Journal of Computational Mathematics
Vol.xx, No.x, 2026, 1–17.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2512-m2025-0168

CUBICALLY CONVERGENT TWO-STEP GAUSS-NEWTON METHOD FOR NONSMOOTH EQUATIONS WITH APPLICATION TO AVE*

Jianhua Peng
Aviation Maintenance NCO School, Air Force Engineering University, Xinyang 464000, China
Jingyong Tang[1)]
College of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China
Email: tangjy@xynu.edu.cn

Abstract

We propose a two-step Gauss-Newton method (TS-GNM) for solving nonsmooth equations. At each iteration, the TS-GNM solves both a Gauss-Newton equation and an approximate Gauss-Newton equation. A second-order derivative-free line search strategy is designed to ensure the global convergence of TS-GNM. Under the nonsingularity condition and the strong semismoothness of the underlying function, we prove that the TS-GNM converges quadratically. Furthermore, we demonstrate that the TS-GNM achieves a cubic convergence rate when the generalized Jacobian is locally Lipschitz continuous at the solutions. Finally, we pay particular attention to the absolute value equation and present some numerical results.

Mathematics subject classification: 90C15, 90C33.
Key words: Nonlinear equations, Nonsmooth analysis, Gauss-Newton method, Cubic convergence, Absolute value equations.

1. Introduction

Pang and Qi [18] presented the motivation and algorithms for the nonsmooth equations

$$\mathrm{F}(\mathbf{x}) = 0, \tag{1.1}$$

where the mapping $\mathrm{F}(x) = (\mathrm{F}_1(\mathbf{x}), \ldots, \mathrm{F}_n(\mathbf{x})) : \mathbf{R}^n \to \mathbf{R}^n$ is assumed to be locally Lipschitzian, but is nonsmooth. The nonsmooth equations (1.1) provide a unified framework for the study of many important problems, including the nonlinear complementarity problem, the variational inequality problem, the Karush-Kuhn-Tucker system, the inequality feasibility problem, the maximal monotone operator, the LC^1 optimization problem, among others; see [18] for more information.

Among the numerical methods for solving (1.1), the generalized Newton method stands as one of the best-known and most powerful algorithms. The generalized Newton method takes the following iteration process: Letting $\mathbf{x}_k$ be the current iteration point, the next iteration point is generated by the formula

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k^{\mathrm{N}},$$

* Received July 1, 2025 / Revised version received October 17, 2025 / Accepted December 1, 2025 / Published online April 9, 2026 /
1) Corresponding author

where the generalized Newton direction $\mathbf{d}_k^{\mathrm{N}}$ is defined by

$$\mathbf{d}_k^{\mathrm{N}} = -\mathrm{V}_{\mathbf{x}_k}^{-1}\mathrm{F}(\mathbf{x}_k), \quad \mathrm{V}_{\mathbf{x}_k} \in \partial\mathrm{F}(\mathbf{x}_k).$$

Here $\partial\mathrm{F}(\mathbf{x}_k)$ denotes the Clarke generalized Jacobian of F at $\mathbf{x}_k$ (see Section 2 for details). Leveraging nonsmooth analysis techniques introduced by Clarke [2], Pang and Qi [18] established the superlinear convergence results for the generalized Newton method under the semismoothness and the BD-regularity of F. Concurrently, Qi and Sun [19] demonstrated the quadratic convergence of the generalized Newton method under the strong semismoothness of F and the uniform nonsingularity of elements in $\partial\mathrm{F}(\mathbf{x})$. Martínez and Qi [14] further investigated convergence properties of the inexact generalized Newton method under assumptions of semismoothness and BD-regularity. Jiang [7] presented the natural extension of the generalized Newton method for solving the Fischer-Burmeister equation of the nonlinear complementarity problem and gave its convergence analysis.

It is noticed that the generalized Newton direction $\mathbf{d}_k^{\mathrm{N}}$ requires the nonsingularity of the generalized Jacobian $\partial\mathrm{F}(\mathbf{x}_k)$. This requirement is simply too strong in general. To address this limitation, the modified Gauss-Newton method (GNM) is investigated, which employs the iterative process

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}},$$

where the Gauss-Newton direction $\mathbf{d}_k^{\mathrm{GN}}$ is defined by

$$\mathbf{d}_k^{\mathrm{GN}} = -\left(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\right)^{-1}\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{F}(\mathbf{x}_k), \quad \mathrm{V}_{\mathbf{x}_k} \in \partial\mathrm{F}(\mathbf{x}_k). \tag{1.2}$$

Here $\lambda_k > 0$ is updated iteratively, and I denotes the $n \times n$ identity matrix. The formulation (1.2) ensures that $\mathbf{d}_k^{\mathrm{GN}}$ remains well-defined without requiring the nonsingularity of $\partial\mathrm{F}(\mathbf{x}_k)$. Notably, the scheme (1.2) is formally identical to the classical Levenberg-Marquardt scheme for solving smooth equations, leading to GNM being frequently referred to as the nonsmooth Levenberg-Marquardt method in many literatures. For example, Qi *et al.* [20] proposed a parameter-self-adjusting Levenberg-Marquardt method (PSA-LMM) for solving a nonlinear system of equations in which the function is semismooth. In recent decades, the GNM has been extensively studied for solving nonsmooth equations reformulated from various optimization problems, including semiinfinite programming problems [10], nonlinear complementarity problems [7,23,25], vertical complementarity problems [24], stochastic linear complementarity problems [9], second-order cone complementarity problems [17], weighted linear complementarity problems [26], etc.

It is well-established that the GNM has the quadratic convergence rate under appropriate conditions, as demonstrated in works such as [25]. Recently, the two-step Levenberg-Marquardt methods for smooth equations have been extensively studied [1, 3, 4, 30, 33, 34], where each iteration incorporates both a standard LM step and an approximate LM step. These algorithms are shown to possess cubic convergence rate under the local error bound condition. Motivated by the work on two-step Levenberg-Marquardt methods, Tang *et al.* [28] proposed an accelerated smoothing Newton method for solving the weighted complementarity problems, combining a Newton step with an approximate Newton step at each iteration. This algorithm enjoys a local cubic convergence rate under proper nonsingularity and strong semismooth assumptions. Lately, Tang and Zhou [27] proposed a two-step Broyden-like method for solving smooth nonlinear equations, which is designed to compute an additional approximate quasi-Newton step at each iteration. The two-step Broyden-like method [27] is shown to have a convergence order $1 + 2t$ $(t \in (0, 1])$ under suitable assumptions.

In this paper, inspired by the aforementioned research, we design a two-step Gauss-Newton method for solving nonsmooth equations (1.1). At every iteration, the TS-GNM computes not only the standard Gauss-Newton direction $\mathbf{d}_k^{\mathrm{GN}}$ given in (1.2), but also an approximate Gauss-Newton direction $\mathbf{d}_k^{\mathrm{AGN}}$ defined by

$$\mathbf{d}_k^{\mathrm{AGN}} = -\big(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)^{-1} \mathrm{V}_{\mathbf{x}_k}^\top \mathrm{F}(\mathbf{w}_k), \quad \mathbf{w}_k := \mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}}, \tag{1.3}$$

where $\mathrm{V}_{\mathbf{x}_k}$ is the same generalized Jacobian element from (1.2). To ensure the global convergence of TS-GNM, a second-order derivative-free line search technique is developed. We prove that the proposed TS-GNM has quadratic convergence under the strong semismoothness of F and the nonsingularity of $\partial\mathrm{F}(\mathbf{x})$ at the solutions. Significantly, we further demonstrate that the TS-GNM can attain cubic convergence rate when $\partial\mathrm{F}(\mathbf{x})$ exhibits local Lipschitz continuity at the solutions. The practical application of TS-GNM is validated through numerical experiments on the absolute value equation. Our numerical results illustrate that the two-step strategy employed in TS-GNM leads to improvements in computational efficiency compared to the traditional GNM [7, 17, 25] and the PSA-LMM [20]. It is worth pointing out that the two-step algorithms discussed in previous papers [1, 3, 4, 27, 28, 30, 33, 34] aim to address smooth optimization problems. In contrast, the proposed TS-GNM is specifically designed to solve nonsmooth equations, a task that relies heavily on nonsmooth analysis. Moreover, the proposed TS-GNM differs from the two-step Gauss-Newton method introduced by Iakymchuk and Shakhno [6], which was designed for solving the nonlinear least squares problem. Using the wider generalized Lipschitz conditions for derivatives of the first and second orders, the authors [6] analyzed the convergence of their method, proving that it achieves a convergence order of $1+\sqrt{2}$ in case of zero residual.

The paper is organized as follows. In Section 2, we present the TS-GNM. The global and local convergence properties of TS-GNM are analyzed in Sections 3 and 4, respectively. The application of TS-GNM to the absolute value equation is given in Section 5. Finally, Section 6 concludes the paper.

2. A Two-step Gauss-Newton Method

In this section, we first review the definitions of the Clarke generalized Jacobian and the semismoothness of functions. Since the function F is locally Lipschitz continuous, F is differentiable almost everywhere in $\mathbf{R}^n$ by Rademacher's theorem. Let $D_{\mathrm{F}} \subset \mathbf{R}^n$ denote the set of points at which F is differentiable. For any $\mathbf{x} \in \mathbf{R}^n$, the set $\partial_B \mathrm{F}(\mathbf{x}) = \{\mathrm{V}_{\mathbf{x}} \in \mathbf{R}^{n\times n} | \exists \{\mathbf{x}_k\} \subset D_{\mathrm{F}} : \mathbf{x}_k \to \mathbf{x}, \mathrm{F}'(\mathbf{x}_k) \to \mathrm{V}_{\mathbf{x}}\}$ is nonempty and it is called the B-subdifferential of F at $\mathbf{x}$. If all members in $\partial_B\mathrm{F}(\mathbf{x})$ are nonsingular, then we say that F is BD-regular at $\mathbf{x}$. The Clarke generalized Jacobian of F at $\mathbf{x}$ is defined by $\partial\mathrm{F}(\mathbf{x}) = \mathrm{conv}(\partial_B\mathrm{F}(\mathbf{x}))$, where "conv" denotes the convex hull.

Proposition 2.1 ([29, Proposition 2.2]). *For all $x \in \mathbf{R}^n$, the following results hold:*

(a) $\partial_B\mathrm{F}(\mathbf{x})$ is nonempty and compact.

(b) $\partial\mathrm{F}(\mathbf{x})$ is nonempty, compact, and convex.

(c) The set-valued mappings $\partial_B\mathrm{F}$ and $\partial\mathrm{F}$ are locally bounded and upper semicontinuous.

Recall form [15] that the function F is said to be semismooth at a point $\mathbf{x} \in \mathbf{R}^n$ if F is directionally differentiable at $\mathbf{x}$ and for any $\mathbf{y} \to \mathbf{x}$ and $\mathrm{V}_{\mathbf{y}} \in \partial \mathrm{F}(\mathbf{y})$,

$$\mathrm{F}(\mathbf{y}) - \mathrm{F}(\mathbf{x}) - \mathrm{V}_{\mathbf{y}}(\mathbf{y} - \mathbf{x}) = o(\|\mathbf{y} - \mathbf{x}\|).$$

F is further said to be strongly semismooth at $\mathbf{x}$ if F is semismooth at $\mathbf{x}$ and for any $\mathbf{y} \to \mathbf{x}$ and $\mathrm{V}_{\mathbf{y}} \in \partial \mathrm{F}(\mathbf{y})$,

$$\mathrm{F}(\mathbf{y}) - \mathrm{F}(\mathbf{x}) - \mathrm{V}_{\mathbf{y}}(\mathbf{y} - \mathbf{x}) = \mathcal{O}(\|\mathbf{y} - \mathbf{x}\|^2).$$

A function F is said to be a semismooth (respectively, strongly semismooth) function if it is semismooth (respectively, strongly semismooth) everywhere in $\mathbf{R}^n$. From [15], we know that the class of semismooth functionals is very broad, including smooth functions, convex functions and piecewise-smooth functions. The sums, differences, products and composites of semismooth functions are semismooth. Moreover, by [19, Corollary 2.4], F is semismooth at $\mathbf{x}$ if each component of F is semismooth at $\mathbf{x}$.

Let F be given in (1.1). Define the merit function $\psi(\mathbf{x}) : \mathbf{R}^n \to \mathbf{R}$ as

$$\psi(\mathbf{x}) := \frac{1}{2}\|\mathrm{F}(\mathbf{x})\|^2.$$

The following lemma gives the continuous differentiability of the function $\psi(\mathbf{x})$.

Lemma 2.1 ([29, Lemma 4.2]). *If* F *is semismooth on* $\mathbf{R}^n$ *and each component function* F_i *of* F *is continuously differentiable on* $\mathbf{R}^n \backslash \mathrm{F}_i^{-1}(0)$*, then the function* $\psi(\mathbf{x}) = \|\mathrm{F}(\mathbf{x})\|^2/2$ *is continuously differentiable at any* $\mathbf{x} \in \mathbf{R}^n$ *with gradient*

$$\nabla\psi(\mathbf{x}) = \mathrm{V}_{\mathbf{x}}^{\top}\mathrm{F}(\mathbf{x}),$$

where $\mathrm{V}_{\mathbf{x}} \in \partial \mathrm{F}(\mathbf{x})$ *is arbitrary.*

To better illustrate the novelty of our proposed method, we first review the classical damped modified Gauss-Newton method for the nonsmooth equations (1.1), as presented in [7, 17, 25].

Algorithm 2.1: Algorithm GNM.

Choose $\mathbf{x}_0 \in \mathbf{R}^n$, $\sigma, \rho \in (0,1)$, $\lambda_0 > 0$, and set $k := 0$.

Step 1: Choose $\mathrm{V}_{\mathbf{x}_k} \in \partial \mathrm{F}(\mathbf{x}_k)$. If $\|\mathrm{V}_{\mathbf{x}_k}^{\top}\mathrm{F}(\mathbf{x}_k)\| = 0$, then stop.

Step 2: Compute the Gauss-Newton direction $\mathbf{d}_k^{\mathrm{GN}}$ by solving the linear system

$$\big(\mathrm{V}_{\mathbf{x}_k}^{\top}\mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)\mathbf{d} = -\mathrm{V}_{\mathbf{x}_k}^{\top}\mathrm{F}(\mathbf{x}_k).$$

Step 3: Let $\alpha_k = \rho^{l_k}$, where l_k is the smallest nonnegative integer l satisfying

$$\psi\big(\mathbf{x}_k + \rho^l \mathbf{d}_k^{\mathrm{GN}}\big) \le \psi(\mathbf{x}_k) + \sigma\rho^l \nabla\psi(\mathbf{x}_k)^{\top}\mathbf{d}_k^{\mathrm{GN}}.$$

Step 4: Choose $\lambda_{k+1} > 0$. Set $\mathbf{x}_{k+1} := \mathbf{x}_k + \alpha_k \mathbf{d}_k^{\mathrm{GN}}$, update $k := k+1$ and go to Step 1.

The GNM achieves superlinear or quadratic convergence under appropriate conditions, as demonstrated in [7,17,25]. It should be noted that the GNM computes only a single search direction per iteration and employs the classical Armijo-type line search to ensure global convergence. To further accelerate convergence, we introduce an additional approximate Gauss-Newton step at each iteration. Specifically, after computing the standard Gauss-Newton direction $\mathbf{d}_k^{\mathrm{GN}}$, we compute an approximate Gauss-Newton direction $\mathbf{d}_k^{\mathrm{AGN}}$ by solving a second linear system using the same generalized Jacobian element and regularization parameter. This two-step strategy is inspired by the success of two-step Levenberg-Marquardt methods for smooth equations [3], which have been shown to achieve cubic convergence under certain conditions. However, since our two-step Gauss-Newton method utilizes both $\mathbf{d}_k^{\mathrm{GN}}$ and $\mathbf{d}_k^{\mathrm{AGN}}$ to form a composite step, the classical Armijo-type line search is no longer suitable for ensuring global convergence. To address this, we design a second-order derivative-free line search strategy.

We now provide a detailed description of the two-step Gauss-Newton method.

Algorithm 2.2: Algorithm TS-GNM.

Choose constants $\rho, \gamma \in (0,1)$, $p_1 > 0$, $p_2 \geq 1$ and $\zeta > 0$. Choose a positive sequence $\{\zeta_k\}$ satisfying $\sum_{k=0}^{\infty} \zeta_k \leq \zeta < \infty$. Choose $\mathbf{x}_0 \in \mathbf{R}^n$. Set $k := 0$.

Step 1: Choose $\mathrm{V}_{\mathbf{x}_k} \in \partial \mathrm{F}(\mathbf{x}_k)$. If $\|\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{F}(\mathbf{x}_k)\| = 0$, then stop.

Step 2: Set

$$\lambda_k := p_1 \|\mathrm{F}(\mathbf{x}_k)\|^{p_2}. \tag{2.1}$$

Compute $\mathbf{d}_k^{\mathrm{GN}} \in \mathbf{R}^n$ by solving the linear system

$$\big(\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)\mathbf{d} = -\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{F}(\mathbf{x}_k). \tag{2.2}$$

Set $\mathbf{w}_k := \mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}}$. Compute $\mathbf{d}_k^{\mathrm{AGN}} \in \mathbf{R}^n$ by solving the linear system

$$\big(\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)\mathbf{d} = -\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{F}(\mathbf{w}_k). \tag{2.3}$$

Step 3: Set $\alpha_k := \rho^{l_k}$, where l_k is the smallest nonnegative integer l satisfying

$$\psi\big(\mathbf{x}_k + \rho^l\big(\mathbf{d}_k^{\mathrm{GN}} + \rho^l \mathbf{d}_k^{\mathrm{AGN}}\big)\big) \leq (1+\zeta_k)\psi(\mathbf{x}_k) - \gamma\big[\rho^l \psi(\mathbf{x}_k)\big]^2. \tag{2.4}$$

Step 4: Set $\mathbf{x}_{k+1} := \mathbf{x}_k + \alpha_k(\mathbf{d}_k^{\mathrm{GN}} + \alpha_k \mathbf{d}_k^{\mathrm{AGN}})$. Set $k := k+1$ and go to Step 1.

Remark 2.1. (i) TS-GNM is well-defined. In fact, if $\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{F}(\mathbf{x}_k) \neq 0$, then $\mathrm{F}(\mathbf{x}_k) \neq 0$ and $\lambda_k > 0$ by (2.1). This guarantees that the matrix $\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{V}_{\mathbf{x}_k} + \lambda_k I$ is symmetric and positive definite. Consequently, the linear systems (2.2) and (2.3) have unique solutions so that $\mathbf{d}_k^{\mathrm{GN}}$ and $\mathbf{d}_k^{\mathrm{AGN}}$ are well-defined. For the line search condition (2.4), as $l \to \infty$, the left-hand side of (2.4) tends to $\psi(\mathbf{x}_k)$ but the right-hand side tends to $(1+\zeta_k)\psi(\mathbf{x}_k)$. This asymptotic behavior implies that (2.4) holds for all sufficiently large $l > 0$. Thus, $\mathbf{x}_{k+1}$ is well generated.

(ii) Although TS-GNM requires solving two linear systems (2.2) and (2.3) at every iteration, the cost of obtaining $\mathbf{d}_k^{\mathrm{AGN}}$ is inexpensive. This is because the decomposition of $\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{V}_{\mathbf{x}_k} + \lambda_k I$ is available after solving (2.2), and only the function value $\mathrm{F}(\mathbf{w}_k)$ is needed to compute in solving (2.3).

3. Global Convergence

In the sequel, we assume $\mathrm{V}_{\mathbf{x}_k}^{\top} \mathrm{F}(\mathbf{x}_k) \neq 0$ for all $k \geq 0$, and establish the global convergence of TS-GNM under the following assumption.

Assumption 3.1. (a) The function F is semismooth.

(b) Each component function F_i of F is continuously differentiable on $\mathbf{R}^n \backslash \mathrm{F}_i^{-1}(0)$.

The conditions presented in Assumption 3.1 were introduced by Ulbrich [29] to establish the global convergence of a class of trust-region methods for bound-constrained semismooth systems of equations. As demonstrated in Lemma 2.1, these conditions guarantee that the merit function $\psi(\mathbf{x})$ is continuously differentiable on $\mathbf{R}^n$. It is worth pointing out that our subsequent analysis can be based directly on the weaker assumption that the merit function $\psi(\mathbf{x})$ is continuously differentiable on $\mathbf{R}^n$. However, since the conditions in Assumption 3.1 are more concrete and easier to verify, we decide to choose them as our working assumptions.

Lemma 3.1 ([8, Lemma 2.2]). *Let $\{a_k\}$ and $\{r_k\}$ be positive sequences satisfying $a_{k+1} \leq (1+r_k)a_k + r_k$ and $\sum_{k=0}^{\infty} r_k < \infty$. Then $\{a_k\}$ converges.*

Lemma 3.2 ([26, Corollary 2.1]). *Let $\{\mathbf{t}_k\} \subset \mathbf{R}^n$ be a sequence converging to $\mathbf{t}$, and $\mathrm{V}_{\mathbf{t}_k} \in \partial \mathrm{F}(\mathbf{t}_k)$ for all $k \geq 0$. Then $\{\mathrm{V}_{\mathbf{t}_k}\}$ is bounded. Moreover, if $\{\mathrm{V}_{\mathbf{t}_k}\}$ converges to $\mathrm{V}_{\mathbf{t}}$, then $\mathrm{V}_{\mathbf{t}} \in \partial \mathrm{F}(\mathbf{t})$.*

Theorem 3.1. *Let $\{\mathbf{x}_k\}$ be the sequence generated by* TS-GNM. *Under Assumption* 3.1, *any accumulation point $\mathbf{x}^*$ of $\{\mathbf{x}_k\}$ satisfies $\nabla \psi(\mathbf{x}^*) = 0$.*

Proof. According to Step 3 of Algorithm 2.2 (TS-GNM), $\psi(\mathbf{x}_{k+1}) \leq (1+\zeta_k)\psi(\mathbf{x}_k)$ holds for all $k \geq 0$. Thus, by Lemma 3.1, there exists a constant $\psi^* \geq 0$ such that $\lim_{k\to\infty} \psi(\mathbf{x}_k) = \psi^*$ and consequently

$$\lim_{k\to\infty} \|\mathrm{F}(\mathbf{x}_k)\| = \sqrt{2\psi^*}, \quad \lim_{k\to\infty} \lambda_k = p_1\big(\sqrt{2\psi^*}\big)^{p_2} =: \lambda^*.$$

In the following, we assume that $\psi^* > 0$. By (2.4), for all $k \geq 0$,

$$0 \leq \gamma[\alpha_k \psi(\mathbf{x}_k)]^2 \leq (1+\zeta_k)\psi(\mathbf{x}_k) - \psi(\mathbf{x}_{k+1}).$$

Since $\lim_{k\to\infty} \zeta_k = 0, \lim_{k\to\infty} \psi(\mathbf{x}_k) = \psi^*$, the above inequality implies that $\lim_{k\to\infty} \alpha_k \psi(\mathbf{x}_k) = 0$, which together with $\psi^* > 0$ gives $\lim_{k\to\infty} \alpha_k = 0$. Since $\mathbf{x}^*$ is an accumulation point of $\{\mathbf{x}_k\}$, there exists a subsequence $\{\mathbf{x}_k\}_{k\in K}$ of $\{\mathbf{x}_k\}$ converging to $\mathbf{x}^*$, where $K \subset \{0,1,\ldots\}$ is an infinite set. According to Lemma 3.2, $\{\mathrm{V}_{\mathbf{x}_k}\}_{k\in K}$ is bounded, implying that it has a convergent subsequence, denoted by $\{\mathrm{V}_{\mathbf{x}_k}\}_{k\in K_1}$ where $K_1 \subset K$ is an infinite set. We assume that $\lim_{k(\in K_1)\to\infty} \mathrm{V}_{\mathbf{x}_k} = \mathrm{V}_{\mathbf{x}^*}$. Then $\mathrm{V}_{\mathbf{x}^*} \in \partial \mathrm{F}(\mathbf{x}^*)$ by Lemma 3.2. Since $\lambda^* = p_1(\sqrt{2\psi^*})^{p_2} > 0$, the matrix $\mathrm{V}_{\mathbf{x}^*}^\top \mathrm{V}_{\mathbf{x}^*} + \lambda^* I$ is positive definite. Then

$$\lim_{k(\in K_1)\to\infty} \big(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)^{-1} = \big(\mathrm{V}_{\mathbf{x}^*}^\top \mathrm{V}_{\mathbf{x}^*} + \lambda^* I\big)^{-1},$$

which together with (2.2) yields

$$\lim_{k(\in K_1)\to\infty} \mathbf{d}_k^{\mathrm{GN}} = -\big(\mathrm{V}_{\mathbf{x}^*}^\top \mathrm{V}_{\mathbf{x}^*} + \lambda^* I\big)^{-1} \mathrm{V}_{\mathbf{x}^*}^\top \mathrm{F}(\mathbf{x}^*) =: \mathbf{d}^*. \tag{3.1}$$

Hence,

$$\lim_{k(\in K_1)\to\infty} \mathbf{w}_k = \lim_{k(\in K_1)\to\infty} \big(\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}}\big) = \mathbf{x}^* + \mathbf{d}^*$$

and by (2.3),

$$\lim_{k(\in K_1)\to\infty} \mathbf{d}_k^{\mathrm{AGN}} = -\big(\mathrm{V}_{\mathbf{x}^*}^\top \mathrm{V}_{\mathbf{x}^*} + \lambda^* I\big)^{-1} \mathrm{V}_{\mathbf{x}^*}^\top \mathrm{F}(\mathbf{x}^* + \mathbf{d}^*).$$

Note that by (2.4) for any sufficiently large $k \in K_1$,

$$\begin{aligned} &\psi\big(\mathbf{x}_k + \rho^{-1}\alpha_k\big(\mathbf{d}_k^{\mathrm{GN}} + \rho^{-1}\alpha_k \mathbf{d}_k^{\mathrm{AGN}}\big)\big) \\ &> (1+\zeta_k)\psi(\mathbf{x}_k) - \gamma\big[\rho^{-1}\alpha_k\psi(\mathbf{x}_k)\big]^2 \\ &\geq \psi(\mathbf{x}_k) - \gamma\big[\rho^{-1}\alpha_k\psi(\mathbf{x}_k)\big]^2, \end{aligned}$$

which gives

$$\frac{\psi\big(\mathbf{x}_k + \rho^{-1}\alpha_k\big(\mathbf{d}_k^{\mathrm{GN}} + \rho^{-1}\alpha_k \mathbf{d}_k^{\mathrm{AGN}}\big)\big) - \psi(\mathbf{x}_k)}{\rho^{-1}\alpha_k} > -\gamma\rho^{-1}\alpha_k\psi(\mathbf{x}_k)^2. \tag{3.2}$$

Hence, according to Lemma 2.1, we have from (3.2) that

$$\begin{aligned} \mathrm{F}(\mathbf{x}^*)^\top \mathrm{V}_{\mathbf{x}^*}\mathbf{d}^* &= \nabla\psi(\mathbf{x}^*)^\top \mathbf{d}^* \\ &= \lim_{k(\in K_1)\to\infty} \frac{\psi\big(\mathbf{x}_k + \rho^{-1}\alpha_k\big(\mathbf{d}_k^{\mathrm{GN}} + \rho^{-1}\alpha_k \mathbf{d}_k^{\mathrm{AGN}}\big)\big) - \psi(\mathbf{x}_k)}{\rho^{-1}\alpha_k} \\ &\geq \lim_{k(\in K_1)} -\gamma\rho^{-1}\alpha_k\psi(\mathbf{x}_k)^2 = 0. \end{aligned}$$

On the other hand, by (2.2) and the condition $\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{F}(\mathbf{x}_k) \neq 0$, it follows that $\mathbf{d}_k^{\mathrm{GN}} \neq 0$. Consequently,

$$\mathrm{F}(\mathbf{x}_k)^\top \mathrm{V}_{\mathbf{x}_k}\mathbf{d}_k^{\mathrm{GN}} = -\big(\mathbf{d}_k^{\mathrm{GN}}\big)^\top \big(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)\mathbf{d}_k^{\mathrm{GN}} < 0,$$

which implies that $\mathrm{F}(\mathbf{x}^*)^\top \mathrm{V}_{\mathbf{x}^*}\mathbf{d}^* \leq 0$. We may therefore conclude that $\mathrm{F}(\mathbf{x}^*)^\top \mathrm{V}_{\mathbf{x}^*}\mathbf{d}^* = 0$. Furthermore, by (3.1) we have

$$\mathrm{F}(\mathbf{x}^*)^\top \mathrm{V}_{\mathbf{x}^*}\big(\mathrm{V}_{\mathbf{x}^*}^\top \mathrm{V}_{\mathbf{x}^*} + \lambda^* I\big)^{-1} \mathrm{V}_{\mathbf{x}^*}^\top \mathrm{F}(\mathbf{x}^*) = 0,$$

which gives $\nabla\psi(\mathbf{x}^*) = \mathrm{V}_{\mathbf{x}^*}^\top \mathrm{F}(\mathbf{x}^*) = 0$, because the matrix $(\mathrm{V}_{\mathbf{x}^*}^\top \mathrm{V}_{\mathbf{x}^*} + \lambda^* I)^{-1}$ is positive definite. Note that if $\psi^* = 0$, then it holds that $\lim_{k\to\infty} \|\mathrm{F}(\mathbf{x}_k)\| = 0$. In this case, Lemma 2.1 also allows us to conclude that $\nabla\psi(\mathbf{x}^*) = 0$. We complete the proof. □

4. Local Cubic Convergence

In this section, we analyze the convergence rate of TS-GNM by using the (strong) semismoothness of the function F. The following assumption is needed in the remaining of the paper.

Assumption 4.1. The sequence $\{\mathbf{x}_k\}$ generated by TS-GNM has an accumulation point $\mathbf{x}^*$, and all $\mathrm{V}_{\mathbf{x}^*} \in \partial\mathrm{F}(\mathbf{x}^*)$ are nonsingular.

Lemma 4.1 ([18, Proposition 3]). *If all $\mathrm{V}_{\mathbf{x}} \in \partial\mathrm{F}(\mathbf{x})$ are nonsingular, then there is a neighborhood $\mathcal{N}$ of $\mathbf{x}$ and a constant $c > 0$ such that for any $\mathbf{y} \in \mathcal{N}$ and $\mathrm{V}_y \in \partial\mathrm{F}(\mathbf{y}), \mathrm{V}_y$ is nonsingular and $\|\mathrm{V}_y^{-1}\| \leq c$. If, furthermore, $\mathrm{F}(\mathbf{x}) = 0$ and $\mathrm{F}(\mathbf{x})$ is semismooth at $\mathbf{x}$, then there is a neighborhood $\tilde{\mathcal{N}}$ of $\mathbf{x}$ and a constant $\alpha > 0$ such that for any $y \in \tilde{\mathcal{N}}$,*

$$\|\mathrm{F}(\mathbf{y})\| \geq \alpha\|\mathbf{y} - \mathbf{x}\|. \tag{4.1}$$

Lemma 4.2. *Under Assumption* 4.1*, for all* $\mathbf{x}_k$ *sufficiently close to* $\mathbf{x}^*$*,*

$$\left\|\mathbf{d}_k^{\mathrm{GN}}\right\| = \mathcal{O}(\|\mathbf{x}_k - \mathbf{x}^*\|). \tag{4.2}$$

Proof. According to Lemma 4.1, for any $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$ and all $\mathrm{V}_{\mathbf{x}_k} \in \partial \mathrm{F}(\mathbf{x}_k)$, $\mathrm{V}_{\mathbf{x}_k}$ is nonsingular and

$$\left\|\mathrm{V}_{\mathbf{x}_k}^{-1}\right\| \le C \tag{4.3}$$

for some positive constant $C > 0$. Notice that for any matrix $\mathrm{M} \in \mathbf{R}^{n\times n}$ and $\lambda > 0$, by the singular value decomposition (SVD) of M, it is easy to verify that

$$\|(\mathrm{M}^\top \mathrm{M} + \lambda I)^{-1}\mathrm{M}^\top \mathrm{M}\| \le 1. \tag{4.4}$$

Thus, by (2.2), (4.3), (4.4) and the Lipschitz continuity of F, for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\begin{aligned}\left\|\mathbf{d}_k^{\mathrm{GN}}\right\| &= \left\|\left(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\right)^{-1}\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{F}(\mathbf{x}_k)\right\| \\ &\le \left\|\left(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\right)^{-1}\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k}\right\|\left\|\mathrm{V}_{\mathbf{x}_k}^{-1}\right\|\|\mathrm{F}(\mathbf{x}_k)\| \\ &\le C\|\mathrm{F}(\mathbf{x}_k)\| = \mathcal{O}(\|\mathbf{x}_k - \mathbf{x}^*\|).\end{aligned}$$

The proof is complete. □

Lemma 4.3. *Under Assumptions* 3.1 *and* 4.1*, if* $\mathrm{F}(\mathbf{x})$ *is semismooth (strongly semismooth, respectively) at* $\mathbf{x}^*$*, then* $\mathrm{F}(\mathbf{x}^*) = 0$ *and for all* $\mathbf{x}_k$ *sufficiently close to* $\mathbf{x}^*$*,*

$$\left\|\mathbf{d}_k^{\mathrm{AGN}}\right\| = o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \left(= \mathcal{O}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2\right)\right). \tag{4.5}$$

Proof. By Lemma 2.1 and Theorem 3.1, we have $\nabla\psi(\mathbf{x}^*) = \mathrm{V}_{\mathbf{x}^*}^\top \mathrm{F}(\mathbf{x}^*) = 0$. This together with the nonsingularity of $\mathrm{V}_{\mathbf{x}^*}$ gives $\mathrm{F}(\mathbf{x}^*) = 0$. Note that we may rewrite (2.2) as

$$\mathrm{F}(\mathbf{x}_k) + \mathrm{V}_{\mathbf{x}_k}\mathbf{d}_k^{\mathrm{GN}} + \lambda_k\left(\mathrm{V}_{\mathbf{x}_k}^\top\right)^{-1}\mathbf{d}_k^{\mathrm{GN}} = 0,$$

which is equivalent to

$$\mathrm{F}(\mathbf{x}_k) - \mathrm{V}_{\mathbf{x}_k}(\mathbf{x}_k - \mathbf{x}^*) + \lambda_k\left(\mathrm{V}_{\mathbf{x}_k}^\top\right)^{-1}\mathbf{d}_k^{\mathrm{GN}} + \mathrm{V}_{\mathbf{x}_k}\left(\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} - \mathbf{x}^*\right) = 0. \tag{4.6}$$

It follows from (4.3) and (4.6) that for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\begin{aligned}\|\mathbf{w}_k - \mathbf{x}^*\| = \left\|\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} - \mathbf{x}^*\right\| &= \left\|\mathrm{V}_{\mathbf{x}_k}^{-1}\left[\mathrm{F}(\mathbf{x}_k) - \mathrm{V}_{\mathbf{x}_k}(\mathbf{x}_k - \mathbf{x}^*) + \lambda_k\left(\mathrm{V}_{\mathbf{x}_k}^\top\right)^{-1}\mathbf{d}_k^{\mathrm{GN}}\right]\right\| \\ &\le C\left[\|\mathrm{F}(\mathbf{x}_k) - \mathrm{V}_{\mathbf{x}_k}(\mathbf{x}_k - \mathbf{x}^*)\| + C\lambda_k\left\|\mathbf{d}_k^{\mathrm{GN}}\right\|\right].\end{aligned} \tag{4.7}$$

Since $\mathrm{F}(\mathbf{x})$ is semismooth (strongly semismooth, respectively) at $\mathbf{x}^*$ and $\mathrm{F}(\mathbf{x}^*) = 0$, for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$ and $\mathrm{V}_{\mathbf{x}_k} \in \partial\mathrm{F}(\mathbf{x}_k)$,

$$\|\mathrm{F}(\mathbf{x}_k) - \mathrm{V}_{\mathbf{x}_k}(\mathbf{x}_k - \mathbf{x}^*)\| = o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \left(= \mathcal{O}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^2\right)\right). \tag{4.8}$$

Moreover, for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$, by the Lipschitz continuity of F,

$$\lambda_k = p_1\|\mathrm{F}(\mathbf{x}_k)\|^{p_2} = \mathcal{O}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^{p_2}\right), \tag{4.9}$$

which together with (4.2) yields

$$\lambda_k\|\mathbf{d}_k^{\mathrm{GN}}\| = \mathcal{O}\left(\|\mathbf{x}_k - \mathbf{x}^*\|^{1+p_2}\right). \tag{4.10}$$

Hence, by (4.7), (4.8), (4.10) and $p_2 \geq 1$, for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\|\mathbf{w}_k - \mathbf{x}^*\| = o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \big(= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^2\big)\big). \tag{4.11}$$

It follows that $\mathbf{w}_k$ is sufficiently close to $\mathbf{x}^*$ when $\mathbf{x}_k$ is sufficiently close to $\mathbf{x}^*$. By the Lipschitz continuity of F and (4.11), for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\|\mathrm{F}(\mathbf{w}_k)\| = O(\|\mathbf{w}_k - \mathbf{x}^*\|) = o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \big(= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^2\big)\big). \tag{4.12}$$

As a result, according to (2.3), (4.3), (4.4) and (4.12) for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\begin{aligned}
\big\|\mathbf{d}_k^{\mathrm{AGN}}\big\| &= \big\|\big(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)^{-1} \mathrm{V}_{\mathbf{x}_k}^\top \mathrm{F}(\mathbf{w}_k)\big\| \\
&\leq \big\|\big(\mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k} + \lambda_k I\big)^{-1} \mathrm{V}_{\mathbf{x}_k}^\top \mathrm{V}_{\mathbf{x}_k}\big\| \big\|\mathrm{V}_{\mathbf{x}_k}^{-1}\big\| \|\mathrm{F}(\mathbf{w}_k)\| \\
&\leq C\|\mathrm{F}(\mathbf{w}_k)\| \\
&= o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \big(= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^2\big)\big).
\end{aligned}$$

The proof is complete. □

Theorem 4.1. *Under Assumptions* 3.1 *and* 4.1*, if* $\mathrm{F}(\mathbf{x})$ *is semismooth (strongly semismooth, respectively) at* $\mathbf{x}^*$*, then* $\mathrm{F}(\mathbf{x}^*) = 0$ *and the whole sequence* $\{\mathbf{x}_k\}$ *converges to* $\mathbf{x}^*$ *with*

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| = o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \big(= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^2\big)\big).$$

Proof. Lemma 4.3 gives $\mathrm{F}(\mathbf{x}^*) = 0$. By (4.5) and (4.11), for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\begin{aligned}
&\big\|\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}} - \mathbf{x}^*\big\| \\
=\ &\big\|\mathbf{w}_k + \mathbf{d}_k^{\mathrm{AGN}} - \mathbf{x}^*\big\| \\
\leq\ &\|\mathbf{w}_k - \mathbf{x}^*\| + \big\|\mathbf{d}_k^{\mathrm{AGN}}\big\| \\
=\ &o(\|\mathbf{x}_k - \mathbf{x}^*\|) \quad \big(= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^2\big)\big).
\end{aligned} \tag{4.13}$$

This implies that $\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}}$ is sufficiently close to $\mathbf{x}^*$ when $\mathbf{x}_k$ is sufficiently close to $\mathbf{x}^*$. Hence, by the Lipschitz continuity of F, (4.1) and (4.13), for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\begin{aligned}
&\psi\big(\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}}\big) \\
=\ &\frac{1}{2}\big\|\mathrm{F}\big(\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}}\big)\big\|^2 \\
=\ &\mathcal{O}\big(\big\|\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}} - \mathbf{x}^*\big\|^2\big) \\
=\ &o\big(\|\mathbf{x}_k - \mathbf{x}^*\|^2\big) \quad \big(= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^4\big)\big) \\
=\ &o\big(\|\mathrm{F}(\mathbf{x}_k)\|^2\big) \quad \big(= \mathcal{O}\big(\|\mathrm{F}(\mathbf{x}_k)\|^4\big)\big) \\
=\ &o\big(\psi(\mathbf{x}_k)\big) \quad \big(= \mathcal{O}\big(\psi(\mathbf{x}_k)^2\big)\big).
\end{aligned}$$

Thus, for all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$,

$$\psi\big(\mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}}\big) + \gamma\psi(\mathbf{x}_k)^2 \leq (1 + \zeta_k)\psi(\mathbf{x}_k),$$

which implies that $\alpha_k = 1$ is always accepted in the Step 3 of Algorithm 2.2 (TS-GNM) and hence

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k^{\mathrm{GN}} + \mathbf{d}_k^{\mathrm{AGN}}. \tag{4.14}$$

This and (4.13) give $\lim_{k\to\infty}\mathbf{x}_k=\mathbf{x}^*$. Furthermore, by (4.2), we get $\lim_{k\to\infty}\mathbf{d}_k^{\rm GN}=0$ and

$$\lim_{k\to\infty}\mathbf{w}_k=\lim_{k\to\infty}\big(\mathbf{x}_k+\mathbf{d}_k^{\rm GN}\big)=\mathbf{x}^*.$$

For all $\mathbf{x}_k$ sufficiently close to $\mathbf{x}^*$, we may rewrite (2.3) as

$$\mathrm{F}(\mathbf{w}_k)+\mathrm{V}_{\mathbf{x}_k}\mathbf{d}_k^{\rm AGN}+\lambda_k\big(\mathrm{V}_{\mathbf{x}_k}^{\top}\big)^{-1}\mathbf{d}_k^{\rm AGN}=0,$$

which gives

$$\mathbf{d}_k^{\rm AGN}=-\mathrm{V}_{\mathbf{x}_k}^{-1}\big[\mathrm{F}(\mathbf{w}_k)+\lambda_k\big(\mathrm{V}_{\mathbf{x}_k}^{\top}\big)^{-1}\mathbf{d}_k^{\rm AGN}\big]. \tag{4.15}$$

Thus, by (4.3), (4.14) and (4.15), for all sufficiently large k and all $\mathrm{V}_{\mathbf{w}_k}\in\partial\mathrm{F}(\mathbf{w}_k)$,

$$\begin{aligned}\|\mathbf{x}_{k+1}-\mathbf{x}^*\|&=\big\|\mathbf{w}_k+\mathbf{d}_k^{\rm AGN}-\mathbf{x}^*\big\|\\&=\big\|\mathrm{V}_{\mathbf{x}_k}^{-1}\big[\mathrm{F}(\mathbf{w}_k)-\mathrm{V}_{\mathbf{x}_k}(\mathbf{w}_k-\mathbf{x}^*)+\lambda_k\big(\mathrm{V}_{\mathbf{x}_k}^{\top}\big)^{-1}\mathbf{d}_k^{\rm AGN}\big]\big\|\\&\le C\big[\|\mathrm{F}(\mathbf{w}_k)-\mathrm{V}_{\mathbf{w}_k}(\mathbf{w}_k-\mathbf{x}^*)\|\\&\qquad+\|\mathrm{V}_{\mathbf{w}_k}-\mathrm{V}_{\mathbf{x}_k}\|\|\mathbf{w}_k-\mathbf{x}^*\|+C\lambda_k\big\|\mathbf{d}_k^{\rm AGN}\big\|\big].\end{aligned} \tag{4.16}$$

Notice that by (4.5) and (4.9), for all sufficiently large k,

$$\lambda_k\big\|\mathbf{d}_k^{\rm AGN}\big\|=o\big(\|\mathbf{x}_k-\mathbf{x}^*\|^{1+p_2}\big)\quad\big(=\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^{2+p_2}\big)\big). \tag{4.17}$$

Moreover, by Lemma 3.2, the sequences $\{\|\mathrm{V}_{\mathbf{w}_k}\|\}$ and $\{\|\mathrm{V}_{\mathbf{x}_k}\|\}$ are bounded because both $\{\mathbf{w}_k\}$ and $\{\mathbf{x}_k\}$ converge to $\mathbf{x}^*$. Thus, by (4.11), (4.16), (4.17) and $p_2\ge 1$, if $\mathrm{F}(\mathbf{x})$ is semismooth at $\mathbf{x}^*$, then

$$\begin{aligned}\|\mathbf{x}_{k+1}-\mathbf{x}^*\|&=o(\|\mathbf{w}_k-\mathbf{x}^*\|)+\mathcal{O}(\|\mathbf{w}_k-\mathbf{x}^*\|)+o\big(\|\mathbf{x}_k-\mathbf{x}^*\|^{1+p_2}\big)\\&=o(\|\mathbf{x}_k-\mathbf{x}^*\|)+o\big(\|\mathbf{x}_k-\mathbf{x}^*\|^{1+p_2}\big)\\&=o(\|\mathbf{x}_k-\mathbf{x}^*\|),\end{aligned}$$

and if $\mathrm{F}(\mathbf{x})$ is strongly semismooth at $\mathbf{x}^*$, then

$$\begin{aligned}\|\mathbf{x}_{k+1}-\mathbf{x}^*\|&=\mathcal{O}\big(\|\mathbf{w}_k-\mathbf{x}^*\|^2\big)+\mathcal{O}(\|\mathbf{w}_k-\mathbf{x}^*\|)+\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^{2+p_2}\big)\\&=\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^4\big)+\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^2\big)+\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^{2+p_2}\big)\\&=\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^2\big).\end{aligned}$$

We complete the proof. □

To obtain the local cubic convergence rate of TS-GNM, we need the Lipschitz continuity of the generalized Jacobian $\partial\mathrm{F}$. Recall from [21, Definition 9.26] that a set-valued mapping $\mathrm{S}:\mathbf{R}^n\to\mathbf{R}^m$ is Lipschitz continuous on $\mathbf{X}$, a subset of $\mathbf{R}^n$, if it is nonempty closed-valued on $\mathbf{X}$ and there exists $\kappa\in\mathbf{R}_{++}$, a Lipschitz constant such that

$$\mathrm{S}(\mathbf{x}')\subset\mathrm{S}(\mathbf{x})+\kappa\|\mathbf{x}'-\mathbf{x}\|\mathbf{B},\quad\forall\,\mathbf{x}',\mathbf{x}\in\mathbf{X},$$

where $\mathbf{B}:=\{\mathbf{x}|\|\mathbf{x}\|\le 1\}$.

Theorem 4.2. *Under Assumptions* 3.1 *and* 4.1*, if* F *is strongly semismooth at* $\mathbf{x}^*$ *and the generalized Jacobian* $\partial\mathrm{F}$ *is locally Lipschitz continuous at* $\mathbf{x}^*$*, i.e., there exist* $\epsilon>0$ *and* $M>0$ *such that*

$$\partial\mathrm{F}(\mathbf{x}')\subset\partial\mathrm{F}(\mathbf{x})+M\|\mathbf{x}'-\mathbf{x}\|\mathbf{B},\quad\forall\,\mathbf{x}',\mathbf{x}\in N(\mathbf{x}^*,\epsilon), \tag{4.18}$$

then

$$\|\mathbf{x}_{k+1}-\mathbf{x}^*\|=\mathcal{O}\big(\|\mathbf{x}_k-\mathbf{x}^*\|^3\big).$$

Proof. Since both $\{\mathbf{x}_k\}$ and $\{\mathbf{w}_k\}$ converge to $\mathbf{x}^*$, when k is sufficiently large, by (4.18), for $\mathrm{V}_{\mathbf{x}_k} \in \partial \mathrm{F}(\mathbf{x}_k)$, there exists $\mathrm{V}_{\mathbf{w}_k} \in \partial \mathrm{F}(\mathbf{w}_k)$ such that

$$\|\mathrm{V}_{\mathbf{x}_k} - \mathrm{V}_{\mathbf{w}_k}\| \le M\|\mathbf{x}_k - \mathbf{w}_k\| = M\big\|\mathbf{d}_k^{\mathrm{GN}}\big\|. \tag{4.19}$$

This and (4.2) yield

$$\|\mathrm{V}_{\mathbf{x}_k} - \mathrm{V}_{\mathbf{w}_k}\| = \mathcal{O}(\|\mathbf{x}_k - \mathbf{x}^*\|). \tag{4.20}$$

Let $\mathrm{V}_{\mathbf{w}_k} \in \partial \mathrm{F}(\mathbf{w}_k)$ be chosen to satisfy (4.19). Since $\mathrm{F}(\mathbf{x})$ is strongly semismooth at $\mathbf{x}^*$, according to (4.11), (4.16), (4.17), (4.20) and $p_2 \ge 1$, we have

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \mathbf{x}^*\| &= \mathcal{O}\big(\|\mathbf{w}_k - \mathbf{x}^*\|^2\big) + \mathcal{O}(\|\mathbf{x}_k - \mathbf{x}^*\|\|\mathbf{w}_k - \mathbf{x}^*\|) + \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^{2+p_2}\big)\\
&= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^4\big) + \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^3\big) + \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^{2+p_2}\big)\\
&= \mathcal{O}\big(\|\mathbf{x}_k - \mathbf{x}^*\|^3\big).
\end{aligned}$$

We complete the proof. □

Remark 4.1. (i) The assumption that $\partial\mathrm{F}$ is locally Lipschitz continuous at $\mathbf{x}^*$ is equivalent to saying that $\partial\mathrm{F}$ is strictly continuous at $\mathbf{x}^*$ (see [21, Proposition 9.29]), i.e., for each $\rho > 0$, there exist $\epsilon > 0$ and $M > 0$ such that

$$\partial\mathrm{F}(\mathbf{x}') \cap \rho\mathbf{B} \subset \partial\mathrm{F}(\mathbf{x}) + M\|\mathbf{x} - \mathbf{x}'\|\mathbf{B}, \quad \forall\, \mathbf{x}', \mathbf{x} \in N(\mathbf{x}^*, \epsilon). \tag{4.21}$$

In fact, (4.18) clearly implies (4.21). Conversely, since $\partial\mathrm{F}$ is bounded, (4.21) implies (4.18) by letting ρ large enough to satisfy $\partial\mathrm{F}(\mathbf{x}') \cap \rho\mathbf{B} = \partial\mathrm{F}(\mathbf{x}')$.

(ii) Since $\partial\mathrm{F}$ is closed and upper semicontinuous, $\partial\mathrm{F}$ is outer semicontinuous by [21, Proposition 5.12]. Thus, by taking into account [21, Theorem 9.38], $\partial\mathrm{F}$ is strictly continuous at $\mathbf{x}^*$ if and only if $\partial\mathrm{F}$ has the Aubin property at $\mathbf{x}^*$ for every $\mathrm{V}_{\mathbf{x}^*} \in \partial\mathrm{F}(\mathbf{x}^*)$. Since the Lipschitz continuity is equivalent to the strict continuity as shown above, we claim that $\partial\mathrm{F}$ is locally Lipschitz continuous at $\mathbf{x}^*$ if and only if $\partial\mathrm{F}$ has the Aubin property at $\mathbf{x}^*$ for every $\mathrm{V}_{\mathbf{x}^*} \in \partial\mathrm{F}(\mathbf{x}^*)$. For the definition of Aubin property, one can see [21, Definition 9.36].

(iii) In Theorems 4.1 and 4.2, we need the strong semismoothness of F and the nonsingularity of $\partial\mathrm{F}(\mathbf{x})$ at the solutions. These assumptions are essential but by no means restrictive. For example, they hold for the absolute value equation, see the next section.

5. Application to the Absolute Value Equation

5.1. Absolute value equation

The absolute value equation (AVE) is defined to find $\mathbf{x} \in \mathbf{R}^n$ such that

$$\mathrm{F}(\mathbf{x}) := A\mathbf{x} - |\mathbf{x}| - b = 0, \tag{5.1}$$

where $A \in \mathbf{R}^{n\times n}$ and $b \in \mathbf{R}^n$ are known and $|\mathbf{x}|$ denotes the vector with absolute values of each component of $\mathbf{x}$. As shown in [13], the general linear complementarity problem which subsumes many mathematical programming problems can be formulated as the AVE. Over the past few decades, the AVE has been extensively studied in the literature, primarily due to its broad applications in mathematical programming and engineering fields [5,11,12,16,22,31,32].

Proposition 5.1. *Let* $\mathrm{F}(\mathbf{x})$ *be given in* (5.1). *Then the following results hold:*

(i) $\mathrm{F}(\mathbf{x})$ *is a globally Lipschitzian function in* $\mathbf{R}^n$*.*

(ii) The generalized Jacobian of F *at* $\mathbf{x}$ *is*

$$\partial\mathrm{F}(\mathbf{x}) = \{A - D_{\mathbf{x}}\},$$

in which $D_{\mathrm{x}} = \mathrm{diag}(\mathbf{d}_{\mathbf{x}})$ *for all* $\mathbf{d}_{\mathbf{x}} \in \mathbf{R}^n$ *such that for each* $i = 1, \ldots, n$*,*

$$(\mathbf{d}_{\mathbf{x}})_i = \begin{cases} 1, & \text{if} \quad \mathbf{x}_i > 0, \\ \xi, & \text{if} \quad \mathbf{x}_i = 0, \\ -1, & \text{if} \quad \mathbf{x}_i < 0, \end{cases}$$

where ξ *is any number in the interval* $[-1, 1]$*.*

(iii) If the interval matrix $[A - I, A - I]$ *is regular, then all* $\mathrm{V}_{\mathbf{x}} \in \partial\mathrm{F}(\mathbf{x})$ *are nonsingular.*

(iv) $\mathrm{F}(\mathbf{x})$ *is strongly semismooth on* $\mathbf{R}^n$*.*

Proof. The result (i) obviously holds. The results (ii) and (iii) come from [32]. The result (iv) holds because $|\mathbf{x}|$ is strongly semismooth on $\mathbf{R}^n$. □

Assume that $\mathbf{x}^*$ is a solution to the nonsmooth equations (5.1), that is, $\mathrm{F}(\mathbf{x}^*) = 0$. If $\mathbf{x}^*$ has no zero components (i.e., $\mathbf{x}_i^* \neq 0$ for all i), then in a sufficiently small neighborhood $N(\mathbf{x}^*, \epsilon)$ of $\mathbf{x}^*$, the sign of each component $\mathbf{x}_i$ of $\mathbf{x} \in N(\mathbf{x}^*, \epsilon)$ remains unchanged. Consequently, $D_{\mathbf{x}}$ is a constant matrix (i.e., the components $(\mathbf{d}_{\mathbf{x}})_i$ are fixed to 1 or -1) for all $\mathbf{x} \in N(\mathbf{x}^*, \epsilon)$. Under this condition, $\partial\mathrm{F}(\mathbf{x})$ is a singleton set for all $\mathbf{x} \in N(\mathbf{x}^*, \epsilon)$, thereby satisfying the local Lipschitz continuity condition (4.18).

5.2. Numerical examples

In this subsection, we apply the proposed TS-GNM to solving some specific AVEs. The parameters used in TS-GNM are set as follows: $p_1 = 10^{-3}, p_2 = 1, \gamma = 10^{-6}, \rho = 0.75, \zeta_k = 0.85^k$. For comparison, we also implement the traditional Gauss-Newton method [7,17,25] (denoted as GNM) and the parameter-self-adjusting Levenberg-Marquardt method [20] (denoted as PSA-LMM) to solve the same test problems.

First, we apply TS-GNM, GNM and PSA-LMM to solve the following four examples. The starting point is chosen as $\mathbf{x}_0 = (0, \ldots, 0)^\top$.

Example 5.1. Let the matrix $A \in \mathbf{R}^{n\times n}$ be given by

$$A = \begin{pmatrix} 4 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 4 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & 1 & 4 \end{pmatrix}.$$

Choose $b = \mathrm{rand}(n, 1)$.

Example 5.2. Let the matrix $A \in \mathbf{R}^{n\times n}$ be given by

$$a_{ii} = 4n, \quad a_{i,i+1} = a_{i+1,i} = n, \quad a_{ij} = 0.5, \quad i = 1, 2, \ldots, n.$$

Choose $b = (10, \ldots, 10)^\top$.

Example 5.3. Let the matrix $A \in \mathbf{R}^{n\times n}$ be given by

$$A = \text{round}\Big(100\big(\text{eye}(n,n) - 0.02\big(2\text{rand}(n,n) - 1\big)\big)\Big).$$

Choose $b = \text{rand}(n, 1)$.

Example 5.4. Consider the following ordinary differential equation (ODE):

$$\frac{d^2\mathbf{x}}{dt^2} - |\mathbf{x}| = (1 - t^2), \quad \mathbf{x}(0) = -1, \quad \mathbf{x}(1) = 0, \quad t \in [0, 1].$$

As explained in [5, Example 4.2], after discretization by using finite difference method, the above ODE can be recast the AVE (5.1) where the matrix $A \in \mathbf{R}^{n\times n}$ is given by

$$a_{i,j} = \begin{cases} -242, & i = j, \\ \ \ 121, & |i - j| = 1, \\ \ \ 0, & \text{otherwise.} \end{cases}$$

Let $e = (1, \ldots, 1)^\top$ and choose $b = Ae - e$.

To observe local convergence behaviors of TS-GNM, GNM and PSA-LMM, we present the results of solving Examples 5.1-5.4 with $n = 1000$. Table 5.1 lists the values of $\|\mathrm{F}(\mathbf{x}_k)\|$ at each iteration, while Figs. 5.1 and 5.2 illustrate the convergence behavior of three algorithms by plotting the logarithm of the residual norm $\|\mathrm{F}(\mathbf{x}_k)\|$ against the iterations. It is clearly observed that TS-GNM converges faster than GNM and PSA-LMM.

Table 5.2 presents the numerical results of TS-GNM, GNM and PSA-LMM in solving Examples 5.1-5.4 with varying sizes. Here, IT and CPU denote the number of iterations and CPU time (in seconds), respectively, while FK represents the value of $\|\mathrm{F}(\mathbf{x}_k)\|$ when the algorithms terminate. The stopping criterion adopted is $\|\mathrm{F}(\mathbf{x}_k)\| < 10^{-10}$. "-" stands for the fact that the

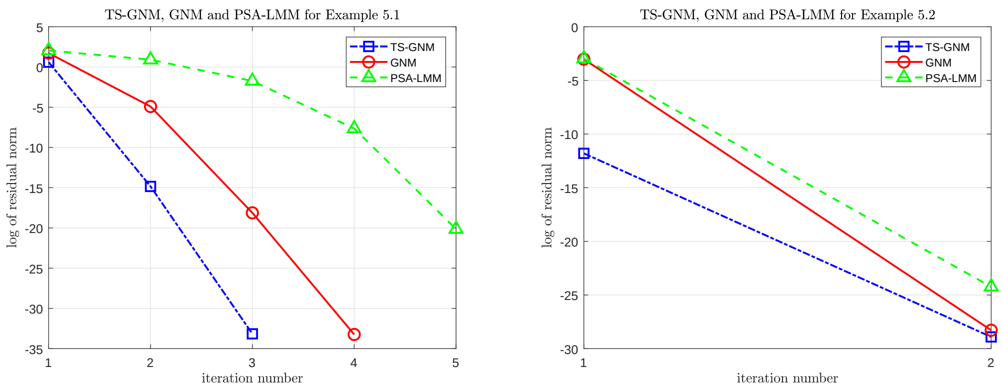

Fig. 5.1. The logarithm of residual norm $\|\mathrm{F}(\mathbf{x}_k)\|$ by iterations.

Table 5.1: The value of $\|F(\mathbf{x}_k)\|$ at the k-th iteration.

		TS-GNM	GNM	PSA-LMM
Example 5.1	$k = 1$	1.8115e+00	5.5846e+00	7.7432e+00
	$k = 2$	3.5813e-07	7.3671e-03	2.4713e+00
	$k = 3$	3.9748e-15	1.3433e-08	1.7564e-01
	$k = 4$	0	3.6740e-15	4.7860e-04
	$k = 5$	0	0	1.7844e-09
Example 5.2	$k = 1$	7.5013e-06	4.8687e-02	5.1055e-02
	$k = 2$	2.7867e-13	5.2011e-13	2.9888e-11
Example 5.3	$k = 1$	2.5215e-03	2.0168e-01	2.4507e-01
	$k = 2$	2.4282e-14	5.7288e-05	9.6122e-05
	$k = 3$	6.5199e-15	2.4684e-14	3.0346e-13
Example 5.4	$k = 1$	2.9775e+01	3.0577e+01	3.5074e+01
	$k = 2$	7.4021e+00	1.8553e+01	3.2129e+01
	$k = 3$	1.3466e+00	1.1735e+00	3.0754e+01
	$k = 4$	3.4717e-03	8.3730e-04	2.4679e+01
	$k = 5$	9.9347e-09	9.8935e-06	1.5402e+01

Table 5.2: Comparison of TS-GNM, GNM and PSA-LMM for Examples 5.1-5.4.

	n	TS-GNM			GNM			PSA-LMM		
		IT	CPU	FK	IT	CPU	FK	IT	CPU	FK
Example 5.1	6000	3	4.75	1.04e-14	4	5.73	9.58e-15	9	9.41	9.65e-15
	7000	3	7.04	1.13e-14	4	8.56	1.02e-14	9	14.30	1.02e-14
	8000	3	10.01	1.18e-14	4	12.12	1.09e-14	9	21.70	1.12e-14
	9000	3	14.40	1.25e-14	4	17.88	1.18e-14	9	33.61	1.16e-14
	10000	3	17.93	1.31e-14	4	23.10	1.22e-14	9	40.40	1.48e-14
Example 5.2	6000	2	2.75	1.11e-13	2	2.90	6.92e-12	2	2.39	1.33e-11
	7000	2	4.69	3.43e-13	2	4.26	5.90e-12	2	3.53	6.43e-12
	8000	2	5.91	7.00e-14	2	6.10	1.17e-12	2	4.97	2.72e-12
	9000	2	8.82	2.06e-12	3	12.46	1.16e-11	3	10.16	1.93e-11
	10000	2	11.90	9.95e-12	3	16.80	1.82e-11	3	13.47	1.02e-11
Example 5.3	2000	2	0.29	1.66e-14	3	0.34	1.92e-14	3	0.24	7.91e-12
	3000	2	0.58	3.03e-14	3	0.75	1.21e-13	4	0.78	2.99e-14
	4000	3	1.75	3.47e-14	3	1.58	3.25e-13	4	1.92	4.77e-14
	5000	3	2.99	7.02e-14	3	2.69	7.20e-12	4	3.15	2.67e-12
	6000	3	4.41	1.51e-13	4	5.96	1.62e-13	6	6.73	1.67e-13
Example 5.4	6000	5	8.66	1.83e-12	7	10.21	1.71e-12	-	-	-
	7000	5	12.26	4.00e-12	7	15.28	1.80e-12	-	-	-
	8000	6	22.01	2.34e-12	7	26.28	1.91e-12	-	-	-
	9000	5	23.74	1.30e-11	7	45.89	1.99e-12	-	-	-
	10000	5	39.10	2.90e-11	7	66.29	2.15e-12	-	-	-

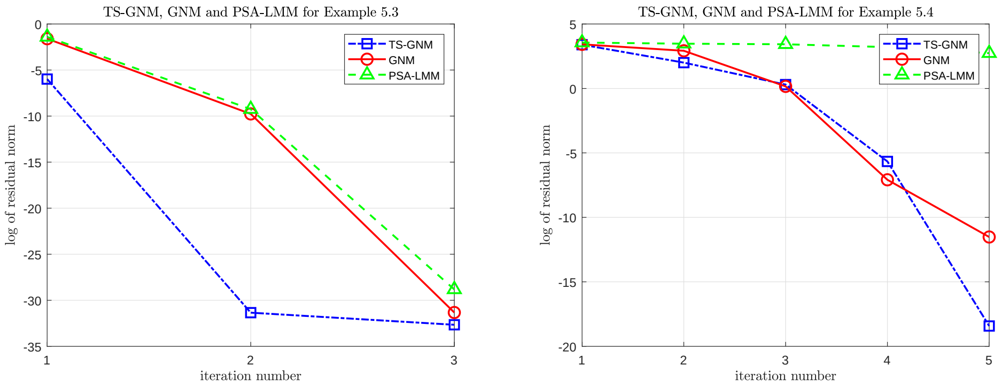

Fig. 5.2. The logarithm of residual norm $\|\mathrm{F}(\mathbf{x}_k)\|$ by iterations.

algorithm fails to solve the problems because of non-convergence. From Table 5.2, it can be observed that TS-GNM can effectively solve all test problems. Moreover, TS-GNM consistently requires fewer iterations than GNM and PSA-LMM to meet the stopping tolerance, and it also frequently consumes less CPU time.

Next, we apply TS-GNM, GNM and PSA-LMM to solving a class of AVE problems with larger condition numbers, respectively. The starting point is also selected as $\mathbf{x}_0 = (0, \ldots, 0)^\top$.

Example 5.5. Create the matrix A with exponentially decaying singular values by the following way. Take $U = \mathrm{orth}(\mathrm{randn}(n)), V = \mathrm{orth}(\mathrm{randn}(n)), s = \exp(-(1:n))$ with $s(1) = 1$ and $s(\mathrm{end}) = 1\mathrm{e}\text{-}15$. Let $A = U * \mathrm{diag}(s) * V^\top$. Let $e = (1, \ldots, 1)^\top$ and choose $b = Ae - e$.

Table 5.3 presents the numerical results of three algorithms for solving Example 5.5, where cond(A) denotes the condition number. We can observe that the proposed TS-GNM still performs very well in solving some challenging AVE problems.

Table 5.3: Comparison of TS-GNM, GNM and PSA-LMM for Example 5.5.

n	cond(A)	TS-GNM			GNM			PSA-LMM		
		IT	CPU	FK	IT	CPU	FK	IT	CPU	FK
500	5.88e+20	3	0.05	9.35e-13	3	0.06	9.35e-13	-	-	-
1000	7.03e+20	3	0.16	3.06e-12	5	0.31	3.04e-15	-	-	-
2000	7.84e+20	4	1.08	3.81e-15	-	-	-	-	-	-
2500	5.42e+21	4	1.74	4.05e-15	-	-	-	-	-	-

6. Conclusions

We proposed a two-step Gauss-Newton method to solve nonsmooth equations. TS-GNM is designed to compute an additional approximate Gauss-Newton step at each iteration. A second-order derivative-free line search technique is introduced in TS-GNM to ensure the global convergence. Under proper assumptions, TS-GNM is shown to possess local cubic convergence rate. Our numerical experiments showed that TS-GNM improves the computational efficiency compared with the traditional Gauss-Newton method and the parameter-self-adjusting Levenberg-Marquardt method in solving the absolute value equation.

Acknowledgements. The authors would like to thank two referees for their valuable suggestions that greatly improved the paper.

References

[1] K. Amini and F. Rostami, A new modified two-step Levenberg-Marquardt method for nonlinear equations, *J. Comput. Appl. Math.*, **288** (2015), 341–350.

[2] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, 1983.

[3] J.Y. Fan, The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence, *Math. Comput.*, **81** (2012), 447–466.

[4] J.Y. Fan, Accelerating the modified Levenberg-Marquardt method for nonlinear equations, *Math. Comput.*, **83** (2014), 1173–1187.

[5] F.K. Haghani, On generalized Traub's method for absolute value equations, *J. Optim. Theory Appl.*, **166** (2015), 619–625.

[6] R. Iakymchuk and S. Shakhno, On the convergence analysis of a two-step modification of the Gauss-Newton method, *PAMM. Proc. Appl. Math. Mech.*, **14** (2014), 813–814.

[7] H.Y. Jiang, Global convergence analysis of the generalized Newton and Gauss-Newton methods of the Fischer-Burmeister equation for the complementarity problem, *Math. Oper. Res.*, **24** (1999), 529–543.

[8] D.H. Li and M. Fukushima, A derivative-free line search and global convergence of Broyden-like method for nonlinear equations, *Optim. Methods Softw.*, **13** (2000), 181–201.

[9] Z. Liu, S. Du, and R. Wang, Nonsmooth Levenberg-Marquardt type method for solving a class of stochastic linear complementarity problems with finitely many elements, *Algorithms*, **9** (2016), 83.

[10] C. Ma and C.Y. Wang, A nonsmooth Levenberg-Marquardt method for solving semi-infinite programming problems, *J. Comput. Appl. Math.*, **230** (2009), 633–642.

[11] O.L. Mangasarian, Primal-dual bilinear programming solution of the absolute value equation, *Optim. Lett.*, **6** (2012), 1527–1533.

[12] O.L. Mangasarian, A hybrid algorithm for solving the absolute value equation, *Optim. Lett.*, **9** (2015), 1469–1474.

[13] O.L. Mangasarian and R.R. Meyer, Absolute value equations, *Linear Algebra Appl.*, **419** (2006), 359–367.

[14] M.J. Martínez and L.Q. Qi, Inexact Newton methods for solving nonsmooth equations, *J. Comput. Appl. Math.*, **60** (1995), 127–145.

[15] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, *SIAM J. Control Optim.*, **15** (1977), 957–972.

[16] M.A. Noor, J. Iqbal, K.I. Noor, and E. Al-Said, On an iterative method for solving absolute value equations, *Optim. Lett.*, **6** (2012), 1027–1033.

[17] S.H. Pan and J.S. Chen, A damped Gauss-Newton method for the second-order cone complementarity problem, *Appl. Math. Optim.*, **59** (2009), 293–318.

[18] J.S. Pang and L.Q. Qi, Nonsmooth equations: Motivation and algorithms, *SIAM J. Optim.*, **3** (1993), 443–465.

[19] L.Q. Qi and J. Sun, A nonsmooth version of Newton's method, *Math. Program.*, **58** (1993), 353–367.

[20] L.Y. Qi, X.T. Xiao, and L.W. Zhang, A parameter-self-adjusting Levenberg-Marquardt method for solving nonsmooth equations, *J. Comput. Math.*, **34** (2016), 317–338.

[21] R.T. Rockafellar and R.B. Wets, *Variational Analysis*, Springer, 1998.

[22] B. Sahey, C.H. Yu, and J.-S. Chen, Numerical comparisons based on four smoothing functions for absolute value equation, *J. Appl. Math. Comput.*, **56** (2018), 131–149.

[23] M. Smietański, On a nonsmooth Gauss-Newton algorithm for solving nonlinear complementarity problems, *Algorithms*, **13** (2020), 190.

[24] L. Song and Y. Gao, A nonsmooth Levenberg-Marquardt method for vertical complementarity problems, *Numer. Algorithms*, **76** (2017), 473–485.

[25] D.F. Sun and R.S. Womersley, A new unconstrained differentiable merit function for box constrained variational inequality problems and a damped Gauss-Newton method, *SIAM J. Optim.*, **9** (1999), 388–413.

[26] J.Y. Tang and J.C. Zhou, A modified damped Gauss-Newton method for non-monotone weighted linear complementarity problems, *Optim. Methods Softw.*, **37** (2022), 1145–1164.

[27] J.Y. Tang and J.C. Zhou, A two-step Broyden-like method for nonlinear equations, *Numer. Algorithms*, **98** (2025), 1085–1105.

[28] J.Y. Tang, J.C. Zhou, and H.C. Zhang, An accelerated smoothing Newton method with cubic convergence for weighted complementarity problems, *J. Optim. Theory Appl.*, **196** (2023), 641–665.

[29] M. Ulbrich, Nonmonotone trust-region methods for bound-constrained semismooth systems of equations with applications to nonlinear mixed complementarity problems, *SIAM J. Optim.*, **11** (2001), 889–917.

[30] H.Y. Wang, Convergence rate of the accelerated modified Levenberg-Marquardt method under Hölderian local error bound, *Pac. J. Optim.*, **18** (2022), 661–677.

[31] D.M. Yu, C.R. Chen, and D.R. Han, A modified fixed point iteration method for solving the system of absolute value equations, *Optimization*, **71** (2022), 449–461.

[32] C. Zhang and Q.J. Wei, Global and finite convergence of a generalized Newton method for absolute value equations, *J. Optim. Theory Appl.*, **143** (2009), 391–403.

[33] L. Zheng, L. Chen, and Y.F. Ma, Convergence rate of the modified Levenberg-Marquardt method under Hölderian local error bound, *Open Math.*, **20** (2022), 998–1012.

[34] W.J. Zhou, On the convergence of the modified Levenberg-Marquardt method with a nonmonotone second-order Armijo-type line search, *J. Comput. Appl. Math.*, **239** (2013), 152–161.