset F(z) + H(z)(z' - z) + \rho_0 \|z' - z\| \mathbb{B}_Y \tag{21}$$

when $z', z \in \Omega_x$ and $\|z' - z\| \leqslant \varepsilon_0$. Additionally, we also assume the following hypotheses:

(i) *it holds that $\tau \in \mathrm{REG}\big(\Psi_x, V_{r,s}(\Psi_x)\big)$, where $\Psi_x(\cdot) := Df(x)(\cdot) + H(x)(\cdot)$ and $V_{r,s}(\Psi_x) = \big\{(v, w) : \|v\| \leqslant r, d\big(w, \Psi_x(v)\big) \leqslant s\big\}$;*

(ii) *$d\big(0,\, f(x) + F(x)\big) < \min\big\{s, \tau^{-1}\varepsilon_0\big\}$;*

(iii) *there exists some constant $\lambda > 0$ such that*

$$\sigma_{H(x),H(z)}(u, \delta) \leqslant \lambda \|x - z\| \delta, \tag{22}$$

for $z \in \bar{\mathbb{B}}_X(x, r)$, $\|u\| \leqslant r$ and $\delta \leqslant \tau s$;

(iv) *$|H(x)|^+ < \infty$ and $\Lambda_H(z, z', t) \leqslant \omega(t)$ for all $z, z' \in \Omega_x$ and $t \leqslant \varepsilon_0$, where $\omega : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a nondecreasing convex function with $\omega(0^+) = \lim\limits_{t \searrow 0} \omega(t) = 0$;*

(v) *$\tau(L + \lambda)r < 1$ and $K_0 := \frac{1}{2}L\varepsilon_0 + \rho_0 + \frac{\omega(\varepsilon_0)}{\varepsilon_0} < \frac{1}{\tau} - (L + \lambda)r$.*

If either $\frac{1-\tau(L+\lambda)r}{1-\tau K_0-\tau(L+\lambda)r}\varepsilon_0 \leqslant r$ or $\frac{1-\tau(L+\lambda)r}{1-\tau K_0-\tau(L+\lambda)r}\tau s \leqslant r$ is valid, then there exists a solution x^* of problem (1) *such that*

$$\big\|x - x^*\big\| \leqslant \frac{1 - \tau(L + \lambda)r}{1 - \tau K_0 - \tau(L + \lambda)r} \min\big\{\tau s, \varepsilon_0\big\} \leqslant r.$$

Furthermore, the scheme (2) *produces a sequence (x_k) which starts at $x_0 = x$ and converges to x^* at least R-linearly, i.e.,*

$$\limsup_{k \to \infty} \big\|x_k - x^*\big\|^{1/k} < 1.$$

For the proof of this theorem, the next technical lemmas will be useful.

Lemma 1 *Keep in mind the assumptions of Theorem 3. For each z in $\mathbb{B}_X(x, r)$ set $\Psi_z(\cdot) := Df(z)(\cdot) + H(z)(\cdot)$. Then one has*

$$\sigma_{\Psi_x,\Psi_z}(u, \delta) \leqslant (L + \lambda)\, \|x - z\|\, \delta, \tag{23}$$

where $u \in r\bar{\mathbb{B}}_X$ and $\delta \leqslant \tau s$.

Proof Fix $u \in r\bar{\mathbb{B}}_X$ and $\delta \leqslant \tau s$. Due to the definition of $\sigma_{\Psi_x,\Psi_z}(u, \delta)$, and by using triangle inequality, we deduce

$$\begin{aligned}
\sigma_{\Psi_x,\Psi_z}(u, \delta) \leqslant & \sup_{\|u-v\| \leqslant \delta} \|[Df(z) - Df(x)](u - v)\| \\
& + \sup_{\zeta \in H(z)(u)} \inf_{\zeta' \in H(x)(u)} \sup_{\substack{\|u-v\| \leqslant \delta, \\ \xi \in H(x)(v)}} \inf_{\xi' \in H(z)(v)} \big\|\zeta - \zeta' + \xi - \xi'\big\| \\
\leqslant & \sup_{\|u-v\| \leqslant \delta} \|Df(z) - Df(x)\|\, \|u - v\| + \sigma_{H(x),H(z)}(u, \delta) \\
\leqslant & L\|z - x\|\delta + \lambda\|x - z\|\delta.
\end{aligned}$$

This implies the conclusion of Lemma 1. □

Lemma 2 *Let $\gamma_1 = \frac{1}{2}\tau L$, $\gamma_2 = \rho_0\tau$ and $\gamma_3 = \tau(L+\lambda)r < 1$, where L, τ, ρ_0, λ and r are in statement of Theorem 3. Let ω be the function appeared in the same theorem. Considering $h(t) = \frac{1}{1-\gamma_3}\big(\gamma_1 t^2 + \gamma_2 t + \tau\omega(t)\big)$ for $t \geqslant 0$. Then under initial constraint $h(\alpha_0) \leqslant \alpha_0$, the recurrence $\alpha_{k+1} = h(\alpha_k)$ produces a sequence converging at least linearly.*

Proof The case $h(\alpha_0) = \alpha_0$ is trivial, because (α_k) is a stationary sequence. Suppose $h(\alpha_0) < \alpha_0$. It is clear that h is convex and increasing, with $h(0) = 0$. The convexity of h implies that $t \in (0, \alpha_0] \longmapsto \frac{h(t)}{t}$ is an increasing function (see [10]). Thus

$$0 < t \leqslant \alpha_0 \Longrightarrow h(t) \leqslant \frac{h(\alpha_0)}{\alpha_0} t = qt,$$

where $q = \frac{h(\alpha_0)}{\alpha_0} \in [0, 1)$. By induction we can prove that (α_k) is well-defined and satisfies $\alpha_k \leqslant q\alpha_{k-1}$, so the linear convergence is followed. □

We are now in a position to prove Theorem 3.

Proof of Theorem 3 We will subdivide the proof in several steps.

Step1: Construction of a Sequence At the beginning, we denote $x_0 = x$, $\Psi_0 = \Psi_x$, $\tau_0 = \tau$, $r_0 = r$ and $s_0 = s$. Then Ψ_0 is metrically regular on the set $V_0 = V_{r,s}(\Psi_x)$ with modulus τ_0. From assumption (3), we can select $y_0 \in F(x_0)$ for which the inequality $\|-f(x_0) - y_0\| < \min\{s, \tau^{-1}\varepsilon_0\}$ holds. Let $z_0 = -f(x_0) - y_0$, we derive

$$d\big(z_0, \Psi_0(0)\big) = d\big(z_0, H(x)(0)\big) = \|z_0\| < s.$$

This means $(0, z_0) \in V_0$, which guarantees the validity of

$$d\big(0, \Psi_0^{-1}(z_0)\big) \leqslant \tau_0 d\big(z_0, \Psi_0(0)\big) = \tau_0 \|z_0\| < \min\{\tau s, \varepsilon_0\}.$$

Hence, there exists an element, says v_0, in the set $\Psi_0^{-1}(z_0)$ such that

$$\|v_0\| < \min\{\tau s, \varepsilon_0\}.$$

Define $x_1 = x_0 + v_0$, we get

$$-f(x_0) - y_0 \in \Psi_0(v_0) = \Psi_0(x_1 - x_0) = Df(x_0)(x_1 - x_0) + H(x_0)(x_1 - x_0),$$

which is equivalent to

$$\begin{aligned} 0 \in f(x_0) + Df(x_0)(x_1 - x_0) + H(x_0)(x_1 - x_0) + y_0 \\ \subset f(x_0) + Df(x_0)(x_1 - x_0) + H(x_0)(x_1 - x_0) + F(x_0). \end{aligned}$$

As a summary, x_1 obeys the scheme (2) and

$$\|x_1 - x_0\| = \|v_0\| < \min\{\tau s, \varepsilon_0\}.$$

Let h be the function defined as in Lemma 2. It follows from assumption (3) that $\frac{h(\varepsilon_0)}{\varepsilon_0} = \frac{1}{1-\tau(L+\lambda)r}\tau K_0 < 1$. Setting $\alpha_0 = \min\{\tau s, \varepsilon_0\} > 0$, the monotonicity of function $\frac{h(t)}{t}$ shows that

$$h(\alpha_0) \leqslant \frac{h(\varepsilon_0)}{\varepsilon_0}\alpha_0 = q\alpha_0, \quad q = \frac{1}{1 - \tau(L+\lambda)r}\tau K_0. \tag{24}$$

Using Lemma 2, the sequence (α_k) produced by $\alpha_{k+1} = h(\alpha_k)$ is well-defined such that $\alpha_k \leqslant q^k\alpha_0$.

We shall continue by induction. Assume that $k+1$ iterations $x_0, x_1, \ldots, x_k$ are known for some $k \geqslant 1$. Moreover, the following conditions are also supposed to be valid:

– $0 \in f(x_j) + Df(x_j)(x_{j+1} - x_j) + H(x_j)(x_{j+1} - x_j) + F(x_j)$, for $j \leqslant k-1$;
– $\|x_{j+1} - x_j\| < \alpha_j$, $j = 0, \ldots, k-1$.

Since

$$\sum_{j=0}^{k-1} \alpha_j \leqslant \sum_{j=0}^{k-1} q^j \alpha_0 < \frac{1}{1-q}\alpha_0 = \frac{1-\tau(L+\lambda)r}{1-\tau(L+\lambda)r-\tau K_0} \min\left\{\tau s, \varepsilon_0\right\} \leqslant r,$$

all known points x_j are in $\mathbb{B}_X(x, r)$. If $x_k = x_{k-1}$ then we simply take $x_{k+1} = x_k$, and pass directly to the second step. Otherwise, we set $\Psi_k(\cdot) = Df(x_k)(\cdot) + H(x_k)(\cdot)$ and write $v_j = x_{j+1} - x_j$ for each $j \leqslant k-1$. By the inductive hypothesis, there is $y_{k-1} \in F(x_{k-1})$ such that $w_{k-1} \in H(x_{k-1})(v_{k-1})$, where

$$w_{k-1} = -f(x_{k-1}) - Df(x_{k-1})(v_{k-1}) - y_{k-1}.$$

Noticing $\alpha_{k-1} \leqslant q^k \alpha_0 \leqslant \varepsilon_0$, the supposition (21) can be used here. This implies that there exist some elements $y_k \in F(x_k)$ and $w'_{k-1} \in H(x_k)(-v_{k-1})$ with

$$y_{k-1} - y_k - w'_{k-1} \in \rho_0 \|v_{k-1}\| \mathbb{B}_Y.$$

Define $z_k = -f(x_k) - y_k$, we have

$$z_k = \left[-f(x_k) + f(x_{k-1}) + Df(x_{k-1})(v_{k-1})\right] + w_{k-1} + \left(y_{k-1} - y_k - w'_{k-1}\right) + w'_{k-1}.$$

The Taylor's expansion at the reference point x_{k-1}

$$f(x_k) = f(x_{k-1}) + \int_0^1 Df(x_{k-1} + t v_{k-1})(v_{k-1})\, dt$$

yields

$$\begin{aligned}
&\|-f(x_k) + f(x_{k-1}) + Df(x_{k-1})(v_{k-1})\| \\
&= \left\| -\int_0^1 \left[Df(x_{k-1} + t v_{k-1}) - Df(x_{k-1})\right](v_{k-1})\, dt \right\| \\
&\leqslant \int_0^1 L \|v_{k-1}\|^2 t\, dt = \frac{1}{2} L \|v_{k-1}\|^2.
\end{aligned}$$

Taking into account $w_{k-1} + w'_{k-1} \in H(x_{k-1})(v_{k-1}) + H(x_k)(-v_{k-1})$, one obtains

$$\begin{aligned}
\|z_k\| &\leqslant \frac{1}{2} L \|v_{k-1}\|^2 + \rho_0 \|v_{k-1}\| + \Lambda_H(x_{k-1}, x_k, \alpha_{k-1}) \\
&\leqslant \frac{1}{2} L \alpha_{k-1}^2 + \rho_0 \alpha_{k-1} + \omega(\alpha_{k-1}).
\end{aligned}$$

By virtue of $x_k \in \mathbb{B}_X(x, r)$, Lemma 1 tells us

$$\sigma_{\Psi_0, \Psi_k}(u, \delta) \leqslant (L+\lambda)\|x_0 - x_k\|\, \delta < (L+\lambda) r \delta.$$

Setting now $\beta_k = \|x_0 - x_k\| < r$, we check that

$$\begin{cases}
\tau(L+\lambda)\beta_k < 1, \\
\frac{\tau}{1-\tau(L+\lambda)\beta_k}\left[\frac{1}{2} L \alpha_{k-1}^2 + \rho_0 \alpha_{k-1} + \omega(\alpha_{k-1})\right] < r, \\
\frac{1}{2} L \alpha_{k-1}^2 + \rho_0 \alpha_{k-1} + \omega(\alpha_{k-1}) < s.
\end{cases} \tag{25}$$

The first inequality of (25) is an immediate consequence of supposition (3). The second one is inferred from the following estimates

$$\frac{\tau}{1-\tau(L+\lambda)\beta_k}\Big(\frac{1}{2}L\alpha_{k-1}^2+\rho_0\alpha_{k-1}+\omega(\alpha_{k-1})\Big)=\frac{1-\tau(L+\lambda)r}{1-\tau(L+\lambda)\beta_k}h(\alpha_{k-1})$$
$$\leqslant h(\alpha_{k-1})=\alpha_k\leqslant q^k\alpha_0\leqslant\min\{\tau s,\varepsilon_0\}<r.$$

Finally, we check the last inequality

$$\frac{1}{2}L\alpha_{k-1}^2+\rho_0\alpha_{k-1}+\omega(\alpha_{k-1})=\frac{1-\tau(L+\lambda)r}{\tau}h(\alpha_{k-1})=\frac{1-\tau(L+\lambda)r}{\tau}\alpha_k$$
$$\leqslant\frac{1-\tau(L+\lambda)r}{\tau}q^k\alpha_0\leqslant\frac{1-\tau(L+\lambda)r}{\tau}\min\{\tau s,\varepsilon_0\}<s.$$

Due to (25), we can add some parameters $r_k>0$ and $s_k>0$ such that

$$r_k+\frac{\tau}{1-\tau(L+\lambda)\beta_k}s_k<r,\ \frac{1}{2}L\alpha_{k-1}^2+\rho_0\alpha_{k-1}+\omega(\alpha_{k-1})<s_k<s.$$

Invoking Theorem 1, the relation $\tau_k\in\mathrm{REG}\,(\Psi_k,V_k)$ holds with $\tau_k=\frac{\tau}{1-\tau(L+\lambda)\beta_k}$ and

$$V_k=\big\{(v,w):\|v\|\leqslant r_k,d\big(w,\Psi_k(v)\big)\leqslant s_k\big\}.$$

Because of $d\big(z_k,\Psi_k(0)\big)=\|z_k\|$, we conclude $(0,z_k)\in V_k$. Consequently,

$$d\big(0,\Psi_k^{-1}(z_k)\big)\ \leqslant\ \tau_k d\big(z_k,\Psi_k(0)\big)=\tau_k\,\|z_k\|$$
$$\leqslant\ \frac{\tau}{1-\tau(L+\lambda)\beta_{k-1}}\Big(\frac{1}{2}L\alpha_{k-1}^2+\rho_0\alpha_{k-1}+\omega(\alpha_{k-1})\Big).$$

Thus, there exists $v_k\in\Psi_k^{-1}(z_k)$ such that

$$\|v_k\|<\frac{\tau}{1-\tau(L+\lambda)r}\Big(\frac{1}{2}L\alpha_{k-1}^2+\rho_0\alpha_{k-1}+\omega(\alpha_{k-1})\Big)=h(\alpha_{k-1})=\alpha_k.$$

Let us define $x_{k+1}=x_k+v_k$. The choice of v_k implies

$$-f(x_k)-y_k=z_k\in\Psi_k(x_{k+1}-x_k)=Df(x_k)(x_{k+1}-x_k)+H(x_k)(x_{k+1}-x_k).$$

In other words, x_{k+1} clearly satisfies (2), since y_k belongs to $F(x_k)$. As a result, the sequence (x_k) is completely determined by algorithm (2).

Step2: Convergence Because $\sum_{k\geqslant 0}\alpha_k$ is a convergent series, the sequence (x_k) converges. Denoting $x^*=\lim\limits_{k\to\infty}x_k$, we claim that x^* solves (1). In fact, recalling

$$0\in f(x_k)+Df(x_k)(v_k)+H(x_k)(v_k)+F(x_k),\ v_k=x_{k+1}-x_k \tag{26}$$

for each index k. Taking into account $\|x-x_k\|=\beta_k<r$ and $\|v_k\|<\alpha_k\leqslant\varepsilon_0$ when k is large enough, hypothesis (3) implies

$$e\big(H(x)(-v_k)+H(x_k)(v_k),0\big)\leqslant\Lambda_H(x_k,x,\alpha_k)\leqslant\omega(\alpha_k)\xrightarrow{k\to\infty}0.$$

(Here, condition $\omega(0^+)=0$ in (3) was used.) Since $|H(x)|^+<+\infty$, the fact that $H(x)(u)\subset|H(x)|^+\,\|u_k\|\,\bar{\mathbb{B}}_Y$ gives us

$$e\big(H(x)(-v_k),0\big)\leqslant|H(x)|^+\,\|v_k\|\xrightarrow{k\to\infty}0.$$

Hence, $e\big(H(x_k)(v_k),0\big)$ tends to 0 as $k\to\infty$. Consequently, by passing to the limit in (26), we obtain $0\in f(x^*)+F(x^*)$.

To finish the proof, we verify $\|x - x^*\| \leqslant r$ as well as $\limsup_{k\to\infty} \|x_k - x^*\|^{1/k} < 1$. Indeed, we infer from the construction

$$\|x_k - x^*\| \leqslant \sum_{j\geqslant k} \|x_j - x_{j+1}\| \leqslant \sum_{j\geqslant k} \alpha_j \leqslant q^k \sum_{j\geqslant 0} q^j \alpha_0 = q^k \frac{\alpha_0}{1-q}. \tag{27}$$

Substituting $k = 0$ in (27), we deduce

$$\|x - x^*\| = \|x_0 - x^*\| \leqslant \frac{\alpha_0}{1-q} = \frac{1 - \tau(L+\lambda)r}{1 - \tau K_0 - \tau(L+\lambda)r} \min\{\tau s, \varepsilon_0\} \leqslant r.$$

Furthermore, extracting the k-th root in (27), and letting $k \to \infty$ we find

$$\limsup_{k\to\infty} \|x_k - x^*\|^{1/k} \leqslant \limsup_{k\to\infty} \left\{ q \left(\frac{\alpha_0}{1-q} \right)^{1/k} \right\} = q < 1,$$

the proof of Theorem 3 is completed. □

We end this section by studying the local behavior of the scheme (2).

Theorem 4 (local convergence) *Suppose that problem* (1) *admits $x^* \in X$ as a solution. Given a map $H : X \longrightarrow \mathcal{PH}(X, Y)$ so that F is pointwise strictly differentiable with respect to H on a neighborhood of x^*. Associated with such a map H, we assume there are two increasing continuous functions $\rho, \varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ and a radius $r > 0$ satisfying*

- $F(x') \subset F(x) + H(x)(x' - x) + \rho(\|x' - x\|)\mathbb{B}_Y$ *whenever* $x, x' \in \mathbb{B}_X(x^*, r)$,
- $\sigma_{H(x^*),H(x)}(u, \varepsilon) \leqslant \varphi(\|x - x^*\|)\,\varepsilon$, *for all* $u \in r\bar{\mathbb{B}}_X$ *and* $\varepsilon \leqslant r$,
- $\theta^* = \limsup_{t\to 0} \frac{\rho(t)}{t} < +\infty$.

In addition, the assumptions below are supposed to be fulfilled:

(i) $\Phi_{x^*}(\cdot) := Df(x^*)(\cdot) + H(x^*)(\cdot)$ *is metrically regular on the neighborhood* $V^* = r\mathbb{B}_X \times s\mathbb{B}_Y$ *with a modulus* $\tau^* > 0$;
(ii) $\tau^*(\theta^* + \varphi(0)) < 1$;
(iii) Df *is Lipschitz continuous while* $|H(\cdot)|^+$ *is finite on the ball* $\mathbb{B}_X(x^*, r)$.

Then, there exists a constant $0 < \alpha < r$ for which the following statement holds. Given $x \in \mathbb{B}_X(x^*, \alpha)$, it has a sequence (x_k) generated by our scheme (2) *such that $x_0 = x$ and $x_k \to x^*$ as $k \to \infty$ at least linearly.*

Before proving this theorem, we note that the function ρ above may quite often appear when F is pointwise strictly differentiable with respect to H. Besides that, the assumption related to the function φ might be valid in many situations, for instance, when H is induced by a set-valued map from X into the space $\mathcal{L}(X, Y)$ of all linear and continuous maps between X and Y.

Let us present the proof of Theorem 4. For this purpose, the next lemmas will be useful.

Lemma 3 *Suppose that all statements of Theorem 4 are satisfied. Given $\bar{x} \in \mathbb{B}_X(x^*, r)$ and set $\Psi_{x^*}(\cdot) := Df(x^*)(\cdot - \bar{x}) + H(x^*)(\cdot - \bar{x})$, $\Psi_{\bar{x}}(\cdot) := Df(\bar{x})(\cdot - \bar{x}) + H(\bar{x})(\cdot - \bar{x})$. If L is a Lipschitz continuity modulus for Df on $\mathbb{B}_X(x^*, r)$, then*

$$\sigma_{\Psi_{x^*},\Psi_{\bar{x}}}(x, \varepsilon) \leqslant \left[L \|\bar{x} - x^*\| + \varphi\left(\|\bar{x} - x^*\| \right) \right] \varepsilon, \forall x \in \mathbb{B}_X(\bar{x}, r), \varepsilon \leqslant r. \tag{28}$$

Proof Based on triangle inequality, the definition of $\sigma_{\Psi_{x^*},\Psi_{\bar{x}}}(x,\varepsilon)$ gives us

$$\begin{aligned}\sigma_{\Psi_{x^*},\Psi_{\bar{x}}}(x,\varepsilon) &\leqslant \sup_{\|x'-x\|\leqslant\varepsilon}\left\|\left[Df(\bar{x})-Df(x^*)\right](x-x')\right\| \\ &\quad+\sup_{\zeta\in H(\bar{x})(x-\bar{x})}\inf_{\zeta'\in H(x^*)(x-\bar{x})}\sup_{\substack{\|x'-x\|\leqslant\varepsilon,\\ \xi\in H(x^*)(x'-\bar{x})}}\inf_{\xi'\in H(\bar{x})(x'-\bar{x})}\left\|\zeta-\zeta'+\xi-\xi'\right\| \\ &= \sup_{\|x'-x\|\leqslant\varepsilon}\left\|\left[Df(\bar{x})-Df(x^*)\right](x-x')\right\|+\sigma_{H(x^*),H(\bar{x})}(x-\bar{x},\varepsilon) \\ &\leqslant \sup_{\|x'-x\|\leqslant\varepsilon}\left\|Df(\bar{x})-Df(x^*)\right\|\left\|x-x'\right\|+\sigma_{H(x^*),H(\bar{x})}(x-\bar{x},\varepsilon) \\ &= \left\|Df(\bar{x})-Df(x^*)\right\|\varepsilon+\sigma_{H(x^*),H(\bar{x})}(x-\bar{x},\varepsilon).\end{aligned}$$

Thanks to the Lipschitz continuity of Df as well as $u=x-\bar{x}\in r\mathbb{B}_X$, we attain (28). □

Lemma 4 *Given $0<\delta<r$ and $0<\delta'<s$. Let $\bar{x}\in\bar{\mathbb{B}}_X(x^*,r-\delta)$ and consider $\Psi_{\bar{x}}(\cdot)=Df(x^*)(\cdot-\bar{x})+H(x^*)(\cdot-\bar{x})$. If $z^*\in\Psi_{\bar{x}}(x^*)$ satisfying $\|z^*\|\leqslant s-\delta'$, then one has $\tau^*\in\mathrm{REG}\big(\Psi_{\bar{x}},V\big)$, for $V=\mathbb{B}_X(x^*,\delta)\times\mathbb{B}_Y(z^*,\delta')$.*

Proof Given $(x,z)\in V$, and set $u=x-\bar{x}$. It is possible to check that $\Psi_{\bar{x}}^{-1}(z)=\bar{x}+\Phi_{x^*}^{-1}(z)$, where Φ_{x^*} is in assumption of Theorem 4. Hence,

$$d\big(x,\Psi_{\bar{x}}^{-1}(z)\big)=d\big(x,\bar{x}+\Phi_{x^*}^{-1}(z)\big)=d\big(u,\Phi_{x^*}^{-1}(z)\big).$$

But (u,z) is in the neighborhood V^*, since $\|x-\bar{x}\|\leqslant\|x-x^*\|+\|x^*-\bar{x}\|<r$ and $\|z\|\leqslant\|z-z^*\|+\|z^*\|<s$. Thus, we obtain

$$d\big(x,\Psi_{\bar{x}}^{-1}(z)\big)=d\big(u,\Phi_{x^*}^{-1}(z)\big)\leqslant\tau^*d\big(z,\Phi_{x^*}(x-\bar{x})\big)=\tau^*d\big(z,\Psi_{\bar{x}}(x)\big),$$

and the proof is done. □

Lemma 5 *Keep in mind all assumptions of Theorem 4. If $x\in\mathbb{B}_X(x^*,r)$ and $\|u\|<r$ then one has*

$$H(x)(u)\subset\Big(\big|H(x^*)\big|^+\varphi(a)\Big)\|u\|\,\bar{\mathbb{B}}_Y,\tag{29}$$

where $\|x-x^*\|<a<r$. In particular, we infer from (29) *the following estimation*

$$\inf_{\upsilon\in H(x^*)(u)}\sup_{\upsilon'\in H(x)(u)}\left\|(\upsilon'-\bar{\upsilon}')-(\upsilon-\bar{\upsilon})\right\|\leqslant\Big(2\big|H(x^*)\big|^+\varphi(a)\Big)\|u\|+\left\|\bar{\upsilon}'-\bar{\upsilon}\right\|\tag{30}$$

for $\bar{\upsilon}\in Y$ and $\bar{\upsilon}'\in Y$.

Proof Pick $b>0$ with $\|u\|<b<r$. Using the condition $\sigma_{H(x^*),H(x)}(u,\varepsilon)\leqslant\varphi(a)\varepsilon$ for $\varepsilon=\frac{1}{2}(b+\|u\|)>\|u\|$, we get

$$\sup_{\eta\in H(x)(u)}\inf_{\eta'\in H(x^*)(u)}\sup_{\substack{\|u'-u\|\leqslant\varepsilon,\\ \xi\in H(x^*)(u')}}\inf_{\xi'\in H(x)(u')}\left\|\eta-\eta'+\xi-\xi'\right\|\leqslant\varphi(a)\varepsilon<\varphi(a)b.$$

Fix $\eta\in H(x)(u)$. Then there is $\eta'\in H(x^*)(u)$ such that

$$\sup_{\|u'-u\|\leqslant\varepsilon,\xi\in H(x^*)(u')}\inf_{\xi'\in H(x)(u')}\left\|\eta-\eta'+\xi-\xi'\right\|<\varphi(a)b.\tag{31}$$

Substituting $u' = 0$, and noticing that $H(x^*)(0) = H(x)(0) = \{0\}$, (31) permits us to write $\|\eta - \eta'\| < \varphi(a)b$. Since b does not depend on η and η', one deduces $\|\eta - \eta'\| \leqslant \varphi(a)\|u\|$ by letting $b \to \|u\|$. In other words,

$$H(x)(u) \subset H(x^*)(u) + \varphi(a)\|u\|\bar{\mathbb{B}}_Y.$$

Finally, taking into account

$$H(x^*)(u) \subset \left|H(x^*)\right|^+ \|u\|\bar{\mathbb{B}}_Y, \tag{32}$$

the inclusion (29) is thereby valid. Upon relation (30), we have just to apply (29) and invoke (32) again. □

Proof of Theorem 4 Under assumptions of Theorem 4, it is possible to take $\theta > \theta^*$ and $0 < \alpha < r/2$ which fulfill the constraints below

$$\begin{cases} \rho(t) \leqslant \theta t, \forall t \leqslant \alpha, \\ \tau^*\left(\frac{3}{2}L\alpha + \varphi(\alpha) + \theta\right) < 1, \\ \left[2L\alpha + 3\varphi(\alpha) + 4\left|H(x^*)\right|^+ + \nu^*\right]\alpha + 2\left(\frac{1}{2}L\alpha^2 + \rho(\alpha)\right) < s, \end{cases} \tag{33}$$

where $\nu^* = \|Df(x^*)\| + |H(x^*)|^+$, and $L > 0$ is a Lipschitz modulus for Df on $\mathbb{B}_X(x^*, r)$.

Let's fix $x \in \mathbb{B}_X(x^*, \alpha)$ and denote $x_0 = x$. Set $u_0 = x_0 - x^*$, $\alpha_0 = \|u_0\| < \alpha$, $\mu_0 = L\alpha_0 + \varphi(\alpha_0)$, and $\Phi_0(\cdot) = Df(x^*)(\cdot - x_0) + H(x^*)(\cdot - x_0)$. By assumptions of Theorem 4,

$$-f(x^*) \in F(x^*) \subset F(x_0) + H(x_0)(x^* - x_0) + \rho\left(\left\|x^* - x_0\right\|\right)\mathbb{B}_Y,$$

so we can select $y_0 \in F(x_0)$ and $w_0 \in H(x_0)(-u_0)$ such that

$$-f(x^*) - y_0 - w_0 \in \rho\left(\left\|x^* - x_0\right\|\right)\mathbb{B}_Y = \rho(\alpha_0)\mathbb{B}_Y. \tag{34}$$

Moreover, observe that the relation

$$\inf_{\xi \in H(x^*)(-u_0)} \sup_{\substack{\|z+u_0\| \leqslant \alpha_0, \\ \upsilon \in H(x^*)(z)}} \inf_{\upsilon' \in H(x_0)(z)} \left\|w_0 - \xi + \upsilon - \upsilon'\right\| \leqslant \sigma_{H(x^*),H(x_0)}(-u_0, \alpha_0) \leqslant \varphi(\alpha_0)\alpha_0$$

holds. Due to this relation, we choose $w_0^* \in H(x^*)(-u_0)$ for which

$$\sup_{\|z+u_0\| \leqslant \alpha_0, \upsilon \in H(x^*)(z)} \inf_{\upsilon' \in H(x_0)(z)} \left\|w_0 - w_0^* + \upsilon - \upsilon'\right\| < \varphi(\alpha)\alpha \tag{35}$$

is valid as well. Both $|H(x^*)|^+$ and $|H(x_0)|^+$ are finite, so $H(x^*)(0) = H(x_0)(0) = \{0\}$. Taking $z = 0$ in the left-hand side of (35), we conclude $\left\|w_0 - w_0^*\right\| < \varphi(\alpha)\alpha$.

Define $z_0 = -f(x_0) - y_0$, $\bar{z}_0 = -Df(x_0)(u_0) + w_0$ and $z_0^* = -Df(x^*)(u_0) + w_0^*$. Since $w_0^* \in H(x^*)(-u_0) \subset |H(x^*)|^+ \|u_0\|\bar{\mathbb{B}}_Y$, one has

$$\left\|z_0^*\right\| \leqslant \left\|Df(x^*)\right\|\|u_0\| + \left|H(x^*)\right|^+\|u_0\| = \nu^*\alpha_0.$$

According to (33), $\nu^*\alpha_0 < \nu^*\alpha < s$. On the other hand, by $\alpha_0 < \alpha < r/2$, it holds that $r - \alpha > \alpha_0$. Using Lemma 4 for $\delta = \alpha$ and $\delta' = s - \nu^*\alpha_0$, the mapping Φ_0 is metrically regular on $V_0 = \mathbb{B}_X(x^*, \alpha) \times \mathbb{B}_Y(z_0^*, s - \nu^*\alpha_0)$ with modulus τ^*.

To obtain the next iteration x_1, we are going to apply Theorem 2. Let $\Psi_0(\cdot) := Df(x_0)(\cdot - x_0) + H(x_0)(\cdot - x_0)$, we get $\bar{z}_0 \in \Psi_0(x^*)$. If $x \in \mathbb{B}_X(x^*, \alpha)$, then we find

$$\begin{aligned}\gamma(x, x_0) &:= \inf_{\upsilon \in \Phi_0(x)} \sup_{\upsilon' \in \Psi_0(x)} \left\| (\upsilon' - \bar{z}_0) - (\upsilon - z_0^*) \right\| \\ &\leqslant \left\| [Df(x_0) - Df(x^*)](x - x^*) \right\| \\ &\qquad + \inf_{v \in H(x^*)(x - x_0)} \sup_{w \in H(x_0)(x - x_0)} \left\| (w - w_0) - (v - w_0^*) \right\|.\end{aligned}$$

Thanks to Lemma 5, one infers from the Lipschitz continuity of Df that

$$\begin{aligned}\gamma(x, x_0) &\leqslant L \left\| x_0 - x^* \right\| \left\| x - x^* \right\| + \left[2 \left| H(x^*) \right|^+ + \varphi(\alpha) \right] \|x - x_0\| \\ &\qquad + \left\| w_0 - w_0^* \right\| \\ &\leqslant 2L\alpha\alpha_0 + 2 \left[2 \left| H(x^*) \right|^+ + \varphi(\alpha) \right] \alpha + \varphi(\alpha)\alpha.\end{aligned} \tag{36}$$

In (36), the last estimation is followed by $\|x - x_0\| \leqslant \|x - x^*\| + \|x^* - x_0\| < 2\alpha$. Furthermore, the choice of z_0 and $\bar{z}_0$ provides

$$z_0 - \bar{z}_0 = f(x^*) - f(x_0) - Df(x_0)(-u_0) - f(x^*) - y_0 - w_0.$$

Writing down the Taylor expansion for f at x_0, we deduce

$$\begin{aligned}f(x^*) &= f(x_0) + \int_0^1 Df(x_0 - tu_0)(-u_0)\, dt \\ &= f(x_0) + Df(x_0)(-u_0) + \int_0^1 \left[Df(x_0 - tu_0) - Df(x_0) \right](-u_0)\, dt.\end{aligned}$$

As a consequence,

$$\begin{aligned}\|z_0 - \bar{z}_0\| &\leqslant \int_0^1 \left\| \left[Df(x_0 - tu_0) - Df(x_0) \right](-u_0) \right\| dt \\ &\qquad + \|-f(x^*) - y_0 - w_0\| \\ &< \int_0^1 L \|u_0\|^2 t\, dt + \rho(\alpha_0) = \tfrac{1}{2} L\alpha_0^2 + \rho(\alpha_0).\end{aligned} \tag{37}$$

In (37), the fact $-f(x^*) - y_0 - w_0 \in \rho(\alpha_0)\mathbb{B}_Y$ was used. We are now going to prove the inequalities below

$$\begin{cases} \mu_0\tau^* = [L\alpha_0 + \varphi(\alpha_0)]\,\tau^* < 1, \\ \frac{\tau^*}{1 - \mu_0\tau^*} \left(\frac{1}{2} L\alpha_0^2 + \rho(\alpha_0) \right) < \alpha, \\ (1 + \mu_0\tau^*) \left(\frac{1}{2} L\alpha_0^2 + \rho(\alpha_0) \right) + \sup_{\|x - x^*\| < \alpha} \gamma(x, x_0) < s - \nu^*\alpha_0. \end{cases} \tag{38}$$

In fact, if we set $\beta = \frac{1}{2} L\alpha^2 + \rho(\alpha)$ and $\beta_0 = \frac{1}{2} L\alpha_0^2 + \rho(\alpha_0)$, the first and second conditions in (33) gives us

$$\tau^*\mu_0 \leqslant \tau^*[L\alpha + \varphi(\alpha)] < 1 - \tau^* \left(\frac{1}{2} L\alpha + \theta \right) \leqslant 1 - \tau^*\alpha^{-1}\beta < 1 - \tau^*\alpha^{-1}\beta_0. \tag{39}$$

Taking into account (33), (36) and (39), the constraints in (38) are valid as well. This permits us to add a parameter $s_0 > \frac{1}{2} L\alpha_0^2 + \rho(\alpha_0)$ satisfying

$$\frac{\tau^*}{1 - \mu_0\tau^*} s_0 < \alpha,\ \left(1 + \mu_0\tau^*\right) s_0 + \sup_{\|x - x^*\| < \alpha} \gamma(x, x_0) < s - \nu^*\alpha_0.$$

Define $\tau_0 = \frac{\tau^*}{1-\mu_0\tau^*}$. Let's apply Theorem 2 with the corresponding data $\bar{x} = x^*$, $\bar{y} = z_0^*$, $V = \mathbb{B}_X(x^*, \alpha) \times \mathbb{B}_Y(z_0^*, s - \nu^*\alpha_0)$, $\bar{z} = \bar{z}_0$, $\nu = s_0$. By virtue of (37), we obtain $\|z_0 - \bar{z}_0\| < s_0$, which implies that

$$d\big(x^*, \Psi_0^{-1}(z_0)\big) \leqslant \tau_0 d\big(z_0, \Psi_0(x^*)\big) \leqslant \tau_0 \|z_0 - \bar{z}_0\| < \frac{\tau^*}{1-\mu_0\tau^*}\left(\frac{1}{2}L\alpha_0^2 + \rho(\alpha_0)\right).$$

Hence, the set $\Psi_0^{-1}(z_0)$ contains some x_1 such that

$$\begin{aligned}\|x^* - x_1\| &< \frac{\tau^*}{1-\mu_0\tau^*}\left(\frac{1}{2}L\alpha_0^2 + \rho(\alpha_0)\right)\\ &= \frac{1}{1-[\tau^*L\alpha_0 + \tau^*\varphi(\alpha_0)]}\left[\frac{1}{2}\tau^*L\alpha_0^2 + \tau^*\rho(\alpha_0)\right].\end{aligned} \tag{40}$$

The function $\psi(t) = \frac{1}{1-[\tau^*Lt+\tau^*\varphi(t)]}\left[\frac{1}{2}\tau^*Lt + \tau^*\theta\right]$ is well-defined and nondecreasing on the interval $[0, \alpha]$. The choice of α gives us $\psi(\alpha) < 1$. Due to (33) and (40), we get

$$\|x^* - x_1\| < \psi(\alpha_0)\alpha_0 \leqslant \psi(\alpha)\alpha_0 \leqslant \alpha_0,$$

which yields $x_1 \in \mathbb{B}_X(x^*, \alpha)$. Consequently, we can apply all previous arguments for the new starting point x_1 instead of x_0.

Recall that $x_1 \in \Psi_0^{-1}(z_0)$, or equivalently, $z_0 = -f(x_0) - y_0 \in \Psi_0(x_1)$. Thus,

$$\begin{aligned}0 \in f(x_0) + y_0 + Df(x_0)(x_1 - x_0) + H(x_0)(x_1 - x_0)&\\ \subset f(x_0) + Df(x_0)(x_1 - x_0) + H(x_0)(x_1 - x_0) + F(x_0).&\end{aligned}$$

In other words, x_1 is generated by (2).

Repeating this process, we obtain a sequence (x_k) which is produced by (2) and

$$\|x^* - x_{k+1}\| \leqslant \psi\left(\|x^* - x_k\|\right)\|x^* - x_k\| \leqslant \psi(\alpha)\|x^* - x_k\|. \tag{41}$$

The proof is thereby done. □

In the previous theorem, the rate of convergence for (x_k) seems to concern the behavior of the function $t \longmapsto \rho(t)$ around $t = 0$. If a stronger condition is imposed on ρ (i.e., on the order of approximation for F), then Theorem 4 can be refined a little bit. The next corollary is in this sense.

Corollary 1 (local convergence revisited) *Suppose all assumptions of Theorem 4 are fulfilled with* $\theta^* = \limsup_{t\to 0} \frac{\rho(t)}{t} = 0$. *Then, the constant* α *can be chosen such that the sequence* (x_k) *in Theorem 4 converges to the solution* x^* *of* (1) *superlinearly.*

Proof Let's write $\rho(t) = \rho_1(t)t$ where $\rho_1 : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ is a function satisfying $\limsup_{t\to 0} \rho_1(t) = 0$. Then there exists $\alpha \in (0, r/2)$ such that

$$\begin{cases}\frac{3}{2}L\alpha + \varphi(\alpha) + \sup_{0\leqslant t\leqslant\alpha}\rho_1(t) < \frac{1}{\tau^*}, \forall t \leqslant \alpha,\\ \left[2L\alpha + 3\varphi(\alpha) + 4\,|H(x^*)|^+ + \nu^*\right]\alpha + 2\left(\frac{1}{2}L\alpha^2 + \rho(\alpha)\right) < s.\end{cases} \tag{42}$$

Here, ν^* is again the quantity $\|Df(x^*)\| + |H(x^*)|^+$. Indeed, (42) can be subsumed as a particular case of (33), where $\sup_{0\leqslant t\leqslant\alpha}\rho_1(t)$ plays the same role as θ.

Now, let x be in $\mathbb{B}_X(x^*, \alpha)$. Repeating the construction as in the proof of Theorem 4, the algorithm (2) generates a sequence (x_k) for which $x_0 = x$ and

$$\|x^* - x_{k+1}\| \leqslant \psi_1\left(\|x^* - x_k\|\right)\|x^* - x_k\|, k = 0, 1, \ldots \tag{43}$$

where $\psi_1(t) := \frac{1}{1-[\tau^* Lt + \tau^* \varphi(t)]}\left[\frac{1}{2}\tau^* Lt + \tau^* \rho_1(t)\right], t \in [0, \alpha]$. By induction, we can prove $\|x^* - x_k\| \leqslant \alpha$ and $0 \leqslant \sup_k \psi_1\left(\|x^* - x_k\|\right) < 1$. Particularly, (43) ensures that x_k converges to x^* as $k \to \infty$. Finally, by taking into account

$$\limsup_{t\to 0} \psi_1(t) = \limsup_{t\to 0}\left\{\frac{1}{1-[\tau^* Lt + \tau^* \varphi(t)]}\left[\frac{1}{2}\tau^* Lt + \tau^* \rho_1(t)\right]\right\} = 0,$$

the superlinear convergence follows, which completes the proof of Corollary 1. □

4 Some Illustrative Examples

Example 1 The purpose of this example is to give a simple comparison between (2) and the well-known Josephy-Newton method. Consider $X = Y = \mathbb{R}$

$$f(x) = -(x-1)^3 + x - 1 \text{ and } F(x) = \begin{cases} [\exp(-2x), +\infty), & \text{if } x \geqslant 0, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{44}$$

Figure 1 depicts the graphs of $-f$ and F, respectively. Choose $H(x)(u) = \{-2u\exp(-2x)\}$ and define

$$\rho(t) = \exp(2t) - 2t - 1, \varrho(t) = 2\left[\exp(2t) - 1\right], \omega(t) = 4t. \tag{45}$$

If we set $g(x) = \exp(-2x)$, then $H(x)(u) = \left\{g'(x)u\right\}$, which permits us to prove that

$$\begin{aligned} &e\big(F(x'), F(x) + H(x)(x' - x)\big) \\ &\quad = e\left(\left[g(x'), +\infty\right), \left[g(x) + g'(x)(x' - x), +\infty\right)\right) \\ &\quad \leqslant \left|g(x') - \left[g(x) + g'(x)(x' - x)\right]\right|; x' \geqslant 0, x \geqslant 0. \end{aligned} \tag{46}$$

Denote $u = x' - x$, we infer from (46) that

$$\begin{aligned} e\big(F(x'), F(x) + H(x)(x' - x)\big) &\leqslant \exp(-2x)\left|\exp(-2u) + 2u - 1\right| \\ &\leqslant \sum_{n\geqslant 2} \frac{(2|u|)^n}{n!} = \exp(2|u|) - 2|u| - 1 \\ &= \rho\left(\left|x' - x\right|\right); x \geqslant 0, x' \geqslant 0. \end{aligned}$$

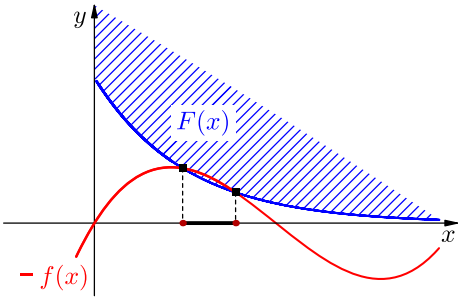

Fig. 1 Graphs of $-f$ and F in Example 1

On the other hand, we have

$$\Lambda_H(x, x', t) = 2\left|\exp(-2x) - \exp(-2x')\right| t \leqslant \omega(t).$$

According to the definition of H, we obtain

$$\sigma_{H(x),H(x')}(u, r) \leqslant \varrho\left(\left|x - x'\right|\right) r, \quad \text{for all } u \geqslant 0 \text{ and } r > 0. \tag{47}$$

Since $\lim\limits_{t\to 0} \frac{\rho(t)}{t^2} = 2$ and $\lim\limits_{t\to 0} \frac{\varrho(t)}{t} = 4$, both the local and semi-local convergence theorems (i.e. Theorems 3 and 4) can be applied to the current case.

Recall that the Josephy-Newton scheme for generalized equation $0 \in f(x) + F(x)$ is of the form

$$0 \in f(x_k) + Df(x_k)(x_{k+1} - x_k) + F(x_{k+1}), k = 0, 1, \ldots, \tag{48}$$

with f and F given in (44). Clearly, (48) is equivalent to solve the following nonlinear inequality

$$\begin{aligned} &\text{find } x \in \mathbb{R}_+ \text{ s.t.} \\ &-(x_k - 1)^3 + x_k - 1 + \left[-3(x_k - 1)^2 + 1\right](x - x_k) + \exp(-2x) \leqslant 0. \end{aligned} \tag{49}$$

On the other hand, the scheme (2) simply reads

$$\text{find } x \in \mathbb{R}_+ \text{ s.t. } f(x_k) + \left[f'(x_k) + g'(x_k)\right](x - x_k) + g(x_k) \leqslant 0. \tag{50}$$

The linear inequality (50) is much easier to solve than the nonlinear one in (49). This example shows that (2) could be much practical to use in some situation (where the computation of H is easy) than the classical Josephy-Newton method. Each approach has its advantages and drawbacks.

Example 2 (An application in finite dimensional spaces) Consider the inclusion (1) with $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ and

$$F(x) = \left\{g(x)w : w \in K\right\} = g(x)K. \tag{51}$$

Here and in what follows, both $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $g : \mathbb{R}^n \longrightarrow \mathbb{R}^{m\times p}$ are assumed to be $C^{1,1}$, while $K \subset \mathbb{R}^p$ is a closed and bounded nonempty subset.

Let $\bar{x} \in \mathbb{R}^n$ be a fixed point and let L_f and L_g be the Lipschitz moduli for f and g, respectively. We set

$$H^z_{\bar{x}}(u) = \left[Dg(z)u - Dg(\bar{x})u\right]K,\ H_{\bar{x}}(z)(u) = [Dg(\bar{x})u]K + H^z_{\bar{x}}(u). \tag{52}$$

Then, one has $H_{\bar{x}}(z) = H_{\bar{x}}(\bar{x}) + H^z_{\bar{x}}$, and hence,

$$\sigma_{H_{\bar{x}}(\bar{x}),H_{\bar{x}}(z)}(u, \epsilon) \leqslant \sup_{\|v-u\|\leqslant\epsilon} e\left(H^z_{\bar{x}}(u), H^z_{\bar{x}}(v)\right). \tag{53}$$

If η is a positive constant such that $K \subset \eta\mathbb{B}$, then by the definition of $H^z_{\bar{x}}$, we derive

$$\begin{aligned} e\left(H^z_{\bar{x}}(u), H^z_{\bar{x}}(v)\right) &\leqslant \sup_{w\in K} \left\|\left(Dg(z) - Dg(\bar{z})\right)(u - v)\right\| \|w\| \\ &\leqslant \eta\|Dg(z) - Dg(\bar{x})\|\|u - v\|. \end{aligned}$$

This implies that

$$\sigma_{H(\bar{x}),H(z)}(u, \epsilon) \leqslant \left(\eta L_g\right)\|z - \bar{x}\|\epsilon. \tag{54}$$

On the other hand, due to the definition of $H_{\bar{x}}(z)$, it is not difficult to verify that

$$|H_{\bar{x}}(z)|^+ \leqslant \eta\|Dg(\bar{x})\| + \eta L_g\|z - \bar{x}\|. \tag{55}$$

As a result, the extended real-valued function $z \longmapsto |H_{\bar{x}}(z)|^+$ is bounded in some neighborhood U of $\bar{x}$. Thus, it is possible to take $\omega(t) = 2\big(\sup_{z\in U} |H_{\bar{x}}(z)|^+\big)t$ where U is a neighborhood of $\bar{x}$.

Finally, let's check that the $H(\cdot)$-differentiability assumptions (Definition 1) are satisfied. Let z and z' be in a neighborhood $U = \mathbb{B}(\bar{x}, r)$. Using Taylor's expansion for g at z, we get

$$g(z') - g(z) - Dg(z)(z' - z) = \int_0^1 [Dg(z + t(z' - z)) - Dg(z)](z' - z)\, dt,$$

which gives us

$$g(z') - g(z) - Dg(z)(z' - z) \in \frac{L_g}{2}\|z' - z\|^2\bar{\mathbb{B}}.$$

Therefore,

$$g(z')K \subset g(z)K + [Dg(z)(z' - z)]K + \frac{1}{2}\eta L_g\|z' - z\|^2\bar{\mathbb{B}}. \tag{56}$$

Nevertheless, by setting $u = z' - z$, we are able to claim that $[Dg(z)(u)]K \subset [Dg(\bar{x})u]K + \big[Dg(z)u - Dg(\bar{x})u\big]K$, and this allows us to write

$$F(z') \subset F(z) + H_{\bar{x}}(z)(z' - z) + \frac{1}{2}\eta L_g\|z' - z\|^2\bar{\mathbb{B}};\ \forall z', z \in U. \tag{57}$$

Consequently, our semi-local result (resp. local result) might be applied by setting $\bar{x} = x_0$ in Theorem 3 (resp. $\bar{x} = x^*$ in Theorem 4 and Corollary 1). All assumptions of those corresponding results are satisfied except the metric regularity property that we must establish for each specific implementation.

As an illustration, let's consider a simple example with $m = n = p = 1$, $K = [1, 2]$ and

$$f(x) = -0.5x^2 - 1,\ g(x) = \exp(x) - 0.5,\ F(x) = g(x)K;\ x \in \mathbb{R}, \tag{58}$$

where $\exp(\cdot)$ indicates the usual exponential function $\exp(x) = e^x,\ x \in \mathbb{R}$. Figure 2 depicts $-f$ and F.

It is easy to prove that $x^* = 0$ is a solution of the inclusion $0 \in f(x) + F(x)$. The mapping $\Psi_{x^*}(v) = g'(x^*)v[1, 2]$ is metrically regular around $(0, 0)$ with a modulus being not less than

$$\tau^* = \max_{\lambda\in[1,2]} \left\{\big|f'(x^*) + \lambda g'(x^*)\big|^{-1}\right\} = \exp(-x^*) = 1.$$

For these aforementioned data, we have $L_f = 1$, $L_g = \exp(x^* + r)$, $r > 0$. The assumptions of Theorem 4 (and also Corollary 1) are satisfied with $\rho(t) = 2L_g t^2$ and $\varphi(t) = 2L_g t$. Some numerical computations are shown in Table 1.

Example 3 (An application in mathematical programming) Let $\phi_0 : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $\phi_1 : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be given smooth maps. Consider the following constrained optimization problem

$$\min\ \ \phi_0(x) \quad \text{s.t.} \quad x \in \mathbb{R}^n, \phi_1(x) \in C \tag{59}$$

where $\emptyset \neq C \subset \mathbb{R}^m$ is a closed convex set. The first-order optimality condition for (59) reads

$$0 \in \nabla\phi_0(x) + N_{\phi_1^{-1}(C)}(x). \tag{60}$$

Here, $\nabla\phi_0$ and N_S respectively denote the Jacobian of ϕ_0 and the normal cone mapping to S. Suppose that x^* is a solution of (60), and that $\nabla\phi_1(x^*)$ is surjective. Then one has $N_{\phi_1^{-1}(C)}(x^*) = \nabla\phi_1(x^*)^T N_C\big(\phi_1(x^*)\big)$ (see [22]), which implies

$$0 \in \nabla\phi_1(x^*)\nabla\phi_0(x^*) + \nabla\phi_1(x^*)\nabla\phi_1(x^*)^T N_C\big(\phi_1(x^*)\big). \tag{61}$$

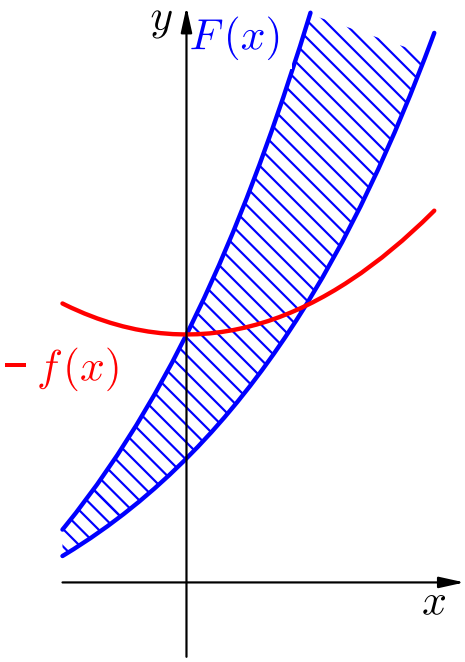

Fig. 2 Graph of $-f$ and F in Example 2

As a result,

$$0 \in \left(\nabla\phi_1(x^*)\nabla\phi_1(x^*)^T\right)^{-1}\nabla\phi_1(x^*)\nabla\phi_0(x^*) + N_C\left(\phi_1(x^*)\right). \tag{62}$$

The map $x \longmapsto \left(\nabla\phi_1(x)\nabla\phi_1(x)^T\right)^{-1}\nabla\phi_1(x)\nabla\phi_0(x)$ is well-defined and bounded around x^*. Thus, it follows from (61) and (62) that x^* can be viewed as a solution of the derived inclusion

$$0 \in \nabla\phi_1(x)\nabla\phi_0(x) + \nabla\phi_1(x)\nabla\phi_1(x)^T\left(N_C\left(\phi_1(x)\right) \cap \eta\mathbb{B}\right) \tag{63}$$

for some suitable $\eta > 0$. If we set $f(x) = \nabla\phi_1(x)\nabla\phi_0(x)$, $g(x) = \nabla\phi_1(x)\nabla\phi_1(x)^T$, $G(x) = N_C\left(\phi_1(x)\right) \cap \eta\mathbb{B}$ and $F(x) = g(x)G(x)$, then (63) is of the form (1).

Note that $G(x) \subset \eta\mathbb{B}$, so it is possible to subsume the problem studied in Example 2 to the current case. In some applications, we are able to obtain the informations related to $G(x^*)$ even if x^* is unknown. An important illustration for such a situation appears when C is defined via a system of equalities and inequalities

$$C = \left\{z \in \mathbb{R}^m \;\middle|\; c_i(z) = 0, i \in \mathcal{E},\ c_j(z) \leqslant 0, j \in \mathcal{I},\right\}$$

associated with smooth functions c_i, $i \in \mathcal{E} \cup \mathcal{I}$. Under certain constraint qualification criteria, the normal cone N_C might be totally determined through the active set $\mathcal{A}(z) := \left\{i \in \mathcal{E} \cup \mathcal{I} : c_i(z) = 0\right\}$ as follows

$$N_C(z) = \left\{v = \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i \nabla c_i(z) \;\middle|\; \lambda_i \leqslant 0, i \in \mathcal{A}(z),\ \lambda_i = 0, i \notin \mathcal{A}(z)\right\}.$$

Table 1 Numerical illustrations, $e_k = |x_k - x^*|$

x_k	e_k
(a)	
8.0000e-02	8.0000e-02
6.1031e-02	6.1031e-02
3.1055e-02	3.1055e-02
6.7082e-03	6.7082e-03
2.7588e-04	2.7588e-04
4.5241e-07	4.5241e-07
–	–
(b)	
−7.6898e-02	7.6898e-02
5.4080e-02	5.4080e-02
2.3311e-02	2.3311e-02
3.6245e-03	3.6245e-03
7.9337e-05	7.9337e-05
3.7380e-08	3.7380e-08
–	–

In other words, the set $G(x^*)$ might be approximated without involving so many informations about x^*. For such a case, let's consider the following correspondence

$$H(x)(u) = [Dg(x_0)(u)]G(x^*) + [Dg(x)(u) - Dg(x_0)(u)]G(x^*). \tag{64}$$

Since $G(x^*) \supset G(x)$ whenever x is nearby x^*, we can mimic the arguments in Example 2 to conclude that $H(\cdot)$ could produce a reasonable approximation for F around x_0. And therefore, a similar strategy to the one in Example 2 can be applied.

Finally, let us also mention that an other choice for the set-valued approximation H which does not depend on the solution x^* is

$$H(x)(u) = Dg(x)(u)G(x),\ x \in X,\ u \in X.$$

For this approximation H, under some suitable assumptions on the metric regularity of the multifunctions $Df(x)(\cdot) + H(x)(\cdot)$ *uniformly* in x , one can establish the convergence properties of the Newton-type scheme (2) for solving (60) as in Theorems 3 and 4.

The preceding examples illustrate the applicability of our theoretical results. There are many applications in practice that are covered by generalized equations, and that can potentially use our Newton-type scheme proposed in this paper. Many open questions need further investigations. As noticed by one of the referees, it should be interesting to test the current schemes in concrete problems dealing for example with variational inequalities or complementarity problems. This is out of the scope of the current paper and will be the subject of an other research project in the future.

Acknowledgments The authors would like to thank the anonymous referees for their insightful comments, which contribute to the improvement of this manuscript.

References

1. Adly, S., Cibulka, R., Ngai, H.V.: Newton's method for solving inclusions using set-valued approximations. SIAM J. Optim. **25**(1), 159–184 (2015)
2. Adly, S., Ngai, H.V., Nguyen, V.V.: Newton's method for solving generalized equations: Kantorovich's and Smale's approaches. J. Math. Anal. Appl. **439**(1), 396–418 (2016)
3. Aragón Artacho, F.J., Dontchev, A.L., Gaydu, M., Geoffroy, M.H., Veliov, V.M.: Metric regularity of Newton's iteration. SIAM J. Control. Optim. **49**, 339–362 (2011)
4. Azé, D.: A unified theory for metric regularity of multifunctions. Journal of Convex Analysis **13**(2), 225–252 (2006)
5. Dias, S., Smirnov, G.: On the Newton method for set-valued maps. Nonlinear Anal. Theory Methods Appl. **75**(3), 1219–1230 (2012)
6. Dontchev, A.L., Lewis, A.S., Rockafellar, R.T.: The radius of metric regularity. Trans. Am. Math. Soc. **355**, 493–517 (2003)
7. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings: a View from Variational Analysis. Springer Monographs in Mathematics. Springer, New York (2009)
8. Dontchev, A.L., Rockafellar, R.T.: Convergence of inexact Newton methods for generalized equations. Math. Program. **139**(1-2), 115–137 (2013)
9. Gaydu, M., Geoffroy, M.H.: A Newton iteration for differentiable set-valued maps. Journal of Applied Mathematical Analysis and Applications **399**(1), 213–224 (2013)
10. Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of Convex Analysis. Grundlehren Text Editions. Springer, Berlin (2001)
11. Ioffe, A.D.: Metric regularity and subdifferential calculus. Russ. Math. Surv. **55**(3), 501–558 (2000)
12. Ioffe, A.D.: On perturbation stability of metric regularity. Set-Valued Analysis **9**, 101–109 (2001)
13. Izmailov, A.F., Kurennoy, A.S., Solodov, M.V.: The josephy–Newton method for semismooth generalized equations and semismooth SQP for optimization. Set-Valued and Variational Analysis **21**(1), 17–45 (2013)
14. Izmailov, A.F., Solodov, M.V.: Newton-Type Methods for Optimization and Variational Problems. Springer Series in Operations Research and Financial Engineering. Springer International Publishing (2014)
15. Izmailov, A.F., Solodov, M.V.: Newton-type methods: a broader view. J. Optim. Theory Appl. **164**(2), 577–620 (2015)
16. Klatte, D., Kummer, B.: Nonsmooth Equations in Optimization: regularity, Calculus, Methods and Applications, Nonconvex Optimization and Its Applications, vol. 60. Kluwer Academic Publishers (2002)
17. Mordukhovich, B.S.: Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. Trans. Am. Math. Soc. 340(1) (1993)
18. Mordukhovich, B.S., Shao, Y.: Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces. Nonlinear Anal. Theory Methods Appl. **25**(12), 1401–1424 (1995)
19. Ngai, H.V., Théra, M.: Error bounds in metric spaces and application to the perturbation stability of metric regularity. SIAM J. Optim. **19**(1), 1–20 (2008)
20. Pang, C.H.J.: Generalized differentiation with positively homogeneous maps: applications in set-valued analysis and metric regularity. Math. Oper. Res. **36**(3), 377–397 (2011)
21. Robinson, S.M.: Normal maps induced by linear transformations. Math. Oper. Res. **17**(3), 691–714 (1992)
22. Rockafellar, R.T., Wets, R.J.B.: Variational Analysis Grundlehren Der Mathematischen Wissenschaften (A Series of Comprehensive Studies in Mathematics), vol. 317. Springer, Berlin (1998)
23. Xiao, B., Harker, P.T.: A nonsmooth Newton method for variational inequalities, I: theory. Math. Program. **65**, 151–194 (1994)
24. Xiao, B., Harker, P.T.: A nonsmooth Newton method for variational inequalities, II: numerical results. Math. Program. **65**, 195–216 (1994)