A New Adaptive Multi-step Levenberg-Marquardt Method and its Improved Convergence Results

Meilan Zeng*, Junchao Zhou

***Abstract*—In this paper, we propose an adaptive multi-step Levenberg-Marquardt (LM) method with a new parameter $\lambda_k = \mu_k \|G_k^T F_k\|^\delta, \delta \in (0,1]$ for solving nonlinear equations and improve convergence results. The global and local convergence theories are established. The local superlinear convergence is proved under the Hölderian local error bound condition, which is weaker than the local error bound. Numerical experiments verify the convergence of our algorithm for singular problems that satisfy the Hölderian local error bound condition.**

***Index Terms*—adaptive multi-step LM method, nonlinear equations, Hölderian local error bound.**

I. Introduction

CONSIDER the following systems of nonlinear equations

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

where $F : R^n \to R^n$ is continuously differentiable. In this paper, we suppose the solution set of (1), represented by X^*, is nonempty. In every situation, $\|\cdot\|$ indicates the Frobenius norm. Let $\phi(x) = \frac{1}{2}\|F(x)\|^2$, then the nonlinear equations (1) can be reformulated as a nonlinear least squares problem

$$\min_{x \in R^n} \phi(x). \tag{2}$$

There are many classical methods [4], [10], [11], [13], [14], [19], [21], [23] for solving nonlinear equations and nonlinear least squares problems. This paper is devoted to the LM methods for solving (1). The LM methods include single-step methods and multi-step methods.

A. Related Works

The single-step LM methods [12], [15] compute a trial step d_k by

$$d_k = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k, \tag{3}$$

in which the LM parameter $\lambda_k \geq 0$ is adjusted iteratively. The single-step LM methods have the quadratic convergence when $J(x)$ is nonsingular at the solution and Lipschitz continuous. Yamashita and Fukushima [20] demonstrated the quadratic convergence of the single-step LM method, if the LM parameter is chosen as $\lambda_k = \|F_k\|^2$ under the following local error bound condition (cf. [20]).

$$c_0 \cdot dist(x, X^*) \leq \|F(x)\|, \forall x \in N(x^*), \tag{4}$$

where c_0 is a positive constant, $dist(x, X^*)$ is the distance from x to X^* and $N(x^*)$ is some neighbourhood of $x^* \in X^*$.

Recently, the single-step LM methods have been further developed (see [1], [5], [6], [8], [10]). Behling et al. [3] proposed a modified LM method with the LM parameter $\lambda_k = \|J_k^T F_k\|^\delta, \delta \in (0,1]$ for solving the nonzero-residue nonlinear least-squares problems. To establish the local convergence, Behling et al. put forward an error bound condition based on the gradient of the nonlinear least-squares function as below (cf. [3]).

$$c_1 \cdot dist(x, X^*) \leq \|J(x)^T F(x)\|, \forall x \in N(x^*), \tag{5}$$

where c_1 is a positive constant, both $dist(x, X^*)$ and $N(x^*)$ have the same meaning as (4).

In order to save Jacobian evaluations, Fan et al. proposed the multi-step LM method [7] and the adaptive multi-step LM method [9] with the parameter $\lambda_k = \mu_k \|F_k\|^\delta, \delta \in [1,2]$. The adaptive technique automatically determines whether an iteration should use the Jacobian matrix at the current iterate to compute a LM step or use the latest evaluated Jacobian matrix for an approximate LM step. The adaptive multi-step LM method achieves the superlinear convergence under the local error bound condition (4).

B. Motivation

Some nonlinear equations fulfill the following Hölderian local error bound condition while may not meet the above local error bound condition (5).

Definition 1.1: We say $\|J(x)^T F(x)\|$ provides a Hölderian local error bound with order $\gamma \in (0,1]$ in some neighbourhood of $x^* \in X^*$, if there exists a constant $c_2 > 0$ such that

$$c_2 \cdot dist(x, X^*) \leq \|J(x)^T F(x)\|^\gamma, \forall x \in N(x^*), \tag{6}$$

where c_2 is a positive constant, $dist(x, X^*)$ and $N(x^*)$ have the same meaning as (4).

For example, Powell singular function ([16]):

$$F(x_1, x_2, x_3, x_4) = (x_1 + 10x_2, \sqrt{5}(x_3 - x_4), (x_2 - 2x_3)^2, \sqrt{10}(x_1 - x_4)^2)^T. \tag{7}$$

Let $0_4 := (0,0,0,0)^T$. It can be seen that $X^* = \{0_4\}$ and $J(0_4)$ is singular. We consider the sequence $\{x_k\}$, where each term is given by $x_k = (0, 0, -\frac{1}{k}, -\frac{1}{k})^T$. Then $\{x_k\} \to 0_4$ and

$$dist(x_k, X^*) = \|x_k\| = \sqrt{2}/k = O(k^{-1}),$$

and $\|J(x_k)^T F(x_k)\| = O(k^{-3})$. The nonlinear function satisfies the Hölderian local error bound condition with order $\frac{1}{3}$ around the zero point, but does not satisfy the local error bound condition (5).

Manuscript received January 14, 2025; revised June 3, 2025. This work was supported by the National Natural Science Foundation of China (No. 12301667) and the Natural Science Foundation of Hubei Province of China (No. 2021CFB275).

M.L. Zeng is a professor of School of Mathematics and Statistics, Hubei Engineering University, Xiaogan, 432000, China (Corresponding author, e-mail: zml571@126.com).

J.C. Zhou is an associate professor of School of Mathematics and Statistics, Hubei Engineering University, Xiaogan, 432000, China (e-mail: zhoujunchao@hbeu.edu.cn).

This paper is devoted to proposing a new adaptive multi-step LM method with the parameter $\lambda_k = \mu_k \|G_k^T F_k\|^\delta, \delta \in (0, 1]$. Local superlinear convergence is proved under the Hölderian local error bound condition (6), which is weaker than the local error bound (5). Numerical results demonstrate the validity of the proposed algorithm.

The structure of the remainder of this paper is as follows. Section 2 introduces the new adaptive multi-step LM algorithm and its global convergence. The local convergence theory is presented under the Hölderian local error bound condition in Section 3. Section 4 presents numerical results concerning test problems, and conclusions are summarized in Section 5.

II. The new adaptive multi-step LM algorithm and its global convergence

In this section, we give the new adaptive multi-step LM algorithm for solving nonlinear equations, and prove its global convergence.

We define the actual reduction as

$$Ared_k = \|F_k\|^2 - \|F(x_k + d_k)\|^2, \tag{8}$$

the predict reduction as

$$Pred_k = \|F_k\|^2 - \|F_k + G_k d_k\|^2. \tag{9}$$

and the radio of actual reduction over predict reduction as

$$r_k = \frac{Ared_k}{Pred_k}. \tag{10}$$

The new adaptive multi-step LM algorithm is presented as follows.

Algorithm 1.

Step 0. Given $x_1 \in R^n, m_1 > 1 > m_2 > 0, \mu_1 \geq \mu_0 > 0, \varepsilon > 0, \delta \in (0, 1], 0 < p_0 < p_1 < p_2 < p_3 < 1, t \geq 1$. Set $k := 1, s := 1, i := 1, k_i = 1, G_1 = J_1$.

Step 1. If $\|G_k^T F_k\| \leq \varepsilon$, stop. Otherwise set

$$\lambda_k = \mu_k \|G_k^T F_k\|^\delta, \tag{11}$$

solve

$$(G_k^T G_k + \lambda_k I)d = -G_k^T F_k \tag{12}$$

to obtain d_k.

Step 2. Compute $r_k = \frac{Ared_k}{Pred_k}$; set

$$x_{k+1} = \begin{cases} x_k + d_k, & if\ r_k \geq p_0, \\ x_k, & otherwise. \end{cases} \tag{13}$$

Step 3. Choose $\mu_{k+1}, \lambda_{k+1}, G_{k+1}$ as

$$\mu_{k+1} = \begin{cases} m_1 \mu_k, & if\ r_k < p_1, \\ \mu_k, & if\ r_k \in [p_1, p_3], \\ max\{m_2 \mu_k, \mu_0\}, & otherwise. \end{cases} \tag{14}$$

$$\lambda_{k+1} = \begin{cases} \lambda_k, & if\ r_k \geq p_2\ and\ s < t, \\ \mu_{k+1} \|G_{k+1}^T F_{k+1}\|^\delta, & otherwise, \end{cases} \tag{15}$$

$$G_{k+1} = \begin{cases} G_k, & if\ r_k \geq p_2\ and\ s < t, \\ J_{k+1}, & otherwise, \end{cases} \tag{16}$$

Step 4. Set $k := k + 1$. If $G_k = G_{k-1}$, set $s := s + 1$, otherwise set $s := 1, i := i + 1, k_i = k$. Go to Step 1.

Let $\bar{S} = \{k_i : i = 1, 2, \cdots\}$ be the set of numbers at which the Jacobians $J(x_{k_i}), i = 1, 2, \cdots$ are needed for computation during the iterations. Let

$$s_i = k_{i+1} - k_i.$$

From step 3 and step 4, it can be observed that $s_i \leq t$.

For any k, there exist k_i and $0 \leq t \leq s_i - 1$ such that

$$k = k_i + t.$$

Note that

$$G_{k_i} = G_{k_i+1} = \cdots = G_{k_i+s_i-1} = J_{k_i}, \tag{17}$$

hence the linear equations (12) can further be rewritten as

$$(G_{k_i}^T G_{k_i} + \lambda_{k_i} I)d = -G_{k_i}^T F_k \tag{18}$$

for $k = k_i, \cdots, k_i + s_i - 1$.

It is easy to verify that d_k also serves as the solution to the trust region subproblem

$$\begin{aligned} &\min_{d \in R^n} \ \|F_k + G_k d\|^2 \\ &s.t.\ \|d\| \leq \Delta_k := \|d_k\|. \end{aligned} \tag{19}$$

Applying Powells result [[17], Theorem 4] gives the following lemma.

Lemma 2.1: Let d_k be computed by (12), then the predicted reduction satisfies

$$Pred_k \geq \|G_k^T F_k\| \min\left\{\|d_k\|, \frac{\|G_k^T F_k\|}{\|G_k^T G_k\|}\right\} \tag{20}$$

for all k.

Lemma 2.1 plays a crucial role in proving the global convergence of Algorithm 1. We introduce the following assumptions to achieve this.

Assumption 2.1 (1) $J(x)$ is Lipschitz continuous, i.e., there exists a positive constant κ_h such that

$$\|J(x) - J(y)\| \leq \kappa_h \|x - y\|, \forall x, y \in R^n. \tag{21}$$

(2) $J(x)$ is bounded, i.e., there is a positive constant κ_b such that

$$\|J(x)\| \leq \kappa_b, \forall x \in R^n. \tag{22}$$

Then it follows from (21) that

$$\|F(y) - F(x) - J(x)(y - x)\| \leq \kappa_h \|y - x\|^2, \forall x, y \in R^n. \tag{23}$$

Theorem 2.1: Under the conditions of Assumption 2.1, the sequence $\{x_k\}$ generated by Algorithm 1 satisfies

$$\lim_{k \to \infty} \inf \|J_k^T F_k\| = 0. \tag{24}$$

Proof: Assume that (24) is not true. Consequently, there exist a positive constant τ and a constant $k_0 \in N$ such that

$$\|J_k^T F_k\| \geq \tau, \forall k \geq k_0. \tag{25}$$

Let us define the set of successful iterations as follows

$$S = \{k | r_k \geq p_0\}. \tag{26}$$

We consider S in two cases.

Case 1: S is finite. Then there exists a $\bar{k} \in N$ such that $r_k < p_0 < p_1 < p_2$ for all $k \geq \bar{k}$. Therefore, for $k \geq \bar{k}$,

$$G_k = J_k, \mu_k \to +\infty. \tag{27}$$

By (25), for $k \geq \bar{k}$,

$$\|G_k^T F_k\| = \|J_k^T F_k\| \geq \tau. \tag{28}$$

Owing to $\lambda_k = \mu_k \|G_k^T F_k\|^\delta$ and $\mu_k \to +\infty$, this gives $\lambda_k \to +\infty$. Thus, by the definition of d_k, we have

$$d_k \to 0. \tag{29}$$

Case 2: S is infinite. Based on Lemma 2.1, (22) and (25), we have

$$\begin{aligned}
\|F_1\|^2 &\geq \sum_{k=1}^{\infty}(\|F_k\|^2 - \|F_{k+1}\|^2) \geq \sum_{k\in S}(\|F_k\|^2 - \|F_{k+1}\|^2) \\
&\geq \sum_{k\in S} p_0 Pred_k \geq \sum_{k\in S\cap\bar{S}} p_0 Pred_k \\
&\geq \sum_{k\in S\cap\bar{S}} p_0 \|G_k^T F_k\| \min\left\{\|d_k\|, \frac{\|G_k^T F_k\|}{\|G_k^T G_k\|}\right\} \\
&\geq \sum_{k\in S\cap\bar{S}} p_0 \tau \min\left\{\|d_k\|, \frac{\tau}{\kappa_b^2}\right\}.
\end{aligned} \tag{30}$$

We assert that the set $S \cap \bar{S}$ is infinite. Otherwise, if it is finite. Let $k_{\bar{i}}$ be the largest index in it. Then, for $i > \bar{i}, r_{k_i} < p_0 < p_2$. By (16), $G_{k_i+1} = J_{k_i+1}$. Thus, $k_{i+1} = k_i + 1$. Moreover, $k_{i+1} \notin S$. Therefore, we conclude that $r_k < p_0$ for all sufficiently large k, which contradicts to the infiniteness of S. Thus, $S \cap \bar{S}$ is infinite.

According to the definition of $\bar{S}$, $k_i \in \bar{S}$. Based on (30), we have $d_{k_i} \to 0$ for $k_i \in S$. Since $d_{k_i} = 0$ for $k_i \notin S$, we obtain $d_{k_i} \to 0$. This, together with $\|G_{k_i}\| \leq \kappa_b, \|G_{k_i}^T F_{k_i}\| \geq \tau$ and (12), gives $\lambda_{k_i} \to +\infty$.

By (22) and (23), for $k = k_i + 1, \cdots, k_i + s_i - 1$,

$$\begin{aligned}
\|d_k\| &= \| - (G_k^T G_k + \lambda_k I)^{-1} G_k^T F_k\| \\
&\leq \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T F_{k_i}\| + \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T J_{k_i} (\sum_{j=k_i}^{k-1} d_j)\| \\
&\quad + \kappa_h \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T\| \|\sum_{j=k_i}^{k-1} d_j\|^2 \\
&\leq \|d_{k_i}\| + \sum_{j=k_i}^{k-1}\|d_j\| + \frac{\kappa_b \kappa_h}{\lambda_{k_i}}(\sum_{j=k_i}^{k-1}\|d_j\|)^2.
\end{aligned}$$

Owing to $\lambda_{k_i} \to +\infty$, we have

$$\|d_{k_i+1}\| \leq 3\|d_{k_i}\| \tag{31}$$

and

$$\begin{aligned}
\|d_{k_i+2}\| &\leq \|d_{k_i}\| + \|d_{k_i}\| + \|d_{k_i+1}\| + \frac{\kappa_b \kappa_h}{\lambda_{k_i}}(\|d_{k_i}\| + \|d_{k_i+1}\|)^2 \\
&\leq 21\|d_{k_i}\|
\end{aligned} \tag{32}$$

for sufficiently large k_i. From induction, it follows that there is a positive constant $\tilde{c}$ such that, for $k = k_i + 1, \cdots, k_i + s_i - 1$,

$$\|d_k\| \leq \tilde{c}\|d_{k_i}\| \tag{33}$$

holds for all sufficiently large k_i. Since $s_i \leq t$ for all i, we have $d_k \to 0$ and

$$\|J_k - G_k\| = \|J_k - J_{k_i}\| \leq \kappa_h \sum_{j=k_i}^{k-1}\|d_j\| \to 0. \tag{34}$$

Therefore, by (25),

$$\|G_k^T F_k\| \geq \|J_k^T F_k\| - \|(J_k - G_k)^T F_k\| \geq \|J_k^T F_k\| - \|J_k - G_k\|\|F_k\| \geq \frac{\tau}{2} \tag{35}$$

holds for sufficiently large k. From the definitions of d_k and λ_k, we have $\lambda_k \to +\infty$ and $\mu_k \to +\infty$.

Hence, whether S is infinite or finite, we obtain

$$d_k \to 0, \lambda_k \to +\infty, \mu_k \to +\infty. \tag{36}$$

Since $\|F_k\|$ is nonincreasing, $\|F_k\| \leq \|F_1\|$. By (21), (22) and (35), we have

$$\begin{aligned}
|r_k - 1| &= \left|\frac{Ared_k - Pred_k}{Pred_k}\right| \\
&= \left|\frac{\|F(x_k + d_k)\|^2 - \|F_k + G_k d_k\|^2}{Pred_k}\right| \\
&\leq \frac{\left|\|F_k + J_k d_k\|^2 - \|F_k + G_k d_k\|^2\right| + 2\kappa_h \|F_k + J_k d_k\|\|d_k\|^2 + \kappa_h^2\|d_k\|^4}{\|G_k^T F_k\| \min\left\{\|d_k\|, \frac{\|G_k^T F_k\|}{\|G_k^T G_k\|}\right\}} \\
&\leq \frac{(\|J_k^T J_k\| + \|G_k^T G_k\|)\|d_k\|^2 + 2\|J_k - G_k\|\|F_k\|\|d_k\|}{\frac{\tau}{2}\min\left\{\|d_k\|, \frac{\tau}{2\kappa_h^2}\right\}} \\
&\quad + \frac{2\kappa_h\|F_1\|\|d_k\|^2 + 2\kappa_h\kappa_b\|d_k\|^3 + \kappa_h^2\|d_k\|^4}{\frac{\tau}{2}\min\left\{\|d_k\|, \frac{\tau}{2\kappa_h^2}\right\}} \\
&\to 0,
\end{aligned}$$

which implies

$$\lim_{k\to+\infty} r_k = 1. \tag{37}$$

By (14), there is a positive constant $\tilde{\mu}$ such that

$$\mu_k < \tilde{\mu} \tag{38}$$

holds for all sufficiently large k. This leads to a contradiction with (36). Thus, (25) cannot be true, (24) is necessarily true. ∎

III. Local convergence rate

This section establishes the local convergence theory of Algorithm 1. Assume that the sequence $\{x_k\}$ lies in some neighborhood of $x^* \in X^*$ and converges to the solution set X^* of (1). We now make a further assumption.

Assumption 3.1. $\|J(x)^T F(x)\|$ provides a Hölderian local error bound with order $\gamma \in (0, 1]$ in some neighborhood of $x^* \in X^*$, i.e., there are constants $\kappa_l > 0$ and $0 < b < 1$ such that

$$\kappa_l \cdot dist(x, X^*) \leq \|J(x)^T F(x)\|^\gamma, \forall x \in N(x^*, b), \tag{39}$$

where $N(x^*, b) = \left\{x \in R^n \middle| \|x - x^*\| \leq b\right\}$.

By (22), we obtain

$$\|F(y) - F(x)\| \leq \kappa_b\|y - x\|, \forall x, y \in R^n. \tag{40}$$

Denote by $\bar{x}_k$ the closest point to x_k in X^*, i.e.,

$$\|\bar{x}_k - x_k\| = dist(x_k, X^*). \tag{41}$$

For all k large enough,

$$\|x_k - \bar{x}_k\| \leq \|x_k - x^*\| \leq \frac{b}{2} \tag{42}$$

and

$$\|\bar{x}_k - x^*\| \leq \|\bar{x}_k - x_k\| + \|x_k - x^*\| \leq b. \tag{43}$$

Thus $\bar{x}_k \in N(x^*, b)$.

In the following, we first show that $\|d_k\|$ is related to $dist(x_{k_i}, X^*)$, then prove that the parameter μ_k is upper bounded, and finally derive the local convergence rate of Algorithm 1 by the singular value decomposition technique.

Lemma 3.1: Given the conditions provided by Assumptions 2.1 and 3.1, there is a constant $c > 0$ such that

$$\|d_k\| \leq c\|\bar{x}_{k_i} - x_{k_i}\|, k = k_i, \cdots, k_i + s_i - 1. \tag{44}$$

Proof: By (11) and (39), we have

$$\lambda_{k_i} = \mu_{k_i}\|G_{k_i}^T F_{k_i}\|^\delta = \mu_{k_i}\|J_{k_i}^T F_{k_i}\|^\delta \geq \mu_0\kappa_l\|\bar{x}_{k_i} - x_{k_i}\|. \tag{45}$$

Define

$$\psi_k(d) = \|F_k + J_k d\|^2 + \lambda_k\|d\|^2. \tag{46}$$

It's obvious that d_{k_i} is the minimizer of $\psi_{k_i}(d)$. From (23), we obtain

$$\begin{aligned}\|F_{k_i} + J_{k_i}(\bar{x}_{k_i} - x_{k_i})\|^2 &= \|F(\bar{x}_{k_i}) - F_{k_i} - J_{k_i}(\bar{x}_{k_i} - x_{k_i})\|^2 \\ &\leq \kappa_h^2\|\bar{x}_{k_i} - x_{k_i}\|^4.\end{aligned} \tag{47}$$

By (45) and (47), we have

$$\begin{aligned}\|d_{k_i}\|^2 &\leq \frac{\psi_{k_i}(d_{k_i})}{\lambda_{k_i}} \\ &\leq \frac{\psi_{k_i}(\bar{x}_{k_i} - x_{k_i})}{\lambda_{k_i}} \\ &= \frac{\|F_{k_i} + J_{k_i}(\bar{x}_{k_i} - x_{k_i})\|^2 + \lambda_{k_i}\|\bar{x}_{k_i} - x_{k_i}\|^2}{\lambda_{k_i}} \\ &\leq \frac{\kappa_h^2}{\mu_0\kappa_l}\|\bar{x}_{k_i} - x_{k_i}\|^4 + \|\bar{x}_{k_i} - x_{k_i}\|^2 \\ &= O(\|\bar{x}_{k_i} - x_{k_i}\|^2).\end{aligned}$$

Thus, $\|d_{k_i}\| \leq c_0\|\bar{x}_{k_i} - x_{k_i}\|$.

For $k = k_i, \cdots, k_i + s_i - 1$,

$$\begin{aligned}\|d_k\| &= \| - (J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T F_k\| \\ &\leq \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T F_{k_i}\| + \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T J_{k_i}(\sum_{j=k_i}^{k-1} d_j)\| \\ &+ \kappa_h\|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T\|\|\sum_{j=k_i}^{k-1} d_j\|^2 \\ &\leq \|d_{k_i}\| + \sum_{j=k_i}^{k-1}\|d_j\| + \kappa_h\|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T\| \left(\sum_{j=k_i}^{k-1}\|d_j\|\right)^2\end{aligned}$$

and

$$\begin{aligned}\|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T\| &= \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1} J_{k_i}^T J_{k_i}(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1}\|^{\frac{1}{2}} \\ &\leq \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1}(J_{k_i}^T J_{k_i} + \lambda_{k_i} I))(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1}\|^{\frac{1}{2}} \\ &= \|(J_{k_i}^T J_{k_i} + \lambda_{k_i} I)^{-1}\|^{\frac{1}{2}} \\ &\leq \frac{1}{\sqrt{\lambda_{k_i}}} \\ &\leq \frac{1}{\sqrt{u_0\kappa_l}}\|\bar{x}_{k_i} - x_{k_i}\|^{-\frac{1}{2}}.\end{aligned}$$

Therefore, it follows that

$$\begin{aligned}\|d_{k_i+1}\| &\leq \|d_{k_i}\| + \|d_{k_i}\| + \frac{\kappa_h}{\sqrt{u_0\kappa_l}}\|d_{k_i}\|^{\frac{3}{2}} \\ &\leq c_1\|d_{k_i}\|.\end{aligned} \tag{48}$$

Similarly,

$$\begin{aligned}\|d_{k_i+2}\| &\leq \|d_{k_i}\| + \|d_{k_i}\| + \|d_{k_i+1}\| + \bar{c}_2\|d_{k_i}\| \\ &\leq c_2\|d_{k_i}\|\end{aligned} \tag{49}$$

and

$$\|d_{k_i+q}\| \leq c_q\|d_{k_i}\|, q = 3, \cdots, s_i - 1. \tag{50}$$

Thus, we get (44). ■

From Algorithm 1, it is known that μ_k is lower bounded. Furthermore, we will demonstrate that μ_k is also upper bounded.

Lemma 3.2: Given the conditions provided by Assumptions 2.1 and 3.1, there is a positive constant $\bar{\mu}$ such that

$$\mu_k \leq \bar{\mu} \tag{51}$$

holds for sufficiently large k.

Proof: We first prove that

$$\|F_k\|^2 - \|F_k + G_k d_k\|^2 \geq \frac{\kappa_l}{2}\|F_k\|\min\{\|d_k\|, \|\bar{x}_k - x_k\|\} \tag{52}$$

holds for sufficiently large k.

According to Lemma 3.1, (21),(23) and $s_i \leq t$, for $k = k_i, \cdots, k_i + s_i - 1$,

$$\begin{aligned}\|F_k + G_k(\bar{x}_k - x_k)\| &\leq \|F_k + J_k(\bar{x}_k - x_k)\| + \|G_k - J_k\|\|\bar{x}_k - x_k\| \\ &\leq \kappa_h\|\bar{x}_k - x_k\|^2 + \kappa_h(\sum_{j=k_i}^{k-1}\|d_j\|)\|\bar{x}_k - x_k\| \\ &\leq \kappa_h\|\bar{x}_k - x_k\|^2 + \kappa_h ct\|\bar{x}_{k_i} - x_{k_i}\|\|\bar{x}_k - x_k\|\end{aligned}$$

Note that $\|\bar{x}_k - x_k\| \to 0$ and $\|\bar{x}_{k_i} - x_{k_i}\| \to 0$, we obtain

$$\|F_k + G_k(\bar{x}_k - x_k)\| \leq \frac{\kappa_l}{2\kappa_b}\|\bar{x}_k - x_k\| \tag{53}$$

holds for sufficiently large k.

Our discussion is divided into two cases.

Case 1: $\|\bar{x}_k - x_k\| \leq \|d_k\|$. By Lemma 3.1, (39) and (53), we have

$$\begin{aligned}\|F_k\| - \|F_k + G_k d_k\| &\geq \frac{1}{\kappa_b}\|J_k^T F_k\| - \|F_k + G_k(\bar{x}_k - x_k)\| \\ &\geq \frac{\kappa_l}{\kappa_b}\|\bar{x}_k - x_k\| - \frac{\kappa_l}{2\kappa_b}\|\bar{x}_k - x_k\| \\ &= \frac{\kappa_l}{2\kappa_b}\|\bar{x}_k - x_k\|\end{aligned} \tag{54}$$

holds for sufficiently large k.

Case 2: $\|\bar{x}_k - x_k\| > \|d_k\|$. By (54), we obtain

$$\begin{aligned}\|F_k\| - \|F_k + G_k d_k\| &\geq \|F_k\| - \|F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}G_k(\bar{x}_k - x_k)\| \\ &= \|F_k\| - \|(1 - \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|})F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}G_k(\bar{x}_k - x_k)\| \\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}(\|F_k\| - \|F_k + G_k(\bar{x}_k - x_k)\|) \\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}\frac{\kappa_l}{2\kappa_b}\|\bar{x}_k - x_k\| \\ &= \frac{\kappa_l}{2\kappa_b}\|d_k\|.\end{aligned} \tag{55}$$

holds for sufficiently large k.

Hence by (54) and (55),

$$\begin{aligned}Pred_k &= (\|F_k\| + \|F_k + G_k d_k\|)(\|F_k\| - \|F_k + G_k d_k\|) \\ &\geq \|F_k\|(\|F_k\| - \|F_k + G_k d_k\|) \\ &\geq \frac{\kappa_l}{2\kappa_b}\|F_k\|\min\{\|d_k\|, \|\bar{x}_k - x_k\|\}\end{aligned} \tag{56}$$

holds for sufficiently large k.

Since $\|F_{k_i} + J_{k_i}d_{k_i}\| \leq \|F_{k_i}\|$, by (39) and (56), we have

$$\begin{aligned}|r_{k_i} - 1| &= \left|\frac{Ared_{k_i} - Pred_{k_i}}{Pred_{k_i}}\right| \\ &= \left|\frac{\|F(x_{k_i} + d_{k_i})\|^2 - \|F_{k_i} + J_{k_i}d_{k_i}\|^2}{Pred_{k_i}}\right| \\ &\leq \left|\frac{2\|F_{k_i} + J_{k_i}d_{k_i}\|\|d_{k_i}\|^2 + \kappa_h^2\|d_{k_i}\|^4}{\frac{\kappa_l}{2\kappa_b}\|F_{k_i}\|\min\{\|d_{k_i}\|, \|\bar{x}_{k_i} - x_{k_i}\|\}}\right| \\ &\to 0.\end{aligned} \tag{57}$$

Therefore,

$$\lim_{k_i \to \infty} r_{k_i} = 1. \tag{58}$$

Note that for $k \notin \bar{S}, r_k \geq p_2 > p_1$ and $G_{k+1} = G_k$, so $\mu_{k+1} \leq \mu_k$. According to the updating rule (14), inequality (51) holds. ■

Next, we derive the local convergence rate of Algorithm 1 by means of the singular value decomposition (SVD) technique and the matrix perturbation theory.

Based on the results obtained by Behling and Iusem in [2], we might assume that $rank(J(\bar{x})) = r$ for all $\bar{x} \in N(x^*, b) \cap X^*$. Suppose that the SVD of $J(\bar{x}_{k_i})$ is

$$\begin{aligned} J(\bar{x}_{k_i}) &= \bar{U}_{k_i} \bar{\Sigma}_{k_i} \bar{V}_{k_i}^T \\ &= (\bar{U}_{k_i,1}, \bar{U}_{k_i,2}) \begin{pmatrix} \bar{\Sigma}_{k_i,1} & \\ & 0 \end{pmatrix} \begin{pmatrix} \bar{V}_{k_i,1}^T \\ \bar{V}_{k_i,2}^T \end{pmatrix} \\ &= \bar{U}_{k_i,1} \bar{\Sigma}_{k_i,1} \bar{V}_{k_i,1}^T, \end{aligned} \tag{59}$$

where $\bar{\Sigma}_{k_i,1} = diag(\bar{\sigma}_{k_i,1}, \cdots, \bar{\sigma}_{k_i,r}) > 0$ and $\bar{U}_{k_i}, \bar{V}_{k_i}$ are orthogonal matrices.

Consequently, the SVD of J_{k_i} is

$$\begin{aligned} J_{k_i} &= U_{k_i} \Sigma_{k_i} V_{k_i}^T \\ &= (U_{k_i,1}, U_{k_i,2}) \begin{pmatrix} \Sigma_{k_i,1} & \\ & \Sigma_{k_i,2} \end{pmatrix} \begin{pmatrix} V_{k_i,1}^T \\ V_{k_i,2}^T \end{pmatrix} \\ &= U_{k_i,1} \Sigma_{k_i,1} V_{k_i,1}^T + U_{k_i,2} \Sigma_{k_i,2} V_{k_i,2}^T, \end{aligned} \tag{60}$$

where $\Sigma_{k_i,1} = diag(\sigma_{k_i,1}, \cdots, \sigma_{k_i,r}) > 0$ and $\Sigma_{k_i,2} = diag(\sigma_{k_i,r+1}, \cdots, \sigma_{k_i,n}) \geq 0$.

By the theory of matrix perturbation [22] and (21), we have

$$\|diag(\Sigma_{k_i,1} - \bar{\Sigma}_{k_i,1}, \Sigma_{k_i,2})\| \leq \|J_{k_i} - J(\bar{x}_{k_i})\| \leq \kappa_h \|\bar{x}_{k_i} - x_{k_i}\|. \tag{61}$$

Lemma 3.3: Under Assumptions 2.1 and 3.1, there are positive constants l_1 and l_2 such that

$$\|d_k\| \leq l_1 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1}, k = k_i, \cdots, k_i + s_i - 1. \tag{62}$$

$$\|F_k + G_k d_k\| \leq l_2 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+2}, k = k_i, \cdots, k_i + s_i - 1. \tag{63}$$

Proof: We prove it by induction. It was shown in [6] that the results hold true for $k = k_i$ and $k = k_i + 1$.

Suppose the results hold true for $k - 1 (k_i + 2 < k < k_i + s_i - 1)$, that is, there exist constants $\bar{l}_1$ and $\bar{l}_2$ such that

$$\|d_{k-1}\| \leq \bar{l}_1 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i}, \tag{64}$$

$$\|F_{k-1} + G_{k-1} d_{k-1}\| \leq \bar{l}_2 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1}. \tag{65}$$

It then follows from (21), (23) and Lemma 3.1 that

$$\begin{aligned} \|F_k\| &= \|F(x_{k-1} + d_{k-1})\| \\ &\leq \|F_{k-1} + J_{k-1} d_{k-1}\| + \kappa_h \|d_{k-1}\|^2 \\ &\leq \|F_{k-1} + G_{k-1} d_{k-1}\| + \|J_{k-1} - G_{k-1}\| \|d_{k-1}\| + \kappa_h \|d_{k-1}\|^2 \\ &= \|F_{k-1} + G_{k-1} d_{k-1}\| + \|J_{k-1} - J_{k_i}\| \|d_{k-1}\| + \kappa_h \|d_{k-1}\|^2 \\ &\leq \bar{l}_2 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1} + \kappa_h (\sum_{j=k_i}^{k-2} \|d_j\|) \|d_{k-1}\| + \kappa_h \bar{l}_1^2 \|\bar{x}_{k_i} - x_{k_i}\|^{2(k-k_i)} \\ &\leq \bar{l}_3 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1}, \end{aligned}$$

where $\bar{l}_3 = \bar{l}_2 + \kappa_h c t \bar{l}_1 + \kappa_h \bar{l}_1^2$. Therefore,

$$\|U_{k_i,1} U_{k_i,1}^T F_k\| \leq \|F_k\| \leq \bar{l}_3 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1}. \tag{66}$$

Moreover, by (39),

$$\|\bar{x}_k - x_k\| \leq \kappa_l^{-1} \|F_k\| \leq \kappa_l^{-1} \bar{l}_3 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1}. \tag{67}$$

Let $\tilde{J}_{k_i} = U_{k_i,1} \Sigma_{k_i,1} V_{k_i,1}^T$ and $\tilde{s}_k = -\tilde{J}_{k_i}^+ F_k$, where $\tilde{J}_k^+$ denotes the pesudo-inverse of $\tilde{J}_{k_i}$. Then $\tilde{s}_k$ becomes the least squares solution to $\min_{s \in R^n} \|F_k + \tilde{J}_{k_i} \tilde{s}_k\|$. From (21) , (61) and (67), we have

$$\begin{aligned} \|U_{k_i,2} U_{k_i,2}^T F_k\| &= \|F_k + \tilde{J}_{k_i} \tilde{s}_k\| \leq \|F_k + \tilde{J}_{k_i} (\bar{x}_k - x_k)\| \\ &\leq \|F_k + J_k (\bar{x}_k - x_k)\| + \|J_{k_i} - J_k\| \|\bar{x}_k - x_k\| + \|\tilde{J}_{k_i} - J_{k_i}\| \|\bar{x}_k - x_k\| \\ &\leq \kappa_h \|\bar{x}_k - x_k\|^2 + \kappa_h (\sum_{j=k_i}^{k-1} \|d_j\|) \|\bar{x}_k - x_k\| + \|U_{k_i,2} \Sigma_{k_i,2} V_{k_i,2}^T\| \|\bar{x}_k - x_k\| \\ &\leq \bar{l}_4 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+2}, \end{aligned}$$

where $\bar{l}_4 = \kappa_l^{-1} \bar{l}_3 \kappa_h (\kappa_l^{-1} \bar{l}_3 + ct + 1)$.

Since $\{x_k\}$ converges to X^*, we may assume that $\kappa_h \|\bar{x}_{k_i} - x_{k_i}\| \leq \frac{\bar{\sigma}}{2}$ holds for all sufficiently large k_i. Then, it follows from (61) that

$$\|(\Sigma_{k_i,1}^2 + \lambda_{k_i} I)^{-1}\| \leq \|\Sigma_{k_i,1}^{-2}\| \leq \frac{1}{(\bar{\sigma} - \kappa_h \|\bar{x}_{k_i} - x_{k_i}\|)^2} \leq \frac{4}{\bar{\sigma}^2}. \tag{68}$$

By (61), we obtain

$$\|\Sigma_{k_i,1}^{-1}\| = \frac{1}{\sigma_{k_i,r}} \leq \frac{1}{|\bar{\sigma}_{k_i,r} - \kappa_h \|\bar{x}_{k_i} - x_{k_i}\||} \leq \frac{2}{\bar{\sigma}_{k_i,r}} \tag{69}$$

holds for sufficiently large k_i. By (45),

$$\|\lambda_{k_i}^{-1} \Sigma_{k_i,2}\| = \frac{\|\Sigma_{k_i,2}\|}{\lambda_{k_i}} \leq \frac{\kappa_h}{\mu_0 \kappa_l}. \tag{70}$$

By the SVD of J_k, we have

$$d_k = -V_{k_i,1} (\Sigma_{k_i,1}^2 + \lambda_{k_i} I)^{-1} \Sigma_{k_i,1} U_{k_i,1}^T F_k - V_{k_i,2} (\Sigma_{k_i,2}^2 + \lambda_{k_i} I)^{-1} \Sigma_{k_i,2} U_{k_i,2}^T F_k.$$

Therefore,

$$\begin{aligned} \|d_k\| &\leq \|\Sigma_{k_i,1}^{-1}\| \|U_{k_i,1} U_{k_i,1}^T F_k\| + \|\lambda_{k_i}^{-1} \Sigma_{k_i,2}\| \|U_{k_i,2} U_{k_i,2}^T F_k\| \\ &\leq \frac{2\bar{l}_3}{\bar{\sigma}_{k_i,r}} \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1} + \frac{\kappa_h \bar{l}_4}{\mu_0 \kappa_l} \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+2} \\ &\leq l_1 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+1}, \end{aligned} \tag{71}$$

where $l_1 = \frac{2\bar{l}_3}{\bar{\sigma}_{k_i,r}} + \frac{\kappa_h \bar{l}_4}{\mu_0 \kappa_l} \|\bar{x}_{k_i} - x_{k_i}\|$.

Note that $\|(\Sigma_{k_i,2}^2 + \lambda_{k_i} I)^{-1}\| \leq \lambda_{k_i}^{-1}$ and $\lambda_{k_i} = \frac{\mu_{k_i} \|F_{k_i}\|}{1 + \|F_{k_i}\|} \leq \mu_{k_i} \|F_{k_i}\| \leq \bar{\mu} \kappa_b \|\bar{x}_{k_i} - x_{k_i}\|$, it follows

$$\begin{aligned} &\|F_k + G_k d_k\| \\ &= \|\lambda_{k_i} U_{k_i,1} (\Sigma_{k_i,1}^2 + \lambda_{k_i} I)^{-1} U_{k_i,1}^T F_k + \lambda_{k_i} U_{k_i,2} (\Sigma_{k_i,2}^2 + \lambda_{k_i} I)^{-1} U_{k_i,2}^T F_k\| \\ &\leq \lambda_{k_i} \|\Sigma_{k_i,1}^{-2}\| \|U_{k_i,1} U_{k_i,1}^T F_k\| + \|U_{k_i,2} U_{k_i,2}^T F_k\| \\ &\leq 4 \bar{\mu} \kappa_b \bar{\sigma}^{-2} \bar{l}_3 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+2} + \bar{l}_4 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+2} \\ &= l_2 \|\bar{x}_{k_i} - x_{k_i}\|^{k-k_i+2}, \end{aligned} \tag{72}$$

where $l_2 = 4 \bar{\mu} \kappa_b \bar{\sigma}^{-2} \bar{l}_3 + \bar{l}_4$. This completes the proof. ■

Theorem 3.1: Under Assumptions 2.1 and 3.1, there exists a constant $l_3 > 0$ such that

$$\|d_{k_{i+1}}\| \leq l_3 \|d_{k_i}\|^{s_i+1}. \tag{73}$$

Consequently, Algorithm 1 converges q-superlinearly to a solution of (1).

Proof: By Lemma 3.1, (39) and (23), we have

$$\begin{aligned}\kappa_l\|\bar{x}_{k_{i+1}} - x_{k_{i+1}}\| &\le \|F(x_{k_{i+1}})\| = \|F(x_{k_i+s_i-1} + d_{k_i+s_i-1})\| \\ &\le \|F_{k_i+s_i-1} + J_{k_i+s_i-1}d_{k_i+s_i-1}\| + \kappa_h\|d_{k_i+s_i-1}\|^2 \\ &\le \|F_{k_i+s_i-1} + G_{k_i+s_i-1}d_{k_i+s_i-1}\| + \|(J_{k_i+s_i-1} - G_{k_i+s_i-1})d_{k_i+s_i-1}\| \\ &+ \kappa_h\|d_{k_i+s_i-1}\|^2 \\ &\le \|F_{k_i+s_i-1} + G_{k_i+s_i-1}d_{k_i+s_i-1}\| + (\sum_{j=k_i}^{k_i+s_i-2}\|d_j\|)\|d_{k_i+s_i-1}\| \\ &+ \kappa_h\|d_{k_i+s_i-1}\|^2 \\ &\le (l_2 + ctl_1 + \kappa_h cl_1)\|\bar{x}_{k_i} - x_{k_i}\|^{s_i+1}.\end{aligned} \tag{74}$$

Note that

$$\|\bar{x}_{k_i} - x_{k_i}\| \le \|\bar{x}_{k_{i+1}} - x_{k_i}\| \le \|\bar{x}_{k_{i+1}} - x_{k_{i+1}}\| + \sum_{j=k_i}^{k_{i+1}-1}\|d_j\|. \tag{75}$$

By (74) and (75),

$$\|\bar{x}_{k_i} - x_{k_i}\| \le 2\sum_{j=k_i}^{k_{i+1}-1}\|d_j\| \le 2s_i c_{s_i-1}\|d_{k_i}\| \tag{76}$$

holds for sufficiently large k_i.

Let $\bar{l}_5 = \kappa_l^{-1}(l_2 + tcl_1 + \kappa_h cl_1)$. From (74), Lemmas 3.1 and 3.3, we have

$$\begin{aligned}\|d_{k_{i+1}}\| &\le c\|\bar{x}_{k_{i+1}} - x_{k_{i+1}}\| \le c\bar{l}_5\|\bar{x}_{k_i} - x_{k_i}\|^{s_i+1} \\ &\le c\bar{l}_5(2s_i c_{s_i-1})^{s_i+1}\|d_{k_i}\|^{s_i+1} \\ &\le c\bar{l}_5(2tc_{t-1})^{t+1}\|d_{k_i}\|^{s_i+1}.\end{aligned} \tag{77}$$

Let $l_3 = c\bar{l}_5(2tc_{t-1})^{t+1}$. We get (73).

By Lemma 3.1, we have $\|d_{k_i}\| \to 0$. Therefore, there exist a positive integer N and $0 < q < 1$ such that $\max\{\|d_{k_i}\|, l_3\|d_{k_i}\|\} \le q < 1$ for all $i \ge N$. Since $s_i \ge 1$, we have

$$\|d_{k_{i+1}}\| \le l_3\|d_{k_i}\|^{s_i+1} \le l_3\|d_{k_i}\|^2 \le q\|d_{k_i}\|, \forall i \ge N. \tag{78}$$

Then,

$$\begin{aligned}\sum_{i=N}^{\infty}\|d_{k_i}\| &\le \|d_{k_N}\| + q\|d_{k_N}\| + q^2\|d_{k_N}\| + \cdots \\ &= \frac{\|d_{k_N}\|}{1-q},\end{aligned} \tag{79}$$

which implies that $\sum_{i=N}^{\infty}\|d_{k_i}\|$ converges.

By $s_i \le t$, we have

$$\|x_{k_{i+1}} - x_{k_i}\| = \|\sum_{j=0}^{s_i-1} d_{k_i+j}\| \le tc_{t-1}\|d_{k_i}\|. \tag{80}$$

Thus, the infinite series $\sum_{i=1}^{\infty}\|x_{k_{i+1}} - x_{k_i}\|$ converges. This implies that the infinite series $\sum_{i=1}^{\infty}(x_{k_{i+1}} - x_{k_i})$ converges. Consequently, the sequence $\{x_{k_i}\}$ converges to some $\hat{x} \in X^*$. Similarly to (74), we have

$$\|x_{k_{i+1}} - \hat{x}\| \le \hat{q}\|x_{k_i} - \hat{x}\|^{s_i+1} \tag{81}$$

for some $\hat{q} > 0$. Therefore, Algorithm 1 converges superlinearly. ■

IV. Numerical experiments

In this section, numerical results for Algorithm 1 in solving nonlinear equations are presented. The test problems are selected as follows. The test problems 1-4 do not satisfy the local error bound condition but instead satisfy the Hölderian local error bound condition near zero point. Problems 5-9, which correspond to the functions from [16], are listed in Table I.

Problem 1 ([24]) $f_1(x) = x_1x_2$,

$$f_2(x) = x_1^2 + x_2^2.$$

Initial point: $x_0 = (1,1)^T$, zero point: $(0,0)^T$.

Problem 2 Powell singular function ([14], [16])

$$f_1(x) = x_1 + 10x_2,$$
$$f_2(x) = \sqrt{5}(x_3 - x_4),$$
$$f_3(x) = (x_2 - 2x_3)^2,$$
$$f_4(x) = \sqrt{10}(x_1 - x_4)^2.$$

Initial point: $x_0 = (3,1,0,1)^T$, zero point: $(0,0,0,0)^T$.

Problem 3 ([22], [24])

$$f_1(x) = x_1 + 10x_2,$$
$$f_2(x) = x_3 - x_4,$$
$$f_3(x) = (x_2 - 2x_3)^{\frac{3}{2}},$$
$$f_4(x) = (x_1 - x_4)^{\frac{3}{2}}.$$

Initial point: $x_0 = (3,1,0,1)^T$, zero point: $(0,0,0,0)^T$.

Problem 4

$$f_1(x) = x_1^2 - x_1x_2,$$
$$f_2(x) = x_2^2 + x_1x_2.$$

Initial point: $x_0 = (1,1)^T$, zero point: $x_0 = (0,0)^T$.

Test problems 5-9 are constructed by adjusting the nonsingular problems from Moré, Garbow, and Hillstrom in [16], and they meet the local error bound condition near the zero point. They have the same form as [18]:

$$\hat{F}(x) = F(x) - J(x^*)A(A^TA)^{-1}A^T(x - x^*),$$

where $F(x)$ is a standard nonsingular test function, $A \in \mathbb{R}^{n\times k}$ $(1 \le k \le n)$ is a matrix which has full column rank, and x^* is a solution of $F(x) = 0$. It is obvious that

$$\hat{J}(x^*) = J(x^*)(I - A(A^TA)^{-1}A^T),$$

is the Jacobian of $\hat{F}(x)$ at x^* with rank $n-k$ and $\hat{F}(x^*) = 0$. Note that, some roots of $\hat{F}(x) = 0$ may be not roots of $F(x)$. Similar to [18], we take

$$A = [1,1,\cdots,1]^T \in \mathbb{R}^{n\times 1},$$

which results in $rank(\hat{J}(x^*)) = n - 1$.

The codes are written in MATLAB R2016 and run on a personal computer 2.9 GHz and 2.9 GHz CPU processor, 16.0 GB RAM, using Windows 11 operation system. Throughout the numerical experiments, We take $p_0 = 10^{-4}, p_1 = 0.25, p_2 = 0.50, p_3 = 0.75, u_1 = 0.01, u_0 = 10^{-8}, m_1 = 4, m_2 = 0.25, t = 10, \delta = \frac{1}{2}$. We stop the program if the number of iterations exceeds $100(n+1)$ or $\|G_k^TF_k\| \le 10^{-5}$.

To consider the global convergence of the algorithms, we run each test problem for three starting points $x_0, 10x_0, 100x_0$, where x_0 is suggested in [16] for problems 5-9. The results are given in Tables II and III. The meaning of notations listed in two tables is as follows:

NF: The numbers of function calculations.

NJ: The numbers of Jacobian calculations.

NT: $NT = NF + NJ * n$.

NS: 'Y' means the algorithm is converged to x^*; 'N' means that algorithm is converged to another solution; '-' shows that the number of iterates is more than $100(n+1)$.

From Table II, we see that Algorithm 1 is effective for singular problems 5-9 under the local error bound condition. Table III reveals that Algorithm 1 converges globally to x^* for the singular problems 1-4 under the weaker Hölderian local error bound condition.

TABLE I
FUNCTIONS CORRESPONDING TO PROBLEMS 5-9

Problem	Function
5	Discrete boundary value function
6	Discrete integral equation function
7	Trigonometric function
8	Broyden tridiagonal function
9	Broyden banded function

TABLE II
RESULTS FOR LARGE SCALE PROBLEMS 5-9 WITH RANK $n-1$

Problem	n	x_0	Algorithm 1	
			NF/NJ/NT	NS
5	1000	1	1/1/1001	Y
		10	12/2/2012	N
		100	21/4/4021	N
6	1000	1	20/4/4020	Y
		10	24/5/5024	Y
		100	21/4/4021	N
7	1000	1	17/9/9017	Y
		10	35/21/21035	Y
		100	69/37/37069	Y
8	1000	1	16/4/4016	Y
		10	26/7/7026	Y
		100	34/9/9034	Y
9	1000	1	23/6/6023	Y
		10	35/9/9035	Y
		100	46/13/13046	Y

TABLE III
RESULTS ON THE PROBLEMS 1-4

Problem	n	x_0	Algorithm 1	
			NF/NJ/NT	NS
1	2	1	14/4/22	Y
		10	22/6/34	Y
		100	29/8/45	Y
2	4	1	18/5/38	Y
		10	25/7/53	Y
		100	32/8/64	Y
3	4	1	19/5/39	Y
		10	21/7/49	Y
		100	36/6/54	Y
4	2	1	14/4/22	Y
		10	21/6/33	Y
		100	29/8/45	Y

V. CONCLUSIONS

In this paper, we propose a new adaptive multi-step LM method with the parameter $\lambda_k = \mu_k \|G_k^T F_k\|^\delta$, $\delta \in (0,1]$ and improve the convergence results. This algorithm combines the advantage of saving Jacobian evaluations with good convergence properties. The convergence rate of our algorithm is proven to be superlinear under mild assumptions. Numerical experiments demonstrate the validity of our proposed algorithm.

REFERENCES

[1] K. Amini, F. Rostami and G. Caristi, "An efficient Levenberg-Marquardt method with a new LM parameter for systems of nonlinear equations," *Optimization*, vol. 67, no. 5, pp. 1-14, 2018.

[2] R. Behling, A. Iusem, "The effect of calmness on the solution set of systems of nonlinear equations," *Math. Program.*, vol.137, no. 1-2, pp. 155-165, 2013.

[3] R. Behling, S. Douglas and A. Sandra, " Local convergence analysis of the Levenberg-Marquardt framework for nonzero-residue nonlinear least-squares problems under an error bound condition," *J. Optim. Theory Appl.*, vol. 183, no. 3, pp. 1099-1122, 2019.

[4] J.E. Dennis, R.B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equation*. Science Press, Beijing, 2009.

[5] J.Y. Fan, "A modified Levenberg-Marquardt algorithm for singular system of nonlinear equations," *J. Comput. Math.*, vol. 21, no. 5, pp. 625-636, 2003.

[6] J.Y. Fan, "The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence," *Math. Comput.*, vol. 81, pp. 447-466, 2012.

[7] J.Y. Fan, "A Shamanskii-like Levenberg-Marquardt method for nonlinear equations," *Comput. Optim. Appl.*, vol. 56, no. 1, pp. 63-80, 2013.

[8] J.Y. Fan, Y.X. Yuan, "On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption," *Computing*, vol. 74, no. 1, pp. 23-39, 2005.

[9] J.Y. Fan, J.C. Huang, J.Y. Pan, "An adaptive multi-step Levenberg-Marquardt method," *Journal of Scientific Computing*, vol. 78, no. 1, pp. 531-548, 2019.

[10] J.Y. Fan, J.Y. Pan, "A modified trust region algorithm for nonlinear equations with new updating rule of trust region radius," *International Journal of Computer Mathematics*, vol. 87, no. 14, pp. 3186-3195, 2010.

[11] M.L. Fang, M. Wang, D.F. Ding, et al, "A new modified nonlinear conjugate gradient method with sufficient descent property for unconstrained optimization," *Engineering Letters*, vol. 31, no. 3, pp. 1036-1044, 2023.

[12] K. Levenberg, "A method for the solution of certain nonlinear problems in least squares," *Quart Appl. Math.*, vol. 2, pp. 164-168, 1944.

[13] D.H. Li, M. Fukushima, "A global and superlinear convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations," *SIAM J. Numer. Anal.*, vol. 37, no. 1, pp. 152-172, 1999.

[14] D.H. Li, M. Fukushima, "A derivative-free line search and global convergence of Broyden-like method for nonlinear equations," *Optimization Methods and Software*, vol. 13, no. 3, pp. 181-201, 2000.

[15] D.W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, pp. 431-441, 1963.

[16] J.J. Moré, B,S. Garbow, K.E. Hillstrom, "Testing unconstrained optimization software," *ACM Trans. Math. Software*, vol. 7, no. 1, pp. 17-41, 1981.

[17] M.J.D. Powell, "Convergence properties of a class of minimization algorithms," *Nonlinear Program.*, vol. 2, pp. 1-27, 1975.

[18] R.B. Schnabel, P.D. Frank, "Tensor methods for nonlinear equations," *SIAM J. Numer. Anal.*, vol. 21, no. 5, pp. 815-843, 1984.

[19] A. Semiu, et al, "A descent conjugate gradient method for optimization problems," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 9, pp. 1765-1775, 2024.

[20] N. Yamashita, M. Fukushima, "On the rate of convergence of the Levenberg-Marquardt method," *Computing*, vol. 15, pp. 239-249, 2001.

[21] Y.X. Yuan, "Trust region algorithm for nonlinear equations," *Information*, vol. 1, pp. 7-21, 1998.

[22] G.W. Stewart, J.G. Sun, *Matrix perturbation theory*. San Diego (CA): Academic Press, Boston, 1990.

[23] J.L. Zhang, Y. Wang, "A new trust region method for nonlinear equations," *Mathematical Methods of Operations Research*, vol. 58, no. 2, pp. 283-298, 2003.

[24] M.L. Zeng, G.H. Zhou, "Improved convergence results of an efficient Levenberg-Marquardt method for nonlinear equations," *Journal of Applied Mathematics and Computing*, vol. 68, pp. 3655-3671, 2022.