Article

Not peer-reviewed version

Hybrid Gauss-Newton Method for Convex Inclusion and Convex-Composite Optimization Problems

Nirjal Shrestha * , Ioannis K Argyros * , Samundra Regmi

Posted Date: 21 November 2025

doi: 10.20944/preprints202511.1628.v1

Keywords: Gauss-Newton method; hybrid method; Banach space; inclusion problem; convex-composite optimization problem; semi-local convergence

Article

Hybrid Gauss-Newton Method for Convex Inclusion and Convex-Composite Optimization Problems

Nirjal Shrestha [1,*], Ioannis K. Argyros [2,*] and Samundra Regmi [3]

[1] Department of Applied Mathematics and Statistics, Colorado School of Mines, Golden, CO, 80401, USA
[2] Department of Computing and Mathematical Sciences, Cameron University, Lawton, OK, 73505, USA
[3] Department of Mathematics, University of Houston, Houston, TX, 77204, USA
* Correspondence: nirjal.shrestha@mines.edu (N.S.); ioannisa@cameron.edu (I.K.A.)

Abstract

In Banach space optimization, solving inclusion and convex-composite problems through iterative methods depends on various convergence conditions. In this work, we develop an extended semi-local convergence analysis for two algorithms used to generate sequences converging to solutions of inclusion and convex-composite optimization problems for Banach space-valued operators. The applicability of these algorithms is extended with benefits: weaker sufficient convergence criteria and tighter error estimates on the distances involved. These advantages are obtained at the same computational cost since the Lipschitz constants used are special cases of the Lipschitz constant used in earlier studies. The implementation issues for these algorithms are also addressed by introducing hybrid algorithms.

Keywords: Gauss-Newton method; hybrid method; Banach space; inclusion problem; convex-composite optimization problem; semi-local convergence

---

1. Introduction

Let X and Y be Banach spaces, and let K represent a closed convex subset of Y. Two inclusion problems have been extensively studied Aragon Artacho et al. (2011); Argyros (2022); Argyros et al. (2024a,b, 2025a); Argyros and Shrestha (2017a); Blum et al. (1998); Burke and Ferris (1995); Chen (2008); Chen and Li (2005); Dedieu and Kim (2002); Dennis and Schnabel (1996); Deuflhard and Heindl (1979); Deuflhard (2004); Dontchev and Rockafellar (2009); Ferreira et al. (2013, 2011a,b); Kantorovich and Akilov (1982); Li and Ng (2007); Li and Wang (2002); Li and Ng (2012); Nocedal and Wright (1999); Proinov (2009); Robinson (1972a,b); Rockafellar (1970, 1967); Smale (1986); Stewart (1969); Wedin (1973). The first problem is defined as

$$F(x) \in K, \tag{1}$$

where $F : X \to Y$ is a differentiable operator in the Fréchet sense. The second problem is a convex composite optimization problem:

$$\min_{x \in X} (h \circ F)(x), \tag{2}$$

where h is a real-valued convex function on Y and F is as defined in (1). If $h(\cdot) = d(\cdot, K)$, the distance function related to K, and (2) is solvable, then (2) reduces to (1). Many optimization and mathematical programming problems can be reduced to solving (1) or (2) Aragon Artacho et al. (2011); Argyros (2022); Argyros et al. (2024a,b, 2025a); Argyros and Shrestha (2017a); Blum et al. (1998); Burke and Ferris (1995); Chen (2008); Chen and Li (2005); Dedieu and Kim (2002); Dennis and Schnabel (1996); Deuflhard and Heindl (1979); Deuflhard (2004); Dontchev and Rockafellar (2009); Ferreira et al. (2013, 2011a,b); Kantorovich and Akilov (1982); Li and Ng (2007); Li and Wang (2002); Li and Ng (2012); Nocedal and Wright (1999); Proinov (2009); Robinson (1972a,b); Rockafellar (1970, 1967); Smale (1986); Stewart (1969); Wedin (1973). Regarding problem (1), Robinson Robinson (1972a,b) developed the

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 November 2025 doi:10.20944/preprints202511.1628.v1

extended Newton method under the conditions that K is a closed convex cone and X is a reflexive space. Let $x_0 \in X$.

Algorithm $A(x_0)$:

Given $x_n, n = 0, 1, 2, \ldots$, compute x_{n+1} as follows: First, define the set

$$D_\infty := \{d \in X : F(x) + F^{'}(x)d \in K \text{ for all } x \in X\}. \tag{3}$$

If $D_\infty(x) \neq \emptyset$, pick $d_n \in D_\infty(x_n)$ to satisfy $\|d_n\| = \min_{d \in D_\infty(x_n)} \|d\|$, and set $x_{n+1} = x_n + d_n$. In the case when $D_\infty(x) = \emptyset$ for some $x \in X$, the algorithm is not well-defined. It is said to be well-defined if at least one sequence is generated by this algorithm. Let T_{x_0} be a convex process given by

$$T_{x_0} d = F^{'}(x_0)d \text{ for all } d \in X. \tag{4}$$

Robinson in Robinson (1972a,b) provided two conditions:

$$\text{Range}(T_{x_0}) = Y, \tag{5}$$

and $F^{'}$ is Lipschitz continuous with modulus H, T_{x_0} is normed with $\|T_{x_0}^{-1}\| < \infty$, and

$$\|x_1 - x_0\| \leq \frac{1}{2H\|T_{x_0}^{-1}\|}. \tag{6}$$

Under these conditions, Algorithm $A(x_0)$ generates a sequence $\{x_n\}$ converging to some x^* solving (1). Numerous authors have presented relevant convergence results lacking the affine invariant property for the operator F Deuflhard (2004). Later, in the elegant work by Li and Ng in Li and Ng (2007), convergence conditions addressing this problem under weaker conditions were introduced. Their study also relied on the new concept of the weak-Robinson condition for convex processes introduced in Section 2, Subsection 2.1. The semi-local convergence of the two algorithms is presented in Section 2, Subsection 2.2. The hybrid methods are studied in Section 2, Subsection 2.3. The concluding remarks appear in Section 3 and the paper ends with discussion in Section 4.

2. Materials and Methods

2.1. Weak-Robinson Condition and Properties of Convex Processes

Let $U(x, \rho)$ denote the open ball in X with center $x \in X$ and radius $\rho > 0$. Also, let $d(x, S)$ denote the distance from x to S. We assume familiarity with the concept of convex process $T : X \to 2^Y$ as defined by Rockafellar Rockafellar (1970, 1967), as well as its properties. We list only some in order to make the study as self-contained as possible. The following sets are needed:

$$D(T) = \{x \in X : T_x \neq \emptyset\},$$
$$R(T) = \bigcup\{T_x : x \in D(T)\},$$

and

$$T^{-1}y = \{x \in X : y \in T_x\} \quad \text{for all } y \in Y.$$

Proposition 1 (Rockafellar (1967))**.** *Let B_1, B_2 be convex processes with $D(B_1) = D(B_2) = X$ and $R(B_1) = Y$. Suppose that $\|B_1^{-1}\| \, \|B_2\| < 1$ and that $(B_1 + B_2)(x)$ is closed for all $x \in X$. Then, the following assertions hold:*

$$R(B_1 + B_2) = Y$$

and

$$\|(B_1 + B_2)^{-1}\| \leq \frac{\|B_1^{-1}\|}{1 - \|B_1^{-1}\| \, \|B_2\|}.$$

Next, we present Lipschitz-type conditions connected to convex processes for comparison.

Definition 1. *Let $E : X \to \mathcal{L}(X,Y)$, where $\mathcal{L}(X,Y)$ denotes the space of linear operators from X to Y. Let $T : Y \to 2^Z$ be a convex process, where Z is also a Banach space. The pair (T,E) is said to be center-Lipschitz continuous on $U(x_0,\rho_0)$ $(\rho_0 > 0)$ if*

$$\|T\big(E(x) - E(x_0)\big)\| \le L_0\|x - x_0\|,$$

for all $x \in U(x_0,\rho_0)$ and some $L_0 > 0$. Set $\rho = \min\{\rho_0, \frac{1}{L_0}\}$.

Definition 2. *Let E and T be as in Definition 1. Then, the pair (T,E) is said to be restricted-Lipschitz continuous on $U(x_0,\rho)$ if*

$$\|T\big(E(x) - E(y)\big)\| \le L\|x - y\|$$

for all $x,y \in U(x_0,\rho)$ and some $L > 0$.

Definition 3. *Let E and T be as in Definition 1. Then, the pair (T,E) is said to be Lipschitz continuous on $U(x_0,\rho_0)$ if*

$$\|T\big(E(x) - E(y)\big)\| \le L_1\|x - y\|,$$

for all $x,y \in U(x_0,\rho_0)$ and some $L_1 > 0$.

Remark 1. *It follows from these Definitions, since $\rho \le \rho_0$, that*

$$L_0 \le L_1 \tag{7}$$

and

$$L \le L_1. \tag{8}$$

The convergence analysis in Argyros et al. (2024b) used L_1. However, the proofs can be repeated with L replacing L_1. Note that $L_0 = L_0\big(U(x_0,\rho_0)\big)$ and $L_1 = L_1\big(U(x_0,\rho_0)\big)$, while $L = L\big(L_0, U(x_0,\rho_0)\big)$. We assume that

$$L_0 \le L. \tag{9}$$

Otherwise, i.e., if $L \le L_0$, the results hold with L_0 replacing L.

Definition 4. *The problem (1) is said to satisfy the Weak-Robinson condition at x_0 on $U(x_0,\rho_0)$ if*

$$-F(x_0) \in R(T_{x_0}) \text{ and } R\big(F'(x)\big) \subseteq R(T_{x_0}) \quad \text{for all} \quad x \in U(x_0,\rho_0). \tag{10}$$

It follows that for the problem (1), the Robinson condition at x_0 implies the Weak-Robinson condition at x_0 on X.

Proposition 2. *Let $x_0 \in X$ and $r \in \left(0, \frac{1}{L_0}\right]$. Suppose that the problem (1) satisfies the Weak-Robinson condition at $x_0 \in U(x_0,r)$ and $(T_{x_0}^{-1}, F')$ is L_0-Lipschitz continuous on $U(x_0,r)$. Then, for all $x \in U(x_0,r)$, the following assertions hold:*

$$R(T_{x_0}) = R(T_x), \quad D(T_x^{-1}F'(x_0)) = X \tag{11}$$

and

$$\|T_x^{-1}F'(x_0)\| \le \frac{1}{1 - L_0\|x - x_0\|}. \tag{12}$$

Additionally, if X is reflexive, then

$$D_\infty(x) \neq \emptyset. \tag{13}$$

Proof. Simply replace L_1 by the smallest L_0 that is actually needed in the proof of Proposition 2.3 in Argyros et al. (2024b). □

2.2. The Gauss-Newton Method for the Problem

We assume in this section that X is reflexive, h is a continuous convex function, and the set K is the set of minimum points defined as

$$K := \arg\min h. \tag{14}$$

For each $\Delta \in (0, +\infty]$ and $x \in X$, let $D_\Delta(x)$ be the set of all $d \in X$ such that $\|d\| \le \Delta$ and

$$h\big(F(x) + F'(x)d\big) = \min\Big\{h\big(F(x) + F'(x)d_0\big) : d_0 \in X, \|d_0\| \le \Delta\Big\}. \tag{15}$$

Then, $d \in D_\Delta(x)$ if and only if d solves the convex minimization problem:

$$\min\Big\{h\big(F(x) + F'(x)d_0\big) : d_0 \in X, \|d_0\| \le \Delta\Big\}. \tag{16}$$

Define

$$D_\Delta(x) = \Big\{d \in X : \|d\| \le \Delta,\ F(x) + F'(x)d \in K\Big\}. \tag{17}$$

Let $r \in [1, +\infty)$, $x_0 \in X$, and $\Delta \in (0, +\infty]$. Then, the Gauss-Newton method for solving the problem (2) is given by the Algorithm:

Algorithm $A(r, \Delta, x_0)$:

Given $x_n, n = 0, 1, \ldots$, compute x_{n+1} as follows: If

$$h\big(F(x_n)\big) = \min\Big\{h\big(F(x_n) + F'(x_n)d\big) : d \in X, \|d\| \le \Delta\Big\}, \tag{18}$$

then stop. Otherwise, select $d_n \in D_\Delta(x_n)$ so that $\|d_n\| \le r\, d\big(0, D_\Delta(x_n)\big)$, and let $x_{n+1} = x_n + d_n$.

Define the parameters β and α by

$$\beta = r\|T_{x_0}^{-1}\big(-F(x_0)\big)\|, \quad \alpha = \frac{r}{1 + (r-1)L\beta}, \quad \text{and} \quad \alpha_1 = \frac{r}{1 + (r-1)L_1\beta}. \tag{19}$$

The majorizing functions for all $t \ge 0$ are given by

$$\begin{aligned} g_0(t) &= \beta - t + \frac{\alpha L_0}{2}t^2, \\ g(t) &= \beta - t + \frac{\alpha L}{2}t^2, \\ \text{and} & \\ g_1(t) &= \beta - t + \frac{\alpha_1 L_1}{2}t^2. \end{aligned}$$

It follows from (19), (7), and these definitions that

$$g_0(t) \le g_1(t) \tag{20}$$

and

$$g(t) \le g_1(t). \tag{21}$$

Define the majorizing sequences $\{s_n\}, \{t_n\}$ by

$$s_0 = 0, \quad s_1 = r, \quad s_{n+1} = s_n - \frac{g(s_n)}{g_0'(s_n)},$$

$$t_0 = 0, \quad t_1 = r, \quad t_{n+1} = t_n - \frac{g(t_n)}{g'(t_n)},$$

and

$$u_0 = 0, \quad u_1 = r, \quad u_{n+1} = u_n - \frac{g_1(u_n)}{g_1'(u_n)}.$$

The sequence $\{u_n\}$ is used in Argyros et al. (2024b). Here, we use $\{t_n\}$ although $\{s_n\}$ can also be used according to the proof of Theorem 3.1. Next, a Kantorovich condition is provided for the convergence of the sequence $\{t_n\}$ also for the sequence $\{s_n\}$.

Lemma 1. *Suppose*

$$\beta(r+1)L \leq 1. \tag{22}$$

Then, the sequence $\{t_n\}$ is increasingly convergent to t^* which is the smallest of the two roots of the function g guaranteed to exist by (22). Moreover for

$$\theta = \frac{1 - L\beta - \sqrt{1 - 2L\beta - (r^2-1)(L\beta)^2}}{\beta r L} < 1, \tag{23}$$

$$t_n = \frac{\sum_{j=0}^{2^n-2} \theta^j}{\sum_{j=0}^{2^n-1} \theta^j} t^*, n = 1, 2, \ldots. \tag{24}$$

Proof. Simply replace L_1, g_1 by L, g in the proof of Lemma 3.1 in Argyros et al. (2024b) . □

Remark 2. *The Kantorovich condition in Argyros et al. (2024b) is*

$$\beta(r+1)L_1 \leq 1. \tag{25}$$

It follows that

$$(25) \implies (22) \tag{26}$$

but not necessarily vice versa unless $L_1 = L$. Moreover,

$$\theta \leq \theta_1, \tag{27}$$

where θ_1 is given by (23) with L replaced by L_1. Thus, the new results are at least as good as those in Argyros et al. (2024b).

It is also worth noting that the improved results are obtained under the same computational effort, since in practice the computation of L_1 requires that of L_0 and L as special cases. Next, we present the semi-local convergence result for the algorithm $A(r, \Delta, x_0)$.

Theorem 1. *Suppose that the problem (1) satisfies the Weak-Robinson condition at x_0 on $U(x_0, t^*)$ and that $(T_{x_0}^{-1}, F^{'})$ is L_0-center-Lipschitz continuous on $U(x_0, \frac{1}{L_0})$ and L-restricted-Lipschitz on $U(x_0, t^*)$. Suppose that*

$$\beta \leq \min\left\{\Delta, \frac{1}{L(r+1)}\right\}. \tag{28}$$

Then, any sequence $\{x_n\}$ generated by the algorithm $A(r, \Delta, x_0)$ is convergent to some x^* solving (1). Moreover, the following assertions hold:

$$\|x_n - x_{n-1}\| \leq t_n - t_{n-1}, \quad n = 1, 2, \ldots, \tag{29}$$

$$F(x_{n-1}) + F^{'}(x_{n-1})(x_n - x_{n-1}) \in K, \quad n = 1, 2, \ldots$$

and

$$\|x^* - x_n\| \leq \frac{\theta^{2^n-1}}{\sum_{j=0}^{2^n-1}\theta^j}t^*. \tag{30}$$

Proof. Simply replace θ_1, L_1, g_1, u_n by θ, L, g, t_n, respectively, in the proof of Theorem 3.1 in Argyros et al. (2024b). □

Theorem 2. *Suppose that the problem (1) satisfies the Weak-Robinson condition at x_0 on $U(x_0, t^*)$, $L_0, L \in (0, +\infty)$, $\beta \geq \|T_{x_0}^{-1}(-F(x_0))\|$, $r = 1$, and $\left(T_{x_0}^{-1}, F'\right)$ is L_0-center-Lipschitz on $U(x_0, \frac{1}{L_0})$, L-Lipschitz continuous on $U(x_0, \rho)$, and $2\beta L \leq 1$. Then, any sequence $\{x_n\}$ generated by the Algorithm $A(x_0)$ converges to a solution t^* of the problem (1). Moreover, the following error assertion holds:*

$$\|x^* - x_n\| \leq \frac{p^{2^n-1}}{\sum_{j=0}^{2^n-1}p^j}t^*,$$

where

$$p = \frac{1 - L\beta - \sqrt{1 - 2\beta L}}{L\beta}.$$

Proof. Replace q_1, L_1, g_1, u_n by p, L, g, t_n in the proof of Theorem 3.2 in Argyros et al. (2024b). □

Remark 3. *(i)* *Theorem 2 can be reduced to a weaker version of Robinson's Robinson (1972a,b) for $\beta = \|T_{x_0}^{-1}(-F(x_0))\|$, $\rho_1 = \frac{1-\sqrt{1-2H\|T_{x_0}^{-1}\|\beta}}{H\|T_{x_0}^{-1}\|}$, and $\|x_1 - x_0\| \leq \frac{1}{2H\|T_{x_0}^{-1}\|}$.*

(ii) *The Kantorovich conditions for the convergence of the sequence $\{t_n\}$ can be weakened further.*

Indeed, let us demonstrate this in the case of Algorithm $A(r, x_0, \Delta)$ and Theorem 3.4. According to the proof of Theorem 3.5:

$$r_0 = 0, \quad r_1 = \beta, \quad r_2 = \beta + \frac{L_0\beta^2}{2(1 - L_0\beta)}, \tag{31}$$

$$r_{n+2} = r_{n+1} + \frac{L(r_{n+1} - r_n)^2}{2(1 - L_0 r_{n+1})}.$$

The sequence majorizing $\{x_n\}$ in Argyros et al. (2024b), i.e., $\{u_n\}$, can be written as:

$$u_0 = 0, \quad u_1 = \beta, \quad u_{n+1} = u_n + \frac{L_1(u_n - u_{n-1})^2}{2(1 - L_1 u_n)} = u_n - \frac{g_1(u_n)}{g_1'(u_n)}. \tag{32}$$

Moreover, the convergence conditions are given in Argyros (2022) and Argyros et al. (2024b), respectively, as:

$$\lambda = \bar{L}\beta \leq 1 \tag{33}$$

and

$$\mu = 2L_1\beta r \leq 1, \tag{34}$$

where $\bar{L} = \frac{1}{4}\left(4L_0 + \sqrt{L_0 L + 8L_0^2} + \sqrt{L_0 L}\right)$. Notice that:

$$\mu \leq 1 \implies \lambda \leq 1, \tag{35}$$

holds and $\frac{\lambda}{\mu} \to 0$ as $\frac{L_0}{L_1}, \frac{L}{L_1}$ approaches 0. Thus, the new convergence condition is infinitely weaker than the original Kantorovich condition (33). Moreover, the following hold:

$$0 \leq r_{n+1} - r_n \leq u_{n+1} - u_n$$

and

$$r^* = \lim_{n\to+\infty} r_n \le t^*.$$

Hence, the error bounds on the distances $\|x_{n+1} - x_n\|$ are at least as tight and the information on the location of the solution is at least as precise. The results can also be rewritten using majorant functions replacing the Lipschitz constants L_0, L, and L_1 (see Argyros (2022); Argyros et al. (2024a,b, 2025a)).

2.3. The Implementation of the Algorithms

The iterates in both algorithms are computed by finding the expensive $F'(x_n)^{-1}$. In the papers Argyros et al. (2024a,b), we developed a hybrid method to address this difficulty. Specifically, we define the method

$$F(x_n) + F_1(x_n)(x_{n+1} - x_n) = 0, \tag{36}$$

where $F_1 = MB_k^{-1}(x_n)$, k is a natural number, $M \in \mathcal{L}(X, Y)$ is an invertible operator, $\Delta = M^{-1}(M - F'(x))$, and $B_k = I + \Delta + \cdots + \Delta^k$. Here, $B_\infty M^{-1} = \lim_{k\to+\infty} B_k M^{-1} = F'$.

Similarly, we introduce hybrid methods to replace both Algorithms. We demonstrate this with Algorithm $A(x_0)$. The method for $A(r, \Delta, x_0)$ follows along the same lines. We take $M = F'(x_0)$ for simplicity. As in Argyros et al. (2024a), define the parameters:

$$a = a(\|x - x_0\|) \ge \|A\|,$$

$$\bar{b_k} = \frac{a(1-a^k)}{1-a}, \quad b_k = \frac{1}{1-\bar{b_k}},$$

$$b_\infty = \frac{1-a}{1-2a}, \quad \lambda(\|x - x_0\|) = \lambda_k(\|x - x_0\|) = b_k b_\infty \frac{a^{k+1}}{1-a}.$$

Define the sequence $\{v_n\}$ with $v_0 = 0$, $v_1 = \beta$,

$$\begin{aligned} v_2 &= \beta + \frac{\frac{L_0}{2}\beta^2 + \lambda(\beta)}{1 - L_0\beta}, \\ v_{n+2} &= v_{n+1} + \frac{\frac{L}{2}(v_{n+1} - v_n)^2 + \lambda_k(v_{n+1})}{1 - L_0 v_{n+1}}. \end{aligned} \tag{37}$$

The sequence $\{v_n\}$ shall be shown to majorizing for the sequence $\{x_n\}$ generated by the Algorithm defined next.

Algorithm $A(x_0)'$:

Given $x_n, n = 0, 1, \ldots,$ compute x_{n+1} as follows: First define the set

$$D'_\infty(x) := \left\{ d \in X : F(x) + F_1(x)d \in K \right\} \quad \text{for all } x \in X. \tag{38}$$

If $D'_\infty(x) \ne \emptyset$, pick $d_n \in D_\infty(x_n)$ to satisfy $\|d_n\| = \min_{d\in D_\infty(x_n)} \|d\|$ and set $x_{n+1} = x_n + d_n$.

Let us provide a convergence criterion for the sequence $\{v_n\}$. Suppose that there exists $0 \le \rho < \frac{1}{L_0}$ such that for each $n = 0, 1, 2, \ldots,$

$$v_n \le \rho. \tag{39}$$

It follows by (39) and (38) that $0 \le v_n \le v_{n+1} \le \rho$, and there exists $v^* \in [0, \rho]$ such that $v^* = \lim_{n\to\infty} v_n$. The limit point v^* is the unique least upper bound of the sequence $\{v_n\}$. We can show the analog of Theorem 2 for the sequence $\{x_n\}$ generated by Algorithm $A(x_0)'$.

 doi:10.20944/preprints202511.1628.v1

Theorem 3. *Suppose that the problem (1) satisfies the weak-Robinson condition at x_0 on $U(x_0, v^*)$, $L_0, L \in (0, +\infty)$, $\beta \geq \|T_{x_0}^{-1}(-F(x_0))\|$, $r = 1$, and $\left(T_{x_0}^{-1}, F'\right)$ is L_0-center Lipschitz on $U(x_0, \frac{1}{L_0})$ and L-Lipschitz on $U(x_0, \rho)$,*

$$a < \frac{1}{2}, \tag{40}$$

and (39) holds. Then, any sequence $\{x_n\}$ generated by Algorithm $A(x_0)'$ converges to a solution x^* of the problem (1). Moreover, the following error estimate holds:

$$\|x^* - x_n\| \leq v^* - v_n. \tag{41}$$

Proof. Notice that for each $x \in U(x_0, v^*)$,

$$\begin{aligned} \|B_k - I\| &= \|\Delta + \Delta^2 + \cdots + \Delta^k\| \\ &\leq \|\Delta\| + \|\Delta\|^2 + \cdots + \|\Delta\|^k \\ &\leq a + a^2 + \cdots + a^k = \frac{a(1-a^k)}{1-a} = \bar{b_k} < 1 \quad (\text{by } (40)). \end{aligned}$$

So, $B_k^{-1} \in \mathcal{L}(Y, X)$ and $\|B_k^{-1}\| \leq \frac{1}{1-\bar{b_k}} = b_k$. Then, we can write

$$\begin{aligned} F_1(x) &= F'(x_0)B_k^{-1}(x) \\ &= F'(x_0)\left(B_k^{-1}(x) - B_\infty^{-1}(x) + B_\infty^{-1}(x)\right) \\ &= F'(x_0)B_\infty^{-1}(x) + F'(x_0)\left(B_k^{-1}(x) - B_\infty^{-1}(x)\right) \\ &= F'(x_0)B_\infty^{-1}(x) + F'(x_0)B_k^{-1}(x)\left(B_\infty(x) - B_k(x)\right)B_\infty^{-1}(x). \end{aligned} \tag{42}$$

We also need the estimate

$$\begin{aligned} \|B_k^{-1}(x)\left(B_\infty(x) - B_k(x)\right)B_\infty^{-1}(x)\| &\leq b_k(a^{k+1} + \cdot)b_\infty \\ &= b_k b_\infty \frac{a^{k+1}}{1-a} \\ &= \lambda_k(\|x - x_0\|). \end{aligned} \tag{43}$$

By replacing F' by F_1 in the proof of Theorem 2 and using (43), we get

$$\|x_{n+1} - x_n\| \leq v_{n+1} - v_n, \tag{44}$$

leading to the existence of $x^* \in U(x_0, v^*)$ solving (1). Then, from (44), we have for each $i = 0, 1, 2, \ldots$

$$\|x_{n+i} - x_n\| \leq v_{n+i} - v_n. \tag{45}$$

Finally, by letting $i \to +\infty$ in (45), we obtain (41). □

Remark 4. *(i) It is worth noticing that the condition (39) is weaker than the Kantorovich condition $2\beta L \leq 1$ used in the Theorem 2 or (33) since both of these conditions imply (39) but not necessarily vice versa. Thus, (39) can replace these stronger conditions in the Theorem 2 provided that the sequence $\{v_n\}$ (for $k = \infty$) replaces the sequence $\{t_n\}$.*

(ii) The results of the Theorem 3 can be presented in a more general setting if M is not necessarily chosen to be $F'(x_0)$ as follows: Simply replace $E(x_0)$ by M in the Definition 1 and $F'(x_0)$ by M in (12) where $M \in \mathcal{L}(X, Y)$ is an invertible operator. Then, the conclusions of the Theorem 3 hold this more general setting.

(iii) The results can be further extended if the Lipschitz continuity is replaced by the generalized continuity of the operator F' along the lines of our work in Argyros (2022); Argyros et al. (2024a,b, 2025a).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 November 2025 doi:10.20944/preprints202511.1628.v1

3. Concluding Remarks

Remark 5. *The results in this work can also be expressed using majorant functions instead of the Lipschitz constants L_0, L, and L_1. This approach allows for even tighter error estimates (see Aragon Artacho et al. (2011); Argyros (2022); Argyros et al. (2024a,b, 2025a)).*

For instance, the constant L_0 in Definition 1 can be replaced with a function $\varphi_0 : [0, +\infty) \to [0, +\infty)$ that is continuous and non-decreasing. Consequently, for all $x \in U(x_0, \rho_0)$, we obtain:

$$\|T(E(x) - E(x_0))\| \leq \varphi_0(\|x - x_0\|). \tag{46}$$

If $\varphi_0(t) = L_0(t)$, (46) reduces to the one stated in Definition 1. Similarly, L_1 and L in Definition 2 and Definition 3 can be replaced accordingly.

Let us introduce some parameters and real functions that play a role in the semi-local convergence of the Gauss-Newton algorithm under generalized continuity conditions used to control the derivative f' of the operator f. Set $A = [0, +\infty)$.
Suppose: (C_1) There exists a function $\psi_0 : A \to A$, which is continuous and nondecreasing and a parameter $\alpha > 0$ such that the equation $\alpha\psi_0(t) - 1 = 0$ has at least one positive solution.Denote the smallest such solution by ξ. Set $A_0 = [0, \xi)$.
(C_2) There exists a function $\psi : A_0 \to A$ which is continuous and nondecreasing. Let $\beta \geq 0$ be a parameter. Define the scalar sequence $\{s_n\}$ for $c > 0$, $s_0 = 0$, $s_1 = \beta$,

$$s_2 = s_1 + \frac{c\int_0^1 \psi_0((1-\theta)(s_1 - s_0))d\theta(s_1 - s_0)}{1 - \psi_0(s_1)} \tag{47}$$

and each $n = 1, 2, \ldots$ by

$$s_{n+2} = s_{n+1} + \frac{c\int_0^1 \psi((1-\theta)(s_{n+1} - s_n))d\theta(s_{n+1} - s_n)}{1 - \psi_0(s_{n+1})}.$$

The sequence $\{s_n\}$ shall be shown to be majorizing in Theorem 6.1. But, first a convergence condition is provided for the sequence $\{s_n\}$.
(C_3) There exists $\xi_0 \in [0, \xi)$ such that for all $n = 0, 1, 2, \ldots$

$$\psi_0(s_n) < 1 \text{ and } s_n \leq \xi_0.$$

It follows by this condition and the definition of the sequence $\{s_n\}$ that for each $n = 0, 1, 2, \ldots$, $0 \leq s_n \leq s_{n+1} \leq \xi_0$ and there exists s^* such that $s^* = \lim_{n\to\infty} s_n$. It is known from calculus that s^* is the least upper bound of the sequence $\{s_n\}$ which is unique.

The scalar functions ψ_0 and ψ connect to the operators on the Gauss-Newton Algorithm. Let $c \in [1, +\infty), \Delta \in (0, +\infty)$ and $x_0 \in X$ be such that $-f(x_0) \in \text{Range}(T(x_0))$. Define β by

$$\beta = c\|T(x_0)^{-1}(-f(x_0))\| \tag{48}$$

(C_4) Let $T : Y \to 2^Z$ be a convex process, where Z is also a Banach space. The pair $(T^{-1}f')$ is center-generalized continuous on $U[x_0, \rho_0)$ for $\rho > 0$, i.e., $\|T^{-1}(f'(x) - f'(x_0))\| \leq \psi_0(\|x - x_0\|)$ for all $x \in U(x_0, \rho_0)$.
Set

$$\bar{\xi}_0 = \min\{\rho_0, \xi\}. \tag{49}$$

(C_5) Let f', T be as in the condition (C_4). The pair (T^{-1}, f') is restricted generalized continuous on $U(x_0, \bar{\xi})$, i.e., $\|T^{-1}(f'(x) - f'(y))\| \leq \psi(\|x - y\|)$ for all $x, y \in U(x_0, \bar{\xi})$.
$(C_6)\beta \leq \Delta$.
$(C_7)U[x_0, s^*] \subset U[x_0, \rho_0]$.

 doi:10.20944/preprints202511.1628.v1

Next, the main semi-local convergence of the Gauss-Newton algorithm is shown under the conditions $(C_1)-(C_7)$.

Theorem 4. *Suppose that the inclusion (1) satisfies the weak-Robinson condition at x_0 on $U(x_0,s^*)$ and (C_1)-(C_7) hold. Then, the Gauss-Newton Algorithm $A(c,\Delta,x_0)$ is well defined and any sequence $\{x_n\}$ converges to some $x^* \in U[x_0,s^*]$ such that $F(x^*) \in K$. Moreover, the following items hold for all $n = 1,2,\ldots$*

$$\|x_n - x_{n-1}\| \leq s_n - s_{n-1}, \tag{50}$$

$$f(x_{n-1}) + f'(x_{n-1})(x_n - x_{n-1}) \in K \tag{51}$$

and

$$\|x^* - x_n\| \leq s^* - s_n. \tag{52}$$

Proof. It follows by the weak-Robinson condition that

$$D_\infty(x_0) = T^{-1}(x_0)(-f(x_0)) \neq 0. \tag{53}$$

So, by the definition of β, (47), (48) and the condition (C_6), we get

$$cd(0,D_\infty(x_0)) = c\left\|T(x_0)^{-1}(-F(x_0))\right\| = \beta \leq s_1 - s_0 \leq \Delta \tag{54}$$

By the reflexivity of X, (54) and since $c \geq 1$,there exists $d \in D_\infty(x_0)$ such that $\|d\| \leq \beta \leq \Delta$. Hence, $d(0,D_\Delta(x_0)) = d(0,D_\infty(x_0))$ and $D_\Delta(x_0) \neq \emptyset$. Thus, the iterate x_1 is well defined and $f(x_0) + f'(x_0)d_0 \in K$; so, (51) holds for $n = 1$.

Moreover, by (54) and the Gauss-Newton Algorithm $A(c,\Delta,x_0) : \|d_0\| \leq cd(0,D_\infty(x_0)) \leq s_1 - s_0$, i.e., $\|x_1 - x_0\| \leq s_1, -s_0$. So,(50) holds sot $n = 1$. Suppose that (50) and (51) hold for $n = 1,2,\ldots,m$. Then, we get

$$\begin{aligned}\|x_m - x_0\| &\leq \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \cdots + \|x_1 - x_0\| \\ &\leq (s_m - s_{m-1}) + (s_{m-1} - s_{m-2}) + \cdots + (s_1 - s_0) = s_m.\end{aligned}$$

Similarly, we have $\|x_{m-1} - x_0\| \leq s_{m-1} \leq s_m$.

Then, by the condition (C_7) $\theta x_m + (1-\theta)x_{m-1} \in U(x_0,s^*) \subset B(x_0,\xi)$ for all $\theta \in [0,1]$, so $D_\infty(x_m) \neq 0$. Hence, the iterate x_{m+1} exists and (51) holds for $n = m+1$. Furthermore, for $\theta = 1$: $\|x_m - x_0\| \leq s^* \leq \xi$. By applying the condition (C_4) $\|T(x_0)^{-1}(f'(x_m) - f'(x_0))\| \leq \psi_0(\|x_m - x_0\|) < \psi_0(s^*) \leq 1$. Then, by the weak-Robinson condition, the Proposition (2) is applicable to x_m replacing x leading to

$$D\Big(T(x_m)^{-1}f'(x_0)\Big) = X \tag{55}$$

and

$$\left\|T^{-1}(x_0)f'(x_0)\right\| \leq \frac{1}{1-\psi_0(\|x_m - x_0\|)} \leq \frac{1}{1-\psi_0(s_m)} \tag{56}$$

Thus, we get

$$T(x_0)^{-1}\int_0^1 \Big((\theta x_m + (1-\theta)x_{m-1}) - f'(x_{m-1})) - x_m\Big)d\theta(x_{m-1} - x_m) \neq 0, \tag{57}$$

and by the induction hypothesis, (47) and the condition (C_s)

$$\begin{aligned}&\|T(x_0)^{-1}[-f(x_m) + f(x_{m-1}) - f'(x_{m-1})(x_{m-1} - x_m)]\| \\ &\leq \int_0^1 \psi((1-\theta)(s_m - s_{m-1}))d\theta(s_{m-1} - s_{m-1}).\end{aligned}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 November 2025 doi:10.20944/preprints202511.1628.v1

Notice that this estimate holds for ψ replaced by ψ_0 if $m=1$. Next, we must show

$$\emptyset \neq \left(T^{-1}(x_m)^{-1}f'(x_0)T(x_0)^{-1}[-f(x_m)+f(x_{m-1})(x_{m-1})(x_m)\right] \subseteq D_\infty(x_m) \tag{58}$$

It follows from (55) and (56) that the set in the middle of (58) is non-empty. To show the rest in (58), set $w=-f(x_m)+f(x_{m-1})+f'(x_{m-1})(x_m-x_{m-1})$ and $d\in\left(T^{-1}(x_m)f'(x_0)\right)T(x_0)^{-1}(w)$, so $d\in\left(T(x_m)^{-1}g'(x_0)\right)v$ for some $v\in T(x_0)^{-1}(w)$. We must show $d\in D_\infty(x_m)$. But we have $f'(x_m)d\in f'(x_0)v+K$ and $f'(x_0)v\in w+K$.

So, $f'(x_m)d\in w+K+K=w+K$, since K is a cone. By the definition of w and the induction hypothesis, it follows that

$$f(x_m)+f'(x_m)d\in f(x_{m-1})+f'(x_{m-1})(x_m-x_{m-1})+K\subset K+K=K,$$

Thus, $d\in D_\infty(x_m)$, showing the right-hand side of the assertion (58). Hence,we can get from (56) and (58) in turn,

$$\begin{aligned} cd(0,D_\infty(x_m)) &\leq \|T(x_m)^{-1}F'(x_0)\|\|T(x_0)^{-1}\left(-F(x_m)+F(x_{m-1})+F'(x_{m-1})(x_m-x_{m-1})\right\| \\ &\leq c\frac{\int_0^1\psi((1-\theta)(s_m-s_{m+1}))d\theta(s_m-s_{m-1})}{1-\psi_0(s_m)} \\ &= s_{m+1}-s_m. \end{aligned} \tag{59}$$

Notice that the sequence $s_{m+1}-s_m$ is decreasing, so $s_{m+1}-s_m\leq s_1-s_0=\beta\leq\Delta$. Thus, by (59)$d(0,D_\infty(x_m))\leq cd(0,D_\infty(x_m))\leq\Delta$, which together with $D_\infty(x_m)\neq 0$ imply that there exists $d_0\in X$ with $\|d_0\|\leq\Delta$ such that $F(x_m)+F'(x_m)d_0\in K$. Hence, we have $D_\Delta(x_m)\neq\emptyset$ and $d(0,D_\Delta(x_m))=d(0,D_\infty(x_m))$. Thus, it follows by (59) that $\|x_{m+1}-x_m\|=\|d_m\|\leq cd(0,D_\Delta(x_m))\leq s_{m+1}-s_m$. Thus, the induction for (50) and (51) is completed. It follows that the sequence $\{x_m\}$ is fundamental in X (since s_m is also fundamental as convergent). Therefore,there exists $x^*\in X$ such that $\lim_{m\to+\infty}x_m=x^*$ and $F(x^*)\in K$. Finally,from (50) for $i=0,1,2\ldots$ and the triangle inequality, we obtain

$$\|x_{m+i}-x_m\|\leq s_{m+i}-s_m$$

leading to (52) if $i\to+\infty$ $\square$

Remark 6. *(i)* *Let us specialize the functions ψ_0 and ψ to be $\psi_0(t)=L_0t$ and $\psi(t)=Lt$. Then, the condition (C_3) becomes $(C_3)'$: $L_0s_n<1$ and $s_n\leq\frac{1}{L_0}$, since $\xi_0=\frac{1}{L_0}$ by the condition (C_1). Notice that the sequence $\{s_n\}$ becomes by the definition of the functions ψ_0 and ψ*

$$\begin{aligned} s_0=0,s_1=\beta,s_2&=s_1+\frac{cL_0(s_1-s_0)^2}{2(1-s_1L_0)} \\ s_{m+1}&=s_m+\frac{2L(s_m-s_{m-1})^2}{2(1-L_0s_m)} \end{aligned} \tag{60}$$

According to the proof of Theorem 6.1, the sequence $\{s_n\}$ given by (60) majorizes $\{x_n\}$ provided that the conditions (C_3) and (C_6) hold. These conditions are weaker than (1). Indeed, this is the case, since

$$1\leq c\leq\alpha(1+(c-1)L\beta)\leq\alpha(1+(c-1)Ls_n),$$

so

$$\frac{c}{\alpha}(1-Ls_n)^{-1}\leq(1-\alpha Ls_n)^{-1}\leq-\psi'(s_n)^{-1}.$$

Hence, the sequence $\{s_n\}$ can be replaced by the less tight $\{t_n\}$ given by the formula (18). Notice that the sequence $\{t_n\}$ converges provided that the $\beta \leq \frac{1}{(c+1)L}$ (see (27)). Consequently, the sequence $\{s_n\}$ given by (60) is tighter than $\{t_n\}$ used in Ferreira et al. (2011b) and the sufficient convergence conditions (C_3) wood (C_6) are weaker than $\xi \leq \frac{1}{(C+1)L_1}$ used in Ferreira et al. (2011b). We conclude that under these choices of the functions ψ_0 and ψ, the results of Theorem 6.1 specialize to the ones in Theorem 1. Clearly, the same is true for Theorem 2 if we take $c = 1$ in Theorem 6.1.

(ii) *It is well known that generalized continuity provides even sharper bounds on the derivative than Lipschitz or Hölder or other conditions.*

4. Discussion

The applicability of two Gauss-Newton methods for solving inclusion and convex-composition optimization problems for Banach space-valued operators is extended using more precise majorizing sequences and under the same computational cost, providing weaker yet sufficient semi-local convergence criteria and more precise error estimates on the distances $\|x_{n+1} - x_n\|$ and $\|x^* - x_n\|$. Moreover, the implementation of these algorithms is addressed by developing the corresponding hybrid Gauss-Newton methods. The Lipschitz constants can be replaced by the generalized continuity of F'.

In future research, the ideas presented in this study shall be used to extend the applicability of other iterative methods used to solve these problems.

Author Contributions: All authors contributed equally.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: Not applicable

Acknowledgments: IWe would like to express our gratitude to the anonymous reviewers for the constructive criticism of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

Aragon Artacho, F.J.; Dontchev, A.L.; Gaydu, M.; Geoffroy, M.H.; Veliov, V.M. 2011. "Metric Regularity of Newton's Iteration." SIAM Journal on Control and Optimization 49:339-362.

Argyros, I.K. 2022. The Theory and Applications of Iteration Methods with Applications. Engineering Series, Second Edition. Boca Raton, FL: CRC Press/Taylor & Francis.

Argyros, I.K.; George, S.; Shakhno, S.; Regmi, S.; Havdiak, M.; Argyros, M.I. 2024. "Asymptotically Newton-Type Methods Without Inverses for Solving Equations." Mathematics 12:1069. doi:10.3390/math12071069.

Argyros, I.K.; George, S.; Regmi, S.; Argyros, C.I. 2024. "Hybrid Newton-like Inverse Free Algorithms for Solving Nonlinear Equations." Algorithms 17:154. doi:10.3390/a17040154.

Argyros, I.K.; Shakhno, S.; Yarmola, H.; Regmi, S.; Shrestha, N. 2025. "Three step inverse free Kurchatov-like methods of convergence order close to four for equations." *Eur. J. Math. Anal.* 5:15-15.

Argyros, I.K.; Shrestha, N. 2017. "Extending the local convergence analysis of Newton's method." *Commun. Appl. Nonlinear Anal.* 24:49-60.

Blum, L.; Cucker, F.; Shub, M.; Smale, S. 1998. Complexity and Real Computation. New York: Springer-Verlag.

Burke, J.V.; Ferris, M.C. 1995. "A Gauss-Newton Method for Convex Composite Optimization." Mathematical Programming 71:179-194.

Chen, J. 2008. "The Convergence Analysis of Inexact Gauss-Newton Methods for Nonlinear Problems." Computational Optimization and Applications 40(1):97-118.

Chen, J.; Li, W. 2005. "Convergence of Gauss-Newton's Method and Uniqueness of the Solution." Applied Mathematics and Computation 170(1):686-705.

Dedieu, J.P.; Kim, M.H. 2002. "Newton's Method for Analytic Systems of Equations with Constant Rank Derivatives." Journal of Complexity 18(1):187-209.

Dennis Jr., J.E.; Schnabel, R.B. 1996. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Classics in Applied Mathematics, vol. 16. Philadelphia, PA: SIAM.

Deuflhard, P.; Heindl, G. 1979. "Affine Invariant Convergence Theorems for Newton's Method and Extensions to Related Methods." SIAM Journal on Numerical Analysis 16:1-10.

Deuflhard, P. 2004. Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms. Berlin; Heidelberg: Springer-Verlag.

Dontchev, A.L.; Rockafellar, R.T. 2009. Implicit Functions and Solution Mappings: A View from Variational Analysis. Springer Monographs in Mathematics. Dordrecht: Springer.

Ferreira, O.P.; Goncalves, M.L.N.; Oliveira, P.R. 2013. "Convergence of the Gauss-Newton Method for Convex Composite Optimization Under a Majorant Condition." SIAM Journal on Optimization 23(3):1757-1783.

Ferreira, O.P.; Goncalves, M.L.N.; Oliveira, P.R. 2011. "Local Convergence Analysis of the Gauss-Newton Method under a Majorant Condition." Journal of Complexity 27:111-125.

Ferreira, O.P.; Goncalves, M.L.N.; Oliveira, P.R. 2011. "Local Convergence Analysis of Inexact Newton-Like Methods under Majorant Condition." Computational Optimization and Applications 48:1-21.

Kantorovich, L.V.; Akilov, G.P. 1982. Functional Analysis. 2nd ed. Oxford: Pergamon Press. (translated from Russian by Howard L. Silcock).

Li, C.; Ng, K.F. 2007. "Majorizing Functions and Convergence of the Gauss-Newton Method for Convex Composite Optimization." SIAM Journal on Optimization 18:613-642.

Li, C.; Wang, X.H. 2002. "On Convergence of the Gauss-Newton Method for Convex Composite Optimization." Mathematical Programming 91:349-356.

Li, C.; Ng, K.F. 2012. "Convergence Analysis of the Gauss-Newton Method for Convex Inclusion and Convex-Composite Optimization Problems." Journal of Mathematical Analysis and Applications 389:469-485.

Nocedal, J.; Wright, S.J. 1999. Numerical Optimization. Springer-Verlag.

Proinov, P.D. 2009. "General Local Convergence Theory for a Class of Iterative Processes and Its Applications to Newton's Method." Journal of Complexity 25(1):38-62.

Robinson, S.M. 1972. "Extension of Newton's Method to Nonlinear Functions with Values in a Cone." Numerische Mathematik 19:341-347.

Robinson, S.M. 1972. "Normed Convex Process." Transactions of the American Mathematical Society 174:127-140.

Rockafellar, R.T. 1970. Convex Analysis. Princeton: Princeton University Press.

Rockafellar, R.T. 1967. Monotone Processes of Convex and Concave Type. Memoirs of the American Mathematical Society, no. 77.

Smale, S. 1986. "Newton's Method Estimates from Data at One Point." In The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics, 185-196. New York: Springer.

Stewart, G.W. 1969. "On the Continuity of the Generalized Inverse." SIAM Journal on Applied Mathematics 17:33-45.

Wedin, P.A. 1973. "Perturbation Theory for Pseudo-Inverses." BIT 13:217-232.