

参赛学生姓名：薛岚心

中学：南京外国语学校

省份：江苏省

国家/地区：中国

指导老师姓名：吴婷、刘鹏杰

指导老师单位：南京大学、中国矿业大学

论文题目：An Accelerated Inexact Newton-Type Projection

Method for Quasar-Monotone Nonlinear Equations

# An Accelerated Inexact Newton-Type Projection Method for Quasar-Monotone Nonlinear Equations

Lanxin Xue

**Abstract** Yin et al. [Numer. Linear Algebra Appl. 33(1), e70067 (2026)] proposed a Newton-type projection method with an iteration-complexity bound of  $O(\epsilon^{-4})$  for computing an  $\epsilon$ -approximate solution. To improve the iteration complexity under mild assumptions, we introduce a well-defined enhanced adaptive line search and, by incorporating a momentum-like acceleration mechanism, develop an accelerated inexact Newton-type projection method for solving unconstrained nonlinear equations with a quasar-monotone mapping, where quasar-monotonicity is weaker than both monotonicity and pseudo-monotonicity. We establish global convergence of the proposed method without requiring Lipschitz continuity of the underlying mapping. Moreover, under an additional local Lipschitz continuity assumption, the proposed method achieves an iteration-complexity bound of  $O(\epsilon^{-2})$ , thereby improving upon the  $O(\epsilon^{-4})$  result in the existing literature. Preliminary numerical experiments on standard nonlinear equations, as well as applications to Tikhonov-regularized logistic regression problems, demonstrate the effectiveness and practical potential of the proposed method.

**Keywords** Quasar-monotone nonlinear equations, Newton-type projection method, Adaptive line search, Global convergence, Iteration complexity

## Contents

|          |                                                                             |           |
|----------|-----------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                                         | <b>2</b>  |
| <b>2</b> | <b>Algorithm</b>                                                            | <b>3</b>  |
| <b>3</b> | <b>Convergence Result</b>                                                   | <b>5</b>  |
| 3.1.     | Global Convergence . . . . .                                                | 6         |
| 3.2.     | Iteration-Complexity Analysis . . . . .                                     | 9         |
| <b>4</b> | <b>Numerical Experiments</b>                                                | <b>11</b> |
| 4.1.     | Numerical Experiments on Standard SoNE Test Problems . . . . .              | 11        |
| 4.2.     | Numerical Experiments on Regularized Logistic Regression Problems . . . . . | 14        |
| <b>5</b> | <b>Conclusions</b>                                                          | <b>17</b> |

# 1 Introduction

In this paper, we consider a system of nonlinear equations (SoNE), which arises in numerous practical applications such as Tikhonov-regularized logistic regression [3, 12, 28], compressed sensing [26, 16], and traffic assignment problems [17, 18]. The SoNE can be formulated as

$$\mathfrak{E}(x) = 0, \quad (1.1)$$

where  $\mathfrak{E} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a continuous mapping. We assume that the solution set of problem (1.1), denoted by  $\Omega$ , is nonempty, and that  $\mathfrak{E}$  is quasar-monotone; that is, for any  $x^* \in \Omega$ ,

$$\mathfrak{E}(x)^\top (x - x^*) \geq 0, \quad \forall x \in \mathbb{R}^n. \quad (1.2)$$

The quasar-monotonicity condition, introduced in [21, 22], is weaker than both monotonicity and pseudo-monotonicity. The mapping  $\mathfrak{E}$  is said to be monotone if

$$\langle \mathfrak{E}(x) - \mathfrak{E}(z), x - z \rangle \geq 0, \quad \forall x, z \in \mathbb{R}^n,$$

and pseudo-monotone if

$$\langle \mathfrak{E}(z), x - z \rangle \geq 0 \implies \langle \mathfrak{E}(x), x - z \rangle \geq 0, \quad \forall x, z \in \mathbb{R}^n.$$

These concepts satisfy the following implication chain:

$$\text{monotone} \implies \text{pseudo-monotone} \implies \text{quasar-monotone},$$

while the converse implications do not generally hold. Counterexamples illustrating this fact can be found in the literature; see, for example, [9, 14] and the references therein for one-dimensional constructions.

Among the various methods for solving problem (1.1), the Newton method and its variants play a fundamental role. Given an iterate  $x_k \in \mathbb{R}^n$ , the Newton method updates the iterate according to  $x_{k+1} = x_k + d_k^N$ , where the search direction  $d_k^N$  is an (exact or approximate) solution of the linear system

$$\nabla \mathfrak{E}_k d = -\mathfrak{E}_k, \quad (1.3)$$

with  $\mathfrak{E}_k := \mathfrak{E}(x_k)$  and  $\nabla \mathfrak{E}_k$  denoting the Jacobian matrix of  $\mathfrak{E}$  at  $x_k$ , assuming that  $\mathfrak{E}$  is differentiable. If the initial point is sufficiently close to a solution  $x^*$  of problem (1.1), at which  $\nabla \mathfrak{E}(x^*)$  is nonsingular, then the sequence generated by the Newton method converges superlinearly or quadratically, depending on additional conditions. Nevertheless, the Newton method may become ineffective or computationally demanding when the Jacobian matrix is singular or nearly singular. In particular, if  $\nabla \mathfrak{E}_k$  is singular, the linear system (1.3) may fail to admit a solution. To address this difficulty, the Gauss–Newton method has been proposed, whose main idea is to solve a sequence of linear least-squares subproblems. Consider the merit function  $\Phi(x) := \frac{1}{2} \|\mathfrak{E}(x)\|^2$ . At each iteration, the search direction  $d_k^{\text{GN}}$  solves  $\min_{d \in \mathbb{R}^n} \frac{1}{2} \|\mathfrak{E}_k + \nabla \mathfrak{E}_k d\|^2$ , which is equivalent to

$$(\nabla \mathfrak{E}_k^\top \nabla \mathfrak{E}_k) d = -\nabla \Phi(x_k), \quad (1.4)$$

where  $\nabla\Phi(x_k) = \nabla\mathfrak{E}_k^\top \mathfrak{E}_k$ . It is straightforward to verify that the system (1.4) is consistent. The iterate is then updated by  $x_{k+1} = x_k + d_k^{\text{GN}}$ . If the initial point  $x_0$  lies sufficiently close to a solution  $x^* \in \Omega$  with a nonsingular Jacobian, the Gauss–Newton method enjoys quadratic convergence. However, similar to the Newton method, the system (1.4) becomes ill-conditioned when  $\nabla\mathfrak{E}_k^\top \nabla\mathfrak{E}_k$  is singular or nearly singular. To mitigate this issue, the Levenberg–Marquardt (LM) method [13, 20] incorporates a regularization strategy to enhance numerical stability. Yin et al. [32] proposed a modified inexact LM method by introducing an adaptive inexact criterion, and numerical results demonstrate its effectiveness and robustness.

For a SoNE problem of the form (1.1), Solodov and Svaiter [23] developed a derivative-free line search and introduced a projection technique, leading to a globally convergent inexact Newton method. Since then, this framework has attracted considerable attention, and a variety of (quasi-)Newton projection-type methods have been proposed; see, e.g., [35, 36, 37, 5, 34, 27]. Under the assumption that  $\mathfrak{E}$  is Lipschitz continuous, the global convergence of these methods has been established in [36, 37, 5, 27]. Yin et al. [31] combined an inertial-relaxed strategy with projection techniques and the line search to develop an inertial-relaxed Newton-type projection method, and derived an iteration-complexity bound of order  $O(\epsilon^{-4})$  for computing an  $\epsilon$ -approximate solution to problem (1.1). To the best of our knowledge, this is the first iteration-complexity result for Newton-type projection methods under their assumptions. Nevertheless, this complexity bound is still suboptimal, which motivates one of the main objectives of the present work.

Motivated by the above observations, this paper aims to improve the iteration complexity for solving unconstrained quasar-monotone SoNE (1.1) by incorporating an enhanced adaptive line search together with a momentum-like acceleration step, thereby proposing an accelerated inexact Newton-type projection (aiNP) method. The proposed method possesses the following theoretical properties: (i) The proposed new line search is shown to be well defined. (ii) Without requiring the Lipschitz continuity of the mapping  $\mathfrak{E}$ , the global convergence of the aiNP method is guaranteed. (iii) Under an additional local Lipschitz continuity assumption, the method achieves an iteration-complexity bound of  $O(\epsilon^{-2})$  for computing an  $\epsilon$ -approximate solution, improving upon the  $O(\epsilon^{-4})$  result reported in [31]. Finally, we demonstrate the numerical performance of the proposed aiNP method on standard nonlinear equations and Tikhonov-regularized logistic regression problems.

The remainder of this paper is organized as follows. Section 2 presents the proposed aiNP method and its iterative procedure. Section 3 establishes the global convergence and derives the iteration-complexity bound of the aiNP method. Numerical experiments are reported in Section 4. Finally, concluding remarks are given in the last section.

## 2 Algorithm

In this section, we first present an enhanced adaptive line search, followed by a detailed description of the proposed aiNP method (see Algorithm 1 below).

For inexact Newton-type projection methods, the choice of a suitable step size  $t_k > 0$  plays a crucial role in both theoretical analysis and numerical performance. Yin et al. [31] employed

the line search introduced in [33] to develop a Newton-type projection method. Specifically, the step size is determined as follows: given parameters  $\gamma, \sigma > 0$  and  $\tau \in (0, 1)$ , compute  $t_k := \max\{\gamma\tau^i : i = 0, 1, 2, \dots\}$  such that

$$-\mathfrak{E}(x_k + t_k d_k)^\top d_k \geq \sigma t_k \|d_k\|^2. \quad (2.1)$$

They established an iteration-complexity bound of order  $O(\epsilon^{-4})$ , which can be further improved. To this end, motivated by the line search in [24], we refine the line search (2.1) and propose an enhanced adaptive line search. Specifically, the step size is chosen as  $t_k := \max\{\gamma\tau^i : i = 0, 1, 2, \dots\}$  such that

$$-\mathfrak{E}(x_k + t_k d_k)^\top d_k \geq \sigma t_k \max\{\varkappa_1 \|\mathfrak{E}(x_k + t_k d_k)\| \|d_k\|, \varkappa_2 \|d_k\|^2\}, \quad (2.2)$$

where  $\varkappa_1 > 0$  and  $\varkappa_2 > 0$  are perturbation parameters whose selection is beyond the scope of this paper.

The following lemma shows that the proposed line search (2.2) is well defined.

**Lemma 2.1** If the search direction satisfies the descent property, then the line search (2.2) is well defined; that is, there exists a step size  $t_k > 0$  satisfying (2.2).

**Proof.** We establish the result by contradiction. Assume that, for all  $i \geq 0$ ,

$$-\mathfrak{E}(x_k + \gamma\tau^i d_k)^\top d_k < \sigma\gamma\tau^i \max\{\varkappa_1 \|\mathfrak{E}(x_k + \gamma\tau^i d_k)\| \|d_k\|, \varkappa_2 \|d_k\|^2\}.$$

Letting  $i \rightarrow \infty$  and using the continuity of  $\mathfrak{E}$ , we obtain  $-\mathfrak{E}(x_k)^\top d_k \leq 0$ , which contradicts the descent property of  $d_k$ . Hence, there exists a step size  $t_k > 0$  satisfying (2.2). This completes the proof.  $\square$

Building on the enhanced adaptive line search (2.2) proposed above and the momentum-like acceleration mechanism discussed in [29, 15], we now formally present the proposed aiNP method as follows.

**Remark 2.1** (i) Let  $\hat{x}_{k+1} = \tilde{x}_k - \frac{\mathfrak{E}(z_k)^\top (\tilde{x}_k - z_k)}{\|\mathfrak{E}(z_k)\|^2} \mathfrak{E}(z_k)$ . The update step (2.6) can be equivalently rewritten as

$$x_{k+1} = (1 - \rho)\tilde{x}_k + \rho \left( \tilde{x}_k - \frac{\mathfrak{E}(z_k)^\top (\tilde{x}_k - z_k)}{\|\mathfrak{E}(z_k)\|^2} \mathfrak{E}(z_k) \right) = (1 - \rho)\tilde{x}_k + \rho\hat{x}_{k+1}. \quad (2.7)$$

(ii) The inexactness criterion (2.3), adapted from [32, 31], is designed to control the approximation error. In particular, the first term on the right-hand side of (2.3) guarantees that  $\|q_k\| \leq p\sqrt{n}$  for all  $k$ , thereby preventing the solution of the system  $\mathfrak{E}(\tilde{x}_k) + (M_k + \nu_k I)d_k = 0$  associated with (2.4) from being excessively inaccurate.

(iii) The search direction generated in Step 2 of the aiNP method satisfies a sufficient descent condition, whose proof can be found in [31, Lemma 1]; that is, for all  $k \geq 0$ , it holds that

$$\mathfrak{E}(\tilde{x}_k)^\top d_k \leq -(1 - p)\nu_k \|d_k\|^2 < 0. \quad (2.8)$$

(iv) For a prescribed tolerance  $\epsilon > 0$  in numerical experiments, one may adopt  $\min\{\|\mathfrak{E}(x_k)\|, \|\mathfrak{E}(\tilde{x}_k)\|, \|\mathfrak{E}(z_k)\|\} \leq \epsilon$  and  $\|d_k\| \leq \epsilon$  as the stopping criteria. In particular, it is reasonable to use  $\|d_k\| \leq \epsilon$  as a termination condition. Indeed, if  $d_k = 0$ , then it follows from (2.3) that  $q_k = 0$ , which, together with the definition of  $q_k$ , implies  $\mathfrak{E}(\tilde{x}_k) = 0$ . As usual, a point  $\bar{x}$  is called an  $\epsilon$ -approximate solution to SoNE (1.1) if it satisfies  $\|\mathfrak{E}(\bar{x})\| \leq \epsilon$ .

---

**Algorithm 1** The aiNP method

---

- 1: **Step 0.** Initialize  $x_0 = x_{-1} \in \mathbb{R}^n$ . Choose parameters  $\alpha, p, \tau \in (0, 1)$ ,  $\lambda, \sigma, \gamma, \varkappa_1, \varkappa_2 > 0$ ,  $0 < \rho < 2$ , and  $\xi_0 \in [\xi_{\min}, \xi_{\max}]$  with  $0 < \xi_{\min} \leq \xi_{\max}$ . Set  $k := 0$ .
- 2: **Step 1.** Compute  $\|\mathfrak{E}(x_k)\|$ . If  $\|\mathfrak{E}(x_k)\| = 0$ , terminate. Select an acceleration parameter  $\alpha_k \in [0, \alpha]$  and compute  $\tilde{x}_k = x_k + \alpha_k(x_k - x_{k-1})$ . If  $\|\mathfrak{E}(\tilde{x}_k)\| = 0$ , terminate.
- 3: **Step 2.** Set  $\nu_k = \xi_k \|\mathfrak{E}(\tilde{x}_k)\|^\lambda$ . Choose a positive semidefinite matrix  $M_k$ , and compute  $d_k \in \mathbb{R}^n$  such that

$$\|q_k\| \leq p \min\{\sqrt{n}, \nu_k \|d_k\|\}, \quad (2.3)$$

where

$$q_k = \mathfrak{E}(\tilde{x}_k) + (M_k + \nu_k I)d_k. \quad (2.4)$$

If  $\|d_k\| = 0$ , terminate.

- 4: **Step 3.** Perform the enhanced adaptive line search. Compute the step size  $t_k := \gamma \tau^{i_k}$ , where  $i_k$  is the smallest nonnegative integer  $i$  such that

$$-\mathfrak{E}(\tilde{x}_k + \gamma \tau^i d_k)^\top d_k \geq \sigma \gamma \tau^i \max\{\varkappa_1 \|\mathfrak{E}(\tilde{x}_k + \gamma \tau^i d_k)\| \|d_k\|, \varkappa_2 \|d_k\|^2\}. \quad (2.5)$$

Set  $z_k = \tilde{x}_k + t_k d_k$ . If  $\|\mathfrak{E}(z_k)\| = 0$ , terminate.

- 5: **Step 4.** Update the iterate by

$$x_{k+1} = \tilde{x}_k - \rho \frac{\mathfrak{E}(z_k)^\top (\tilde{x}_k - z_k)}{\|\mathfrak{E}(z_k)\|^2} \mathfrak{E}(z_k). \quad (2.6)$$

- 6: **Step 5.** Choose  $\xi_{k+1} \in [\xi_{\min}, \xi_{\max}]$ , set  $k := k + 1$ , and return to Step 1.
- 

### 3 Convergence Result

In this section, we establish the global convergence and derive the iteration-complexity bound of the aiNP method. We first introduce the following assumption.

**Assumption 3.1** The sequences  $\{x_k\}$  and  $\{\alpha_k\}$  generated by the aiNP method satisfy  $\sum_{k=1}^{\infty} \alpha_k \|x_k - x_{k-1}\|^2 < \infty$ .

The following theorem provides a sufficient condition to guarantee that Assumption 3.1 holds by imposing mild and easily verifiable conditions on the sequence  $\{\alpha_k\}$ . Since its proof can be derived from [1, Theorem 2] and [30, Theorem 3.3], we omit the details.

**Theorem 3.1** Let  $\{x_k\}$  and  $\{\alpha_k\}$  be the sequences generated by the aiNP method. Assume that there exists a constant  $\alpha \in [0, 1)$  such that, for some  $\vartheta \in (0, 1)$ ,  $0 \leq \alpha_k \leq \alpha_{k+1} \leq \alpha < \vartheta < 1$  for all  $k \geq 0$ , and  $\rho = \rho(\vartheta) := \frac{2(\vartheta-1)^2}{2(\vartheta-1)^2 + 3\vartheta - 1} \in (0, 2)$ . Then  $\sum_{k=1}^{\infty} \alpha_k \|x_k - x_{k-1}\|^2 \leq \sum_{k=1}^{\infty} \|x_k - x_{k-1}\|^2 < \infty$ .

### 3.1. Global Convergence

**Lemma 3.1** Let  $\{x_k\}$ ,  $\{\tilde{x}_k\}$ , and  $\{z_k\}$  be the sequences generated by the aiNP method. Then, for all  $x^* \in \Omega$ ,

$$\|x_{k+1} - x^*\|^2 + \mu_{k+1} \leq \|\tilde{x}_k - x^*\|^2, \quad \forall k \geq 0, \quad (3.1)$$

where

$$\mu_{k+1} := (2 - \rho) \max \left\{ \rho \sigma^2 \kappa_1^2 t_k^2 \|\tilde{x}_k - z_k\|^2, \rho \sigma^2 \kappa_2^2 \frac{\|\tilde{x}_k - z_k\|^4}{\|\mathfrak{E}(z_k)\|^2}, \rho^{-1} \|x_{k+1} - \tilde{x}_k\|^2 \right\}. \quad (3.2)$$

**Proof.** From (2.5) and  $z_k = \tilde{x}_k + t_k d_k$ , it follows that

$$\mathfrak{E}(z_k)^\top (\tilde{x}_k - z_k) \geq \sigma t_k^2 \max\{\kappa_1 \|\mathfrak{E}(z_k)\| \|d_k\|, \kappa_2 \|d_k\|^2\} > 0. \quad (3.3)$$

Setting  $x = z_k$  in (1.2) yields  $\mathfrak{E}(z_k)^\top (z_k - x^*) \geq 0$ , or equivalently,

$$\mathfrak{E}(z_k)^\top (x^* - z_k) \leq 0. \quad (3.4)$$

This implies that  $\Omega \subseteq H_k^- := \{x \in \mathbb{R}^n \mid \mathfrak{E}(z_k)^\top (x - z_k) \leq 0\}$ . Combining (3.3) and (3.4), we deduce that the hyperplane  $H_k := \{x \in \mathbb{R}^n \mid \mathfrak{E}(z_k)^\top (x - z_k) = 0\}$  strictly separates  $\tilde{x}_k \notin H_k^-$  from the solution set  $\Omega \subseteq H_k^-$ . By the definition of  $\hat{x}_{k+1}$  in (2.7),  $\hat{x}_{k+1}$  is the projection of  $\tilde{x}_k$  onto  $H_k^-$ . Since  $x^* \in \Omega \subseteq H_k^-$ , it follows from [2, Lemma 2.2] that  $(\hat{x}_{k+1} - \tilde{x}_k)^\top (x^* - \hat{x}_{k+1}) \geq 0$ . A direct computation gives

$$-\frac{1}{2} \|\hat{x}_{k+1} - x^*\|^2 + \frac{1}{2} (\|\tilde{x}_k - x^*\|^2 - \|\hat{x}_{k+1} - \tilde{x}_k\|^2) \geq 0,$$

i.e.,

$$\|\tilde{x}_k - x^*\|^2 - \|\hat{x}_{k+1} - x^*\|^2 \geq \|\hat{x}_{k+1} - \tilde{x}_k\|^2. \quad (3.5)$$

It follows from (2.7) that  $x_{k+1} - x^* = (1 - \rho)(\tilde{x}_k - x^*) + \rho(\hat{x}_{k+1} - x^*)$ . Hence,

$$\begin{aligned} \|x_{k+1} - x^*\|^2 &= \|(1 - \rho)(\tilde{x}_k - x^*) + \rho(\hat{x}_{k+1} - x^*)\|^2 \\ &= (1 - \rho) \|\tilde{x}_k - x^*\|^2 + \rho \|\hat{x}_{k+1} - x^*\|^2 - \rho(1 - \rho) \|\hat{x}_{k+1} - \tilde{x}_k\|^2. \end{aligned}$$

Combining this with the definition of  $\hat{x}_{k+1}$ , (3.3), (3.5), and the fact that  $\rho \in (0, 2)$ , we obtain

$$\begin{aligned} \|\tilde{x}_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 &= \rho(\|\tilde{x}_k - x^*\|^2 - \|\hat{x}_{k+1} - x^*\|^2) + \rho(1 - \rho) \|\hat{x}_{k+1} - \tilde{x}_k\|^2 \\ &\geq \rho \|\hat{x}_{k+1} - \tilde{x}_k\|^2 + \rho(1 - \rho) \|\hat{x}_{k+1} - \tilde{x}_k\|^2 \\ &= \rho(2 - \rho) \|\hat{x}_{k+1} - \tilde{x}_k\|^2 \\ &= \rho(2 - \rho) \left\| \frac{\mathfrak{E}(z_k)^\top (\tilde{x}_k - z_k)}{\|\mathfrak{E}(z_k)\|^2} \mathfrak{E}(z_k) \right\|^2 \end{aligned} \quad (3.6)$$

$$\begin{aligned} &= \rho(2 - \rho) \left( \frac{\mathfrak{E}(z_k)^\top (\tilde{x}_k - z_k)}{\|\mathfrak{E}(z_k)\|} \right)^2 \\ &\geq \rho(2 - \rho) \left( \frac{\sigma t_k^2 \max\{\kappa_1 \|\mathfrak{E}(z_k)\| \|d_k\|, \kappa_2 \|d_k\|^2\}}{\|\mathfrak{E}(z_k)\|} \right)^2. \end{aligned} \quad (3.7)$$

We consider two cases for (3.7).

- (i) When  $\max\{\varkappa_1\|\mathfrak{E}(z_k)\| \|d_k\|, \varkappa_2\|d_k\|^2\} = \varkappa_1\|\mathfrak{E}(z_k)\| \|d_k\|$ , combining (3.7) with  $z_k = \tilde{x}_k + t_k d_k$ , we obtain

$$\|\tilde{x}_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 \geq \rho(2 - \rho)\sigma^2 \varkappa_1^2 t_k^4 \|d_k\|^2 = \rho(2 - \rho)\sigma^2 \varkappa_1^2 t_k^2 \|\tilde{x}_k - z_k\|^2.$$

- (ii) When  $\max\{\varkappa_1\|\mathfrak{E}(z_k)\| \|d_k\|, \varkappa_2\|d_k\|^2\} = \varkappa_2\|d_k\|^2$ , combining (3.7) with  $z_k = \tilde{x}_k + t_k d_k$ , we obtain

$$\|\tilde{x}_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 \geq \rho(2 - \rho)\sigma^2 \frac{t_k^4 \varkappa_2^2 \|d_k\|^4}{\|\mathfrak{E}(z_k)\|^2} = \rho(2 - \rho)\sigma^2 \varkappa_2^2 \frac{\|\tilde{x}_k - z_k\|^4}{\|\mathfrak{E}(z_k)\|^2}.$$

On the other hand, from (2.6), (3.6), and the fact that  $\rho \in (0, 2)$ , we have

$$\|\tilde{x}_k - x^*\|^2 - \|x_{k+1} - x^*\|^2 \geq \rho^{-1}(2 - \rho)\|x_{k+1} - \tilde{x}_k\|^2. \quad (3.8)$$

Combining the above two cases with (3.8), and the definition of  $\mu_k$  in (3.2), we conclude that (3.1) holds.  $\square$

Based on Lemma 3.1, Lemma 3.2 can be established, and its proof is similar to that of [29, Lemma 3.2] and [15, Lemma 3.2]. Building on this result, we further derive Theorem 3.2, which is frequently used in the convergence analysis of inertial-type algorithms; see, for example, [29, 1, 15] and the references therein. The proofs of Lemma 3.2 and Theorem 3.2 are therefore omitted.

**Lemma 3.2** Let  $\{x_k\}$ ,  $\{\tilde{x}_k\}$ ,  $\{z_k\}$ , and  $\{\alpha_k\}$  be the sequences generated by the aiNP method. For all  $x^* \in \Omega$ , define  $\psi_k := \|x_k - x^*\|^2$  ( $k \geq -1$ ) and  $\delta_k := \alpha_k(1 + \alpha_k)\|x_k - x_{k-1}\|^2$  ( $k \geq 0$ ), and let  $\mu_k$  be defined in (3.2). Then  $\psi_0 = \psi_{-1}$  and  $\psi_{k+1} - \psi_k + \mu_{k+1} \leq \alpha_k(\psi_k - \psi_{k-1}) + \delta_k$  for all  $k \geq 0$ .

**Theorem 3.2** Let  $\{x_k\}$ ,  $\{\tilde{x}_k\}$ ,  $\{z_k\}$ , and  $\{\alpha_k\}$  be the sequences generated by the aiNP method. Suppose that Assumption 3.1 is satisfied. Then, for all  $x^* \in \Omega$ , (i)–(ii) hold.

- (i) The limit  $\lim_{k \rightarrow \infty} \|x_k - x^*\|$  exists.
- (ii) The sequence  $\{\mu_k\}$  defined in (3.2) satisfies  $\sum_{k=1}^{\infty} \mu_k < \infty$ .

**Corollary 3.1** Suppose that the assumptions of Theorem 3.2 hold. Then the following statements are valid:

- (i)  $\lim_{k \rightarrow \infty} t_k \|\tilde{x}_k - z_k\| = \lim_{k \rightarrow \infty} \frac{\|\tilde{x}_k - z_k\|^2}{\|\mathfrak{E}(z_k)\|} = \lim_{k \rightarrow \infty} \|x_k - \tilde{x}_{k-1}\| = 0$ .
- (ii) The sequences  $\{x_k\}$  and  $\{\tilde{x}_k\}$  are bounded.

**Proof.** Statement (i) follows from Theorem 3.2(ii) and the definition of  $\mu_k$  in (3.2). Moreover, Theorem 3.2(i) implies that  $\{x_k\}$  is bounded. Combined with the fact that  $\lim_{k \rightarrow \infty} \|x_k - \tilde{x}_{k-1}\| = 0$ , this further shows that  $\{\tilde{x}_k\}$  is also bounded.  $\square$

**Theorem 3.3** Let  $\{x_k\}$ ,  $\{\tilde{x}_k\}$ , and  $\{z_k\}$  be the sequences generated by the aiNP method. Suppose that Assumption 3.1 holds and that there exists a constant  $Q_0 > 0$  such that  $\|M_k\| \leq Q_0$  for all  $k \geq 0$ . Then the sequences  $\{x_k\}$ ,  $\{\tilde{x}_k\}$ , and  $\{z_k\}$  all converge to the same solution of problem (1.1).

**Proof.** We divide the proof into two parts.

(i) We show that  $\liminf_{k \rightarrow \infty} \|\mathfrak{E}(\tilde{x}_k)\| = 0$ . Suppose, by contradiction, that  $\liminf_{k \rightarrow \infty} \|\mathfrak{E}(\tilde{x}_k)\| > 0$ . Then there exists a constant  $Q_1 > 0$  such that

$$\|\mathfrak{E}(\tilde{x}_k)\| \geq Q_1, \quad \forall k \geq 0. \quad (3.9)$$

By the continuity of  $\mathfrak{E}$  and the boundedness of  $\{\tilde{x}_k\}$ , there exists a constant  $Q_2 \geq Q_1$  such that

$$\|\mathfrak{E}(\tilde{x}_k)\| \leq Q_2, \quad \forall k \geq 0. \quad (3.10)$$

Combining (3.9) and (3.10) with the definition of  $\nu_k$  and  $0 < \xi_{\min} \leq \xi_k \leq \xi_{\max}$ , we obtain

$$\xi_{\min} Q_1^\lambda \leq \nu_k = \xi_k \|\mathfrak{E}(\tilde{x}_k)\|^\lambda \leq \xi_{\max} Q_2^\lambda, \quad \forall k \geq 0. \quad (3.11)$$

It follows from (2.3), (2.4), the triangle inequality, and  $\|M_k\| \leq Q_0$  that

$$\begin{aligned} \|\mathfrak{E}(\tilde{x}_k)\| &\leq \|(M_k + \nu_k I)d_k\| + \|q_k\| \leq (\|M_k\| + \nu_k)\|d_k\| + p\nu_k\|d_k\| \\ &\leq (Q_0 + (1+p)\xi_{\max}Q_2^\lambda)\|d_k\| =: Q_3\|d_k\|, \end{aligned} \quad (3.12)$$

where  $Q_3 := Q_0 + (1+p)\xi_{\max}Q_2^\lambda$ . Hence, by (3.9),

$$\|d_k\| \geq Q_1/Q_3, \quad \forall k \geq 0. \quad (3.13)$$

On the other hand, from (2.8) and the Cauchy–Schwarz inequality, we have

$$\|d_k\| \leq \frac{\|\mathfrak{E}(\tilde{x}_k)\|}{(1-p)\nu_k} \leq \frac{Q_2}{(1-p)\xi_{\min}Q_1^\lambda}, \quad (3.14)$$

where the last inequality follows from (3.10) and (3.11). Hence,  $\{d_k\}$  is bounded. Together with  $z_k = \tilde{x}_k + t_k d_k$ ,  $0 < t_k \leq \gamma$ , and the boundedness of  $\{\tilde{x}_k\}$ , this implies that  $\{z_k\}$  is bounded. By continuity of  $\mathfrak{E}$ , the sequence  $\{\mathfrak{E}(z_k)\}$  is also bounded, and thus there exists a constant  $Q_4 > 0$  such that  $\|\mathfrak{E}(z_k)\| \leq Q_4$  for all  $k \geq 0$ , which implies that  $\frac{\|\tilde{x}_k - z_k\|^2}{Q_4} \leq \frac{\|\tilde{x}_k - z_k\|^2}{\|\mathfrak{E}(z_k)\|^2}$ . Combining this with Corollary 3.1(i), we deduce that

$$\lim_{k \rightarrow \infty} \|\tilde{x}_k - z_k\| = \lim_{k \rightarrow \infty} t_k \|d_k\| = 0. \quad (3.15)$$

Together with (3.13) and  $z_k = \tilde{x}_k + t_k d_k$ , this yields  $\lim_{k \rightarrow \infty} t_k = 0$ . By the adaptive line search (2.5), we have

$$-\mathfrak{E}(\tilde{x}_k + \tau^{-1} t_k d_k)^\top d_k < \sigma \tau^{-1} t_k \max\{\varkappa_1 \|\mathfrak{E}(\tilde{x}_k + \tau^{-1} t_k d_k)\| \|d_k\|, \varkappa_2 \|d_k\|^2\}. \quad (3.16)$$

Since  $\{\tilde{x}_k\}$  and  $\{d_k\}$  are bounded, there exists a subsequence  $\{k_i\}$  such that

$$\tilde{x}_{k_i} \rightarrow \tilde{x}' \text{ and } d_{k_i} \rightarrow d'. \quad (3.17)$$

Taking  $k = k_i$  in (3.16), passing to the limit, and using  $\lim_{k \rightarrow \infty} t_k = 0$  and (3.17), we obtain

$$-\mathfrak{E}(\tilde{x}')^\top d' \leq 0. \quad (3.18)$$

On the other hand, by (2.8), (3.11), and (3.13), we have

$$-\mathfrak{E}(\tilde{x}_k)^\top d_k \geq (1-p)\nu_k \|d_k\|^2 \geq (1-p)\xi_{\min} Q_1^\lambda \|d_k\|^2 \geq (1-p)\xi_{\min} Q_1^\lambda \frac{Q_1^2}{Q_3^2} > 0,$$

which contradicts (3.18). Therefore,  $\liminf_{k \rightarrow \infty} \|\mathfrak{E}(\tilde{x}_k)\| = 0$ .

(ii) We show that  $\{x_k\}$ ,  $\{\tilde{x}_k\}$ , and  $\{z_k\}$  all converge to the same solution of problem (1.1). By the boundedness of  $\{\tilde{x}_k\}$ , the continuity of  $\mathfrak{E}$ , and the fact that  $\liminf_{k \rightarrow \infty} \|\mathfrak{E}(\tilde{x}_k)\| = 0$ , there exists a subsequence  $\{k_i\}$  such that  $\tilde{x}_{k_i} \rightarrow \tilde{x}'$  and  $\|\mathfrak{E}(\tilde{x}_{k_i})\| \rightarrow 0$ . Hence, by continuity of  $\mathfrak{E}$ , it follows that  $\mathfrak{E}(\tilde{x}') = 0$ , that is,  $\tilde{x}' \in \Omega$ . Combining  $\tilde{x}_{k_i} \rightarrow \tilde{x}'$  with Corollary 3.1(i), we deduce that  $\tilde{x}'$  is also an accumulation point of  $\{x_k\}$ , and  $x_{k_i+1} \rightarrow \tilde{x}'$ . Setting  $x^* := \tilde{x}'$  in Theorem 3.2, we obtain  $\lim_{k \rightarrow \infty} \|x_k - \tilde{x}'\| = \lim_{i \rightarrow \infty} \|x_{k_i+1} - \tilde{x}'\| = 0$ . Therefore, the entire sequence  $\{x_k\}$  converges to  $\tilde{x}' \in \Omega$ . By Corollary 3.1(i) and (3.15), it follows that  $\{\tilde{x}_k\}$  and  $\{z_k\}$  also converge to  $\tilde{x}'$ . This completes the proof.  $\square$

### 3.2. Iteration-Complexity Analysis

Next, we derive an iteration-complexity bound for computing an  $\epsilon$ -approximate solution to problem (1.1). For this purpose, we impose an additional assumption.

**Assumption 3.2** The mapping  $\mathfrak{E}$  in problem (1.1) is locally Lipschitz continuous.

From Theorem 3.2(ii) and the definition of  $\mu_k$ , it follows that  $\sum_{k=0}^{\infty} t_k^2 \|\tilde{x}_k - z_k\|^2 < \infty$ . Consequently, there exists a constant  $Q_5 > 0$  such that

$$\sum_{i=0}^{k-1} t_i^2 \|\tilde{x}_i - z_i\|^2 \leq Q_5, \quad \forall k \geq 1. \quad (3.19)$$

**Lemma 3.3** Suppose that the assumptions of Theorem 3.3 and Assumption 3.2 hold. For a prescribed tolerance  $\epsilon > 0$ , if  $\|\mathfrak{E}(\tilde{x}_k)\| > \epsilon$  for all  $k \geq 0$ , then the following statements hold:

- (i) There exists a constant  $Q_6 \in (0, \gamma]$  such that  $t_k \geq Q_6$  for all  $k \geq 0$ .
- (ii) There exists a constant  $Q_7 > 0$  such that  $\sum_{i=0}^{k-1} \|d_i\|^2 \leq Q_7$  for all  $k \geq 1$ .

**Proof.** (i) Since  $\|\mathfrak{E}(\tilde{x}_k)\| > \epsilon$  for all  $k \geq 0$ , it follows from the definition of  $\nu_k$  and  $\xi_k \geq \xi_{\min} > 0$  that

$$\nu_k = \xi_k \|\mathfrak{E}(\tilde{x}_k)\|^\lambda > \xi_{\min} \epsilon^\lambda, \quad \forall k \geq 0. \quad (3.20)$$

Using arguments analogous to those leading to (3.14), we conclude that  $\{d_k\}$  is bounded. Let  $\hat{t}_k := \tau^{-1} t_k$  and  $\hat{z}_k := \tilde{x}_k + \hat{t}_k d_k$ . By the boundedness of  $\{\tilde{x}_k\}$ , together with  $t_k \in (0, \gamma]$ , it follows that both  $\{z_k\}$  and  $\{\hat{z}_k\}$  are bounded. Hence, there exists a compact set  $\Lambda \subset \mathbb{R}^n$  such that  $\{x_k, \hat{z}_k\} \subseteq \Lambda$ . By Assumption 3.2 and [6, Theorem 2.1.6], there exists  $L > 0$  such that

$$\|\mathfrak{E}(x) - \mathfrak{E}(y)\| \leq L \|x - y\|, \quad \forall x, y \in \Lambda. \quad (3.21)$$

If  $t_k = \gamma$ , then the desired result follows immediately. Otherwise,  $\hat{t}_k$  does not satisfy the line search (2.5), i.e.,

$$-\mathfrak{E}(\hat{z}_k)^\top d_k < \sigma \hat{t}_k \max\{\kappa_1 \|\mathfrak{E}(\hat{z}_k)\| \|d_k\|, \kappa_2 \|d_k\|^2\}. \quad (3.22)$$

We distinguish two possible situations.

- (i-a) Suppose that the maximum is attained at the first term, i.e.,  $\max\{\varkappa_1 \|\mathfrak{E}(\hat{z}_k)\| \|d_k\|, \varkappa_2 \|d_k\|^2\} = \varkappa_1 \|\mathfrak{E}(\hat{z}_k)\| \|d_k\|$ . Then, by (3.22),

$$-\mathfrak{E}(\hat{z}_k)^\top d_k < \sigma \hat{t}_k \varkappa_1 \|\mathfrak{E}(\hat{z}_k)\| \|d_k\|.$$

Combining this with  $t_k \in (0, \gamma]$ , (2.8), (3.12), (3.20), and (3.21), we derive

$$\begin{aligned} (1-p)\xi_{\min}\epsilon^\lambda \|d_k\|^2 &< (1-p)\nu_k \|d_k\|^2 \leq -\mathfrak{E}(\tilde{x}_k)^\top d_k = (\mathfrak{E}(\hat{z}_k) - \mathfrak{E}(\tilde{x}_k))^\top d_k - \mathfrak{E}(\hat{z}_k)^\top d_k \\ &< L\hat{t}_k \|d_k\|^2 + \sigma \hat{t}_k \varkappa_1 \|\mathfrak{E}(\hat{z}_k)\| \|d_k\| \\ &= L\hat{t}_k \|d_k\|^2 + \sigma \hat{t}_k \varkappa_1 (\|\mathfrak{E}(\hat{z}_k)\| - \|\mathfrak{E}(\tilde{x}_k)\| + \|\mathfrak{E}(\tilde{x}_k)\|) \|d_k\| \\ &\leq L\hat{t}_k \|d_k\|^2 + \sigma \hat{t}_k \varkappa_1 L\gamma \|d_k\|^2 + \sigma \hat{t}_k \varkappa_1 Q_3 \|d_k\|^2 \\ &= (L + \sigma \varkappa_1 L\gamma + \sigma \varkappa_1 Q_3) \tau^{-1} t_k \|d_k\|^2, \end{aligned}$$

Hence,  $t_k > \frac{(1-p)\tau\xi_{\min}\epsilon^\lambda}{L + \sigma \varkappa_1 L\gamma + \sigma \varkappa_1 Q_3}$ .

- (i-b) Alternatively, if the maximum is realized by the second term, namely,  $\max\{\varkappa_1 \|\mathfrak{E}(\hat{z}_k)\| \|d_k\|, \varkappa_2 \|d_k\|^2\} = \varkappa_2 \|d_k\|^2$ , then (3.22) yields

$$-\mathfrak{E}(\hat{z}_k)^\top d_k < \sigma \hat{t}_k \varkappa_2 \|d_k\|^2.$$

Combining this with (2.8), (3.20), and (3.21), we obtain

$$\begin{aligned} (1-p)\xi_{\min}\epsilon^\lambda \|d_k\|^2 &< (1-p)\nu_k \|d_k\|^2 \leq -\mathfrak{E}(\tilde{x}_k)^\top d_k = (\mathfrak{E}(\hat{z}_k) - \mathfrak{E}(\tilde{x}_k))^\top d_k - \mathfrak{E}(\hat{z}_k)^\top d_k \\ &< L\hat{t}_k \|d_k\|^2 + \sigma \hat{t}_k \varkappa_2 \|d_k\|^2 \\ &= (L + \sigma \varkappa_2) \tau^{-1} t_k \|d_k\|^2, \end{aligned}$$

which implies  $t_k > \frac{(1-p)\tau\xi_{\min}\epsilon^\lambda}{L + \sigma \varkappa_2}$ .

In both cases, we conclude that there exists  $Q_6 > 0$  such that  $t_k \geq Q_6$  for all  $k \geq 0$ .

- (ii) Combining (3.19), part (i), and  $z_i = \tilde{x}_i + t_i d_i$ , we obtain

$$Q_6^4 \sum_{i=0}^{k-1} \|d_i\|^2 \leq \sum_{i=0}^{k-1} t_i^4 \|d_i\|^2 = \sum_{i=0}^{k-1} t_i^2 \|\tilde{x}_i - z_i\|^2 \leq Q_5,$$

which yields  $\sum_{i=0}^{k-1} \|d_i\|^2 \leq Q_5/Q_6^4 := Q_7$  for all  $k \geq 1$ . This completes the proof.  $\square$

Based on Lemma 3.3, we establish an iteration-complexity bound for computing an  $\epsilon$ -approximate solution to problem (1.1).

**Theorem 3.4** Suppose that the assumptions of Lemma 3.3 hold. Then the proposed aiNP method requires at most  $\left\lceil \frac{Q_3^2 Q_7}{\epsilon^2} \right\rceil$  iterations to compute an  $\epsilon$ -approximate solution to (1.1), where  $\lceil \cdot \rceil$  denotes the ceiling function.

**Proof.** Let  $N := \left\lceil \frac{Q_3^2 Q_7}{\epsilon^2} \right\rceil$ . Then

$$N\epsilon^2 = \epsilon^2 \left\lceil \frac{Q_3^2 Q_7}{\epsilon^2} \right\rceil \geq Q_3^2 Q_7. \quad (3.23)$$

Suppose, by contradiction, that the proposed aiNP method does not terminate within  $N$  iterations. Then  $\|\mathfrak{E}(\tilde{x}_k)\| > \epsilon$  for all  $k = 0, 1, \dots, N-1$ . Combining this with (3.12) and Lemma 3.3(ii), we obtain  $\frac{N\epsilon^2}{Q_3^2} < \frac{1}{Q_3^2} \sum_{i=0}^{N-1} \|\mathfrak{E}(\tilde{x}_i)\|^2 \leq \sum_{i=0}^{N-1} \|d_i\|^2 \leq Q_7$ . This implies that  $N\epsilon^2 < Q_3^2 Q_7$ , which contradicts (3.23). Therefore, the proof is complete.  $\square$

## 4 Numerical Experiments

In this section, we present a series of numerical experiments to assess the performance of the proposed aiNP method on standard SoNE benchmark problems and regularized logistic regression tasks. All experiments were conducted in MATLAB R2025a on a 64-bit workstation equipped with an Intel(R) Core(TM) Ultra 9 275HX processor operating at 2.70 GHz, 32 GB of RAM, and running the Windows 11 operating system.

Throughout this section, we compare the proposed aiNP method with two existing inertial methods, namely IRNM [31, Algorithm 2] and NM [31, Algorithm 2]. The parameter settings of these methods are adopted from their original references. All methods are terminated whenever one of the following conditions is satisfied:  $\|\mathfrak{E}(x_k)\| \leq 10^{-6}$ ,  $\|\mathfrak{E}(\tilde{x}_k)\| \leq 10^{-6}$ ,  $\|\mathfrak{E}(z_k)\| \leq 10^{-6}$ ,  $\|d_k\| \leq 10^{-7}$ , or  $\text{Itr} > 500$ , where “Itr” denotes the iteration number. If  $\text{Itr} > 500$ , the method is deemed to have failed and the corresponding result is reported as “NaN”. In the inner iteration, the linear conjugate gradient (CG) method is employed to compute the search direction in [3, Algorithm 5.2], with initialization similar to that in [32]. Following [8], we choose  $\nu_k = \min\{\xi_k \|\mathfrak{E}(\tilde{x}_k)\|^\lambda, \zeta\}$  with  $\zeta > 0$ . Unless otherwise specified, the parameters of aiNP are set to  $\alpha = 0.001$ ,  $p = 0.9$ ,  $\tau = 0.5$ ,  $\lambda = 1$ ,  $\sigma = 0.01$ ,  $\gamma = 1$ ,  $\varkappa_1 = 10^{-5}$ ,  $\varkappa_2 = 0.007$ ,  $\rho = 1.765$ ,  $\xi_k \equiv 0.34$  and  $\zeta = 0.001$ . For the iNP variant without the inertial step, we set  $\alpha_k \equiv 0$  and  $\rho = 1.6$ .

### 4.1. Numerical Experiments on Standard SoNE Test Problems

In this subsection, we employ several benchmark instances to evaluate the numerical performance of the proposed aiNP method. The four standard SoNE test problems considered are listed below, where  $\mathfrak{E}$  in (1.1) is defined by  $\mathfrak{E}(x) = (\mathfrak{E}_1(x), \mathfrak{E}_2(x), \dots, \mathfrak{E}_n(x))^\top$ . For each method and each test problem, the initial point  $x_0$  is generated by the MATLAB command “rand( $n, 1$ ) \* 11 - 1”. In the experiments, the problem dimensions are chosen as  $n = 100, 300, 500, 800$  and 1000. All reported results are averaged over five independent runs.

**Problem 1.** [25]. Set  $\mathfrak{E}_i(x) = e^{x_i} - 1$  for  $i = 1, 2, \dots, n$ .

**Problem 2.** Modified from [7]. Set  $\mathfrak{E}_i(x) = 2x_i - \sin(|x_i - 1|)$  for  $i = 1, 2, \dots, n$ .

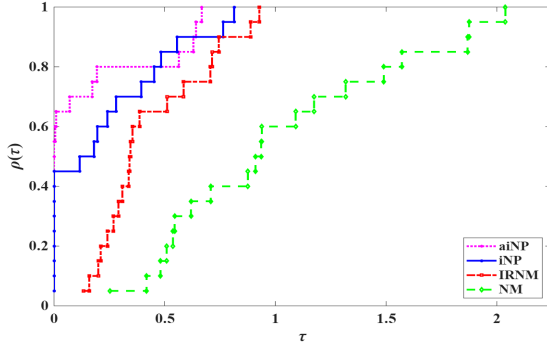


Figure 1: Performance profiles on Time

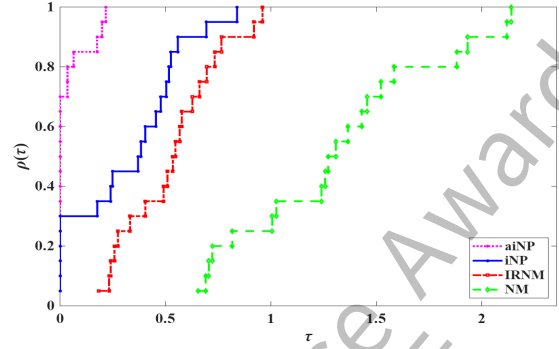


Figure 2: Performance profiles on Itr

**Problem 3.** [29]

$$\begin{aligned}\mathfrak{E}_1(x) &= e^{x_1} - 1, \\ \mathfrak{E}_i(x) &= e^{x_i} - x_i - 1 \text{ for } i = 2, 3, \dots, n.\end{aligned}$$

**Problem 4.** [5]  $\mathfrak{E}(x) = Ax + f(x)$ , where  $f(x) = (2e^{x_1} - 1, 3e^{x_2} - 1, \dots, 3e^{x_{n-1}} - 1, 2e^{x_n} - 1)^\top$  and

$$A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}.$$

The numerical results are reported in Table 1, where “ $n$ ” denotes the problem dimension, “Itr” represents the number of iterations, “Intr” denotes the number of inner CG iterations, “NF” denotes the number of function evaluations, “Time” represents the computing time, and “ $\|\mathfrak{E}^*\|$ ” denotes the final residual norm  $\|\mathfrak{E}(\cdot)\|$  at termination.

Following the methodology proposed by Dolan and Moré [10], we employ performance profiles to assess the numerical performance of the proposed methods. These profiles provide a comprehensive statistical comparison among the competing methods, where the uppermost curve indicates the method exhibiting the best overall performance. Figs. 1–2 present the performance profiles of “Time” and “Itr”.

The results reported in Table 1 and Figs. 1–2 demonstrate that the proposed aiNP method, together with its corresponding iNP variant, significantly outperforms the other two methods in terms of computational time, iteration number, and solution quality. These results further confirm the effectiveness and robustness of the proposed approaches. Moreover, they clearly indicate that the inertial-relaxed technique is capable of substantially accelerating the convergence of the Newton-type projection method.

Table 1: Numerical comparison of Problems 1–4

| Problem | $n$  | aiNP                                    | iNP                                     | IRNM                                    | NM                                      |
|---------|------|-----------------------------------------|-----------------------------------------|-----------------------------------------|-----------------------------------------|
|         |      | Itr/Initr/NF/Time/ $\ \mathfrak{E}^*\ $ | Itr/Initr/NF/Time/ $\ \mathfrak{E}^*\ $ | Itr/Initr/NF/Time/ $\ \mathfrak{E}^*\ $ | Itr/Initr/NF/Time/ $\ \mathfrak{E}^*\ $ |
| 1       | 100  | 21.0/328.4/85.8/0.003/6.17e-07          | 24.8/352.0/113.6/0.002/6.89e-07         | 27.8/384.0/129.6/0.002/6.61e-07         | 42.2/468.2/138.6/0.002/8.30e-07         |
|         | 300  | 17.8/370.4/73.0/0.009/3.78e-07          | 24.4/393.2/110.4/0.009/6.81e-07         | 29.6/461.0/153.6/0.011/7.73e-07         | 65.6/635.6/285.8/0.018/7.39e-07         |
|         | 500  | 18.6/368.2/75.0/0.013/5.71e-07          | 27.4/419.8/137.0/0.018/7.62e-07         | 36.2/520.4/201.2/0.025/7.47e-07         | 80.8/696.2/397.4/0.049/6.71e-07         |
|         | 800  | 20.0/384.6/82.2/0.043/4.66e-07          | 35.8/479.2/187.6/0.076/7.51e-07         | 34.0/529.6/183.0/0.072/7.41e-07         | 76.4/738.4/364.2/0.159/6.67e-07         |
|         | 1000 | 20.4/392.0/89.0/0.070/4.99e-07          | 27.0/430.2/123.2/0.093/7.47e-07         | 38.6/580.8/217.2/0.130/6.39e-07         | 90.0/774.6/468.8/0.289/7.07e-07         |
| 2       | 100  | 10.0/37.6/45.4/0.001/4.09e-08           | 14.4/50.0/66.6/0.001/6.35e-07           | 16.2/48.4/75.6/0.001/6.69e-07           | 30.0/64.6/112.4/0.001/7.10e-07          |
|         | 300  | 11.8/45.8/53.2/0.003/4.00e-07           | 15.4/55.6/71.2/0.003/7.00e-07           | 16.8/57.0/78.8/0.003/6.48e-07           | 30.4/68.4/115.0/0.005/6.07e-07          |
|         | 500  | 11.2/47.8/51.0/0.005/2.26e-07           | 15.6/56.8/72.4/0.007/3.85e-07           | 16.6/57.0/77.8/0.008/6.13e-07           | 30.2/71.0/114.2/0.013/7.50e-07          |
|         | 800  | 11.0/47.4/49.4/0.021/3.45e-07           | 15.6/56.6/72.4/0.031/6.77e-07           | 17.4/58.2/81.2/0.034/5.64e-07           | 30.2/69.4/114.6/0.058/8.18e-07          |
|         | 1000 | 11.0/47.0/48.4/0.032/6.11e-07           | 17.8/64.4/81.0/0.055/5.13e-07           | 17.0/60.0/79.2/0.053/9.24e-07           | 31.6/73.4/118.8/0.096/9.00e-07          |
| 3       | 100  | 13.0/232.4/45.2/0.001/4.93e-07          | 16.8/248.0/60.2/0.001/6.92e-07          | 19.0/257.0/70.4/0.001/7.78e-07          | 31.4/309.2/85.6/0.001/7.29e-07          |
|         | 300  | 13.0/272.0/44.4/0.005/5.05e-07          | 18.6/296.0/67.4/0.006/7.80e-07          | 19.4/303.4/72.0/0.006/7.94e-07          | 32.2/371.8/88.4/0.008/7.29e-07          |
|         | 500  | 13.8/277.6/48.2/0.009/4.50e-07          | 16.4/293.8/57.4/0.010/5.28e-07          | 19.4/305.4/72.0/0.011/7.10e-07          | 32.6/379.6/89.8/0.017/7.04e-07          |
|         | 800  | 13.8/279.2/48.4/0.030/3.52e-07          | 15.6/293.6/53.4/0.034/8.09e-07          | 20.0/308.2/74.0/0.043/7.08e-07          | 33.0/383.8/91.0/0.068/6.41e-07          |
|         | 1000 | 16.6/282.8/59.8/0.057/5.76e-07          | 16.2/292.2/55.2/0.057/6.58e-07          | 19.6/305.2/72.6/0.066/7.08e-07          | 33.0/379.8/91.0/0.108/6.81e-07          |
| 4       | 100  | 31.4/517.6/108.2/0.003/7.88e-07         | 27.8/500.8/96.4/0.002/8.19e-07          | 35.0/556.2/132.8/0.002/6.84e-07         | 49.0/700.2/135.4/0.003/8.33e-07         |
|         | 300  | 35.6/615.2/123.2/0.024/8.18e-07         | 31.0/589.8/109.8/0.021/8.62e-07         | 35.2/623.6/135.0/0.023/7.97e-07         | 48.8/801.8/137.2/0.028/7.72e-07         |
|         | 500  | 32.0/597.6/111.2/0.087/8.73e-07         | 30.6/604.0/108.0/0.083/8.87e-07         | 36.6/647.0/140.8/0.105/6.80e-07         | 49.4/828.4/139.0/0.121/8.62e-07         |
|         | 800  | 32.8/618.0/114.6/0.310/8.56e-07         | 32.0/623.6/113.8/0.308/7.80e-07         | 37.8/675.6/145.4/0.381/7.72e-07         | 52.8/874.6/149.0/0.438/8.16e-07         |
|         | 1000 | 37.2/642.8/128.6/0.545/8.63e-07         | 32.0/621.4/114.2/0.483/8.75e-07         | 37.6/657.6/144.8/0.592/7.10e-07         | 52.2/859.6/147.6/0.675/7.62e-07         |

## 4.2. Numerical Experiments on Regularized Logistic Regression Problems

In this subsection, we consider the following Tikhonov-regularized logistic regression problem [19]:

$$\min_{x \in \mathbb{R}^n} g(x) = \frac{1}{N} \sum_{i=1}^N \log \left( 1 + \exp \left( -c_i d_i^\top x \right) \right) + \frac{\eta}{2} \|x\|^2, \quad (4.1)$$

where  $(d_i, c_i) \in \mathbb{R}^n \times \{-1, 1\}$  ( $i = 1, 2, \dots, N$ ) are generated from a given dataset, and  $\eta > 0$  denotes the regularization parameter. In all experiments, we set  $\eta = 0.02$ . As shown in [11], the unique optimal solution of problem (4.1) can be characterized by the following first-order optimality condition:

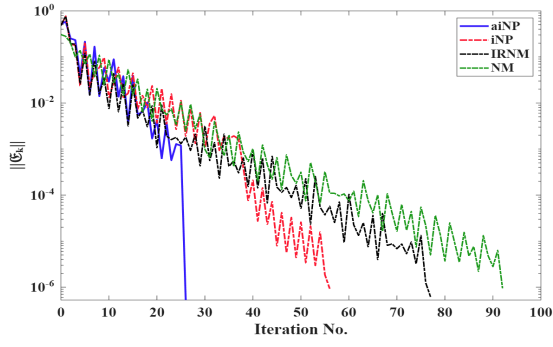
$$\mathfrak{E}(x) := \nabla g(x) = \frac{1}{N} \sum_{i=1}^N \frac{-c_i \exp(-c_i d_i^\top x) d_i}{1 + \exp(-c_i d_i^\top x)} + \eta x = 0. \quad (4.2)$$

Therefore, to obtain the solution of problem (4.1), we employ the proposed aiNP method to solve the nonlinear system (4.2). By direct computation, we obtain

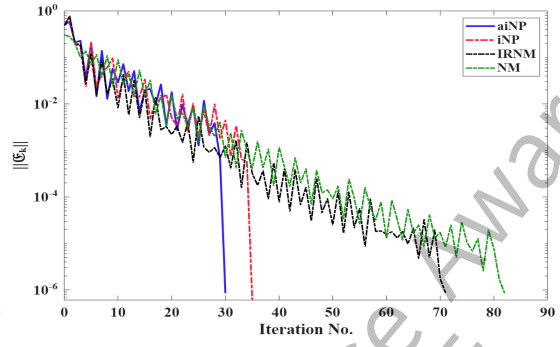
$$J(x) := \nabla^2 g(x) = \frac{1}{N} \sum_{i=1}^N \frac{\exp(-c_i d_i^\top x) d_i d_i^\top}{(1 + \exp(-c_i d_i^\top x))^2} + \eta I.$$

In this subsection, we set  $\rho = 1.6$  for the aiNP method and  $\rho = 1.5$  for the iNP method. The parameter settings of all competing methods, including the stopping criteria, are chosen according to their original references. Moreover, the origin is selected as the initial point for all experiments. The test datasets are taken from [4]. The statistics of the employed datasets are reported in Table 2, while the corresponding numerical results are presented in Table 3. In addition, the evolution of  $\|\mathfrak{E}\|$  with respect to the iteration number is illustrated in Fig. 3, and the corresponding bar charts are displayed in Figs. 4–5.

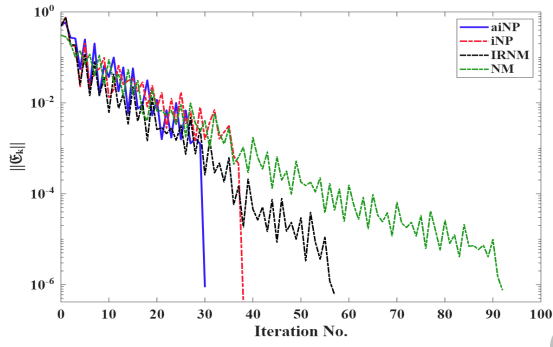
The results demonstrate that the proposed aiNP method converges more rapidly toward the solution than the other competing methods. Although IRNM exhibits competitive performance in certain cases, the aiNP method achieves the best overall performance in most instances. Furthermore, the iNP method without the inertial step also performs competitively, which further confirms the effectiveness of the proposed framework. Combining the observations from both the tables and figures with those obtained in the previous subsection, it can be concluded that the proposed aiNP method possesses strong efficiency and robustness for solving the considered problems.



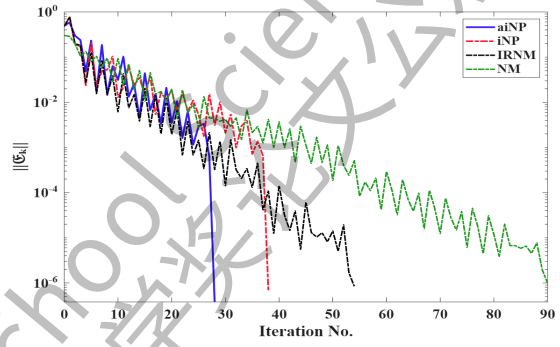
(a) a1a.t



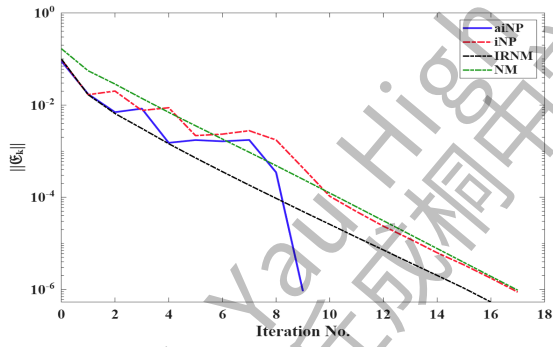
(b) a5a.t



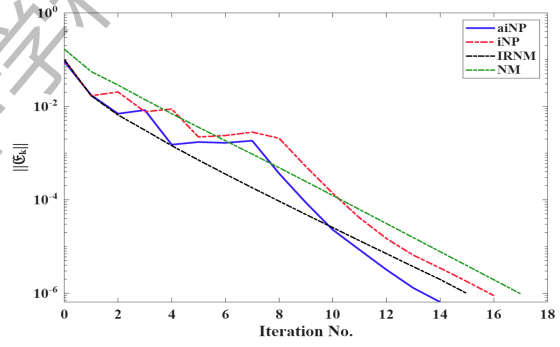
(c) a6a.t



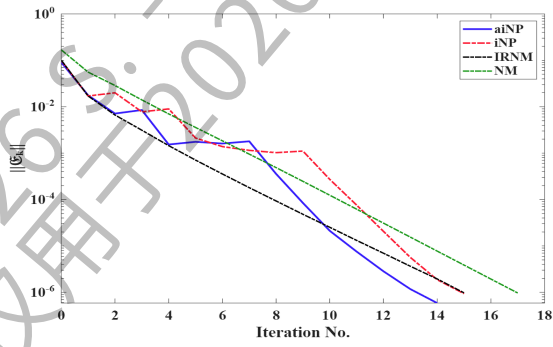
(d) a7a.t



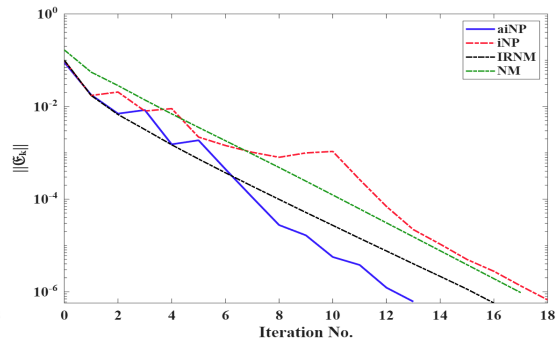
(e) w1a.t



(f) w5a.t



(g) w6a.t



(h) w8a.t

Figure 3: Change of  $\|\mathcal{E}\|$  with respect to the iteration.

Table 2: Statistics of the LIBSVM test datasets

| No. | Dataset | Data points N | Variables n |
|-----|---------|---------------|-------------|
| 1   | a1a.t   | 30,956        | 123         |
| 2   | a5a.t   | 26,147        | 123         |
| 3   | a6a.t   | 21,341        | 123         |
| 4   | a7a.t   | 16,461        | 123         |
| 5   | w1a.t   | 47,272        | 300         |
| 6   | w5a.t   | 39,861        | 300         |
| 7   | w6a.t   | 32,561        | 300         |
| 8   | w8a.t   | 14,951        | 300         |

Table 3: Numerical results on the LIBSVM test instances

| Dataset | aiNP                                 | iNP                                  | IRNM                                 | NM                                   |
|---------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
|         | Itr/Initr/Time/ $\ \mathfrak{E}^*\ $ | Itr/Initr/Time/ $\ \mathfrak{E}^*\ $ | Itr/Initr/Time/ $\ \mathfrak{E}^*\ $ | Itr/Initr/Time/ $\ \mathfrak{E}^*\ $ |
| a1a.t   | 27/118/1.164/5.253e-07               | 57/303/2.472/9.835e-07               | 70/259/3.122/8.152e-07               | 93/333/3.754/8.042e-07               |
| a5a.t   | 31/128/1.058/8.510e-07               | 36/153/1.223/6.065e-07               | 67/256/2.410/8.831e-07               | 90/341/2.911/9.132e-07               |
| a6a.t   | 31/134/0.874/8.797e-07               | 39/159/1.109/4.168e-07               | 63/253/1.835/6.317e-07               | 93/341/2.431/8.073e-07               |
| a7a.t   | 29/124/0.606/3.801e-07               | 39/163/0.807/6.886e-07               | 56/226/1.213/5.903e-07               | 92/323/1.801/7.438e-07               |
| w1a.t   | 10/39/0.542/9.194e-07                | 18/78/1.006/8.784e-07                | 17/72/0.968/5.339e-07                | 18/73/0.897/9.633e-07                |
| w5a.t   | 15/72/0.694/6.496e-07                | 17/71/0.789/8.907e-07                | 16/66/0.762/9.883e-07                | 18/71/0.763/9.734e-07                |
| w6a.t   | 15/72/0.559/5.879e-07                | 16/68/0.594/9.368e-07                | 16/66/0.616/9.738e-07                | 18/73/0.629/9.676e-07                |
| w8a.t   | 14/67/0.271/6.126e-07                | 19/88/0.360/6.687e-07                | 17/72/0.319/5.701e-07                | 18/74/0.308/9.584e-07                |

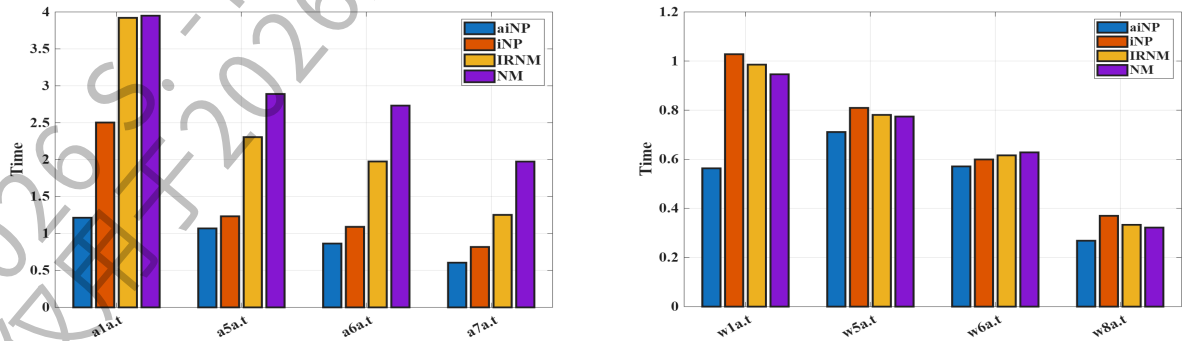


Figure 4: Comparison of Methods in Terms of Time

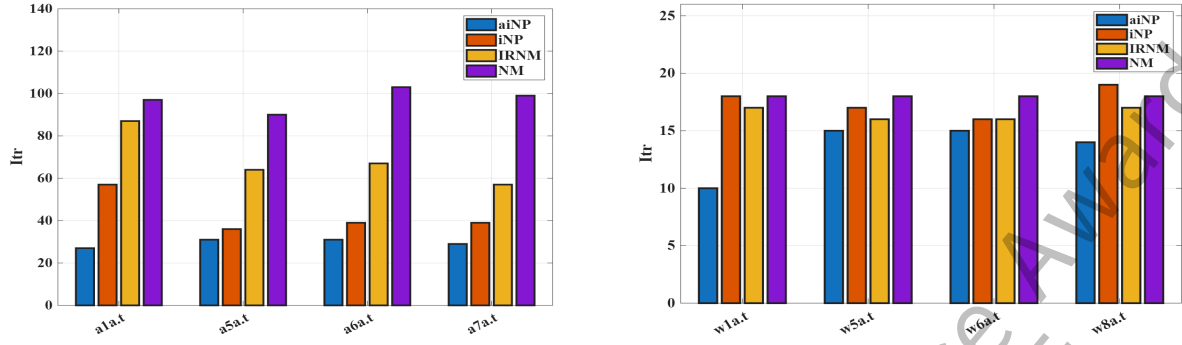


Figure 5: Comparison of Methods in Terms of Itr

## 5 Conclusions

In this paper, we propose an accelerated inexact Newton-type projection method for solving quasar-monotone SoNE (1.1). Under mild assumptions, we establish the global convergence of the proposed method and derive an iteration-complexity bound for computing an  $\epsilon$ -approximate solution. Numerical experiments on standard nonlinear equations and Tikhonov-regularized logistic regression demonstrate the effectiveness and efficiency of the proposed method.

## References

- [1] Alves, M.M., Eckstein, J., Geremia, M., Melo, J.G.: Relative-error inertial-relaxed inexact versions of Douglas-Rachford and ADMM splitting algorithms. *Comput. Optim. Appl.* 75, 389–422 (2020)
- [2] Bagirov, A., Karmitsa, N., Mäkelä, M.M.: *Introduction to Nonsmooth Optimization: Theory, Practice and Software*. Springer, New York (2014)
- [3] Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3(1), 1–122 (2010)
- [4] Chang C.C., Lin C.J.: LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2(3), 1–27 (2011)
- [5] Chen, Z., Cheng, W., Li, X.: A global convergent quasi-Newton method for systems of monotone equations. *J. Appl. Math. Comput.* 44, 455–465 (2014)
- [6] Cobzas, S., Miculescu, R., Nicolae, A.: Lipschitz Functions. *Lect. Notes Math.* 2241, 365–556 (2019)
- [7] Cruz, W.L.: A spectral algorithm for large-scale systems of nonlinear monotone equations. *Numer. Algorithms* 76, 1109–1130 (2017)
- [8] Dan H., Yamashita N., Fukushima M.: Convergence properties of the inexact Levenberg-Marquardt method under local error bound conditions. *Optim. Methods Softw.* 17(4), 605–626 (2002)
- [9] Dang, C.D., Lan, G.: On the convergence properties of non-Euclidean extragradient methods for variational inequalities with generalized monotone operators. *Comput. Optim. Appl.* 60, 277–310 (2015)
- [10] Dolan E.D., Moré J.J.: Benchmarking optimization software with performance profiles. *Math. Program.* 91(2), 201–213 (2002)
- [11] Jian, J.B., Yin, J.H., Tang, C.M., Han, D.L.: A family of inertial derivative-free projection methods for constrained nonlinear pseudo-monotone equations with applications. *Comput. Appl. Math.* 41, 309 (2022)
- [12] Jian, J.B., Ren, Z.W., Yin, J.H., Jiang, X.Z.: A family of hybrid acceleration DFPMs and applications for pseudo-monotone nonlinear equations with convex constraints. *J. Optim. Theory Appl.* 209, 26 (2026)
- [13] Levenberg, K.: A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* 2(2), 164–168 (1944)

- [14] Liu, M.R., Mroueh, Y., Ross, J., Zhang, W., Cui, X.D., Das, P., Yang, T.B.: Towards better understanding of adaptive gradient algorithms in generative adversarial nets. *ICLR*, 1–23 (2020)
- [15] Liu P.J., Li J.Y., Shao H., Wu T.: An accelerated derivative-free memoryless Davidon-Fletcher-Powell method and its iteration-complexity analysis. *J. Sci. Comput.* **105**(2), 43 (2025)
- [16] Liu, P.J., Shao, H., Yuan, Z.H., Zhou, J.H.: A family of inertial-based derivative-free projection methods with a correction step for constrained nonlinear equations and their applications. *Numer. Linear Algebra Appl.* **31**(2), e2533 (2024)
- [17] Liu, P.J., Li, L.H., Shao, H., Liu, M.X., Fan, J.X.: An inertial-type CG projection method with restart for pseudo-monotone costs with application to traffic assignment. *Netw. Spat. Econ.* **25**(1), 147–172 (2025)
- [18] Liu, P.J., Li, L.H., Shao, F., Li, K.: An accelerated hybrid three-term derivative-free projection method with applications. *Optim. Lett.* (2025)
- [19] Luo H.: Accelerated primal-dual methods for linearly constrained convex optimization problems. *arXiv:2109.12604* (2021)
- [20] Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **11**(2), 431–441 (1963)
- [21] Mertikopoulos, P., Zhou, Z.: Learning in games with continuous action sets and unknown payoff functions. *Math. Program. Ser. A* **173**, 465–507 (2019)
- [22] Nguyen, N.H., Trinh, T.Q., Vuong, P.T.: The projected heavy-ball method for quasar-monotone variational inequalities and quasar-convex optimization. *J. Optim. Theory Appl.* **208**, 83 (2026)
- [23] Solodov, M.V., Svaiter, B.F.: A globally convergent inexact Newton method for systems of monotone equations. In: Fukushima, M., Qi, L. (eds.) *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*. *Appl. Optim.*, pp. 355–369. Springer, Dordrecht (1998)
- [24] Ullah, N., Sabi'u, J., Shah, A.: A derivative-free scaling memoryless Broyden-Fletcher-Goldfarb-Shanno method for solving a system of monotone nonlinear equations. *Numer. Linear Algebra Appl.* **28**(5), e2374 (2021)
- [25] Wang C.W., Wang Y.J., Xu C.L.: A projection method for a system of nonlinear monotone equations with convex constraints. *Math. Methods Oper. Res.* **66**, 33–46 (2007)
- [26] Xiao, Y.H., Wang, Q.Y., Hu, Q.J.: Non-smooth equations based method for  $\ell_1$ -norm problems with applications to compressed sensing. *Nonlinear Anal. Theory Methods Appl.* **74**(11), 3570–3577 (2011)

- [27] Xiao, X., Li, Y., Wen, Z., Zhang, L.: A regularized semismooth Newton method with projection steps for composite convex programs. *J. Sci. Comput.* 76, 364–389 (2018)
- [28] Yin, J.H., Huang, X.Y., Jian, J.B., Huang, Q.X., Gao, P.T.: A family of derivative-free projection methods with an adaptive Armijo line search for solving nonlinear monotone equations. *J. Comput. Appl. Math.* 491, 117995 (2027)
- [29] Yin, J.H., Jian, J.B., Jiang, X.Z., Wu, X.D.: A family of inertial-relaxed DFPM-based algorithms for solving large-scale monotone nonlinear equations with application to sparse signal restoration. *J. Comput. Appl. Math.* 419, 114674 (2023)
- [30] Yin, J.H., Huang, Q.X., Jian, J.B., Han, D.L.: An inertial-relaxed three-term conjugate gradient projection method for large-scale unconstrained nonlinear pseudo-monotone equations with applications. *J. Comput. Math.* (2025)
- [31] Yin J.H., Li Y.B., Jian D., Jiang X.Z.: An inertial-relaxed Newton-type projection method for solving nonlinear equations with applications. *Numer. Linear Algebra Appl.* 33(1), e70067 (2026)
- [32] Yin, J.H., Jian, J.B., Ma, G.D.: A modified inexact Levenberg-Marquardt method with the descent property for solving nonlinear equations. *Comput. Optim. Appl.* 87(1), 289–322 (2024)
- [33] Zhang, L., Zhou, W.: Spectral gradient projection method for solving nonlinear monotone equations. *J. Comput. Appl. Math.* 196(2), 478–484 (2006)
- [34] Zhang, B., Zhu, Z.: A modified quasi-Newton diagonal update algorithm for total variation denoising problems and nonlinear monotone equations with applications in compressive sensing. *Numer. Linear Algebra Appl.* 22, 500–522 (2015)
- [35] Zhou, G., Toh, K.C.: Superlinear convergence of a Newton-type algorithm for monotone equations. *J. Optim. Theory Appl.* 125, 205–221 (2005)
- [36] Zhou, W.J., Li, D.H.: Limited memory BFGS method for nonlinear monotone equations. *J. Comput. Math.* 25, 89–96 (2007)
- [37] Zhou, W.J., Li, D.H.: A globally convergent BFGS method for nonlinear monotone equations without any merit functions. *Math. Comput.* 77, 2231–2240 (2008)

## Acknowledgments

First and foremost, I would like to express my sincere gratitude to Prof. Ting Wu and Associate Prof. Pengjie Liu for their generous and unpaid guidance throughout this research project. Their valuable advice and continuous encouragement played an important role in the completion of this work. After discussing possible research directions with my advisors, I decided to investigate an accelerated inexact Newton-type projection method for quasar-monotone nonlinear equations. Under their guidance, I reviewed the relevant literature and gradually identified the main research question and technical route. Prof. Ting Wu mainly provided guidance on algorithm design and numerical experiments, while Associate Prof. Pengjie Liu mainly offered suggestions concerning the theoretical analysis and interpretation of the numerical results. Their guidance helped me develop the initial idea into a complete research project.

I was responsible for the main research work throughout the project. In the algorithm-design stage, I studied Newton-type and projection-based methods and developed the proposed accelerated inexact Newton-type projection method. I was responsible for deriving the iterative scheme, implementing the algorithm, and testing its computational performance. In the theoretical-analysis stage, I worked on the global convergence and iteration-complexity analysis of the proposed method under the quasar-monotone setting. Through repeated discussions with my advisors and careful revision of the proofs, I gradually resolved these difficulties.

For the numerical study, I independently implemented the relevant algorithms, prepared the test problems, conducted the experiments, collected and analyzed the computational results, and prepared the tables and figures. The experiments included standard systems of nonlinear equations and regularized logistic regression problems. During this stage, I encountered difficulties such as parameter sensitivity, unstable computational behavior on some test instances, and relatively high computational costs. I addressed these issues by checking the implementation, conducting additional tests, adjusting the experimental settings, and discussing the results with my advisors.

I also undertook the main responsibility for writing the research report. I prepared the initial drafts of all major sections, including the introduction, algorithm, convergence analysis, numerical experiments, and conclusions. Prof. Ting Wu and Associate Prof. Pengjie Liu reviewed the manuscript and provided suggestions concerning the mathematical arguments, logical structure, numerical analysis, and academic writing. I subsequently revised the manuscript based on their comments. Therefore, although the research was conducted under their academic guidance, the data processing, algorithm implementation, theoretical derivations, numerical experiments, and initial manuscript preparation were primarily completed by me.

Finally, I would like to thank my advisors once again for their patient guidance and encouragement throughout this research. I am also grateful to my family for their continuous support during the competition and throughout my studies. This research experience has taught me that completing a mathematical research project requires not only a good idea, but also careful theoretical analysis, reliable computational experiments, repeated verification, and clear academic writing. I am deeply grateful for the opportunity to undertake this research and for everything I have learned from the process.