The Convergence Conditions of Interval Newton's Method Based on Point Estimates

Zhe Li[1,2(✉)], Baocheng Wan[2,3], and Shugong Zhang[2]

[1] School of Science, Changchun University of Science and Technology,
Changchun 130022, China
lizhe@amss.ac.cn
[2] School of Mathematics, Key Laboratory of Symbolic Computation
and Knowledge Engineering, Jilin University, Changchun 130012, China
wanbaocheng@163.com, sgzh@jlu.edu.cn
[3] Jilin Agricultural University, Changchun 130118, China

Abstract. Both Smale's alpha theory and Rump's interval theorem provide the conditions which guarantee the existence of a simple solution of a square nonlinear system. In this paper, we generalize the conclusion provided by Rall to reveal the relationship between Smale's alpha theory and Rump's interval theorem. By point estimates, we propose the conditions under which the condition of Rump's interval theorem holds. Furthermore, using only the information of the given system at the initial approximate point, we give the convergence conditions of interval Newton's algorithm proposed by Rump.

Keywords: Newton iteration · Alpha theory · Interval algorithm · Point estimate · Verification

1 Introduction

Solving a nonlinear system in the form $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ with $\boldsymbol{f} = (f_1, f_2, \ldots, f_n)^T$ and $\boldsymbol{x} = (x_1, \ldots, x_n)^T$ is one of the most fundamental problems in scientific computing. In this paper, we assume that $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $f_1, f_2, \ldots, f_n$ have all order continuous partial derivatives. Denote the Jacobian matrix of $\boldsymbol{f}$ at $\boldsymbol{x}$ by $\boldsymbol{f}'(\boldsymbol{x})$.

Newton's method and its modifications have long played an important role in solving nonlinear systems. Under certain conditions, Newton's method constructs a sequence of iteration points that will converge to a solution of the given nonlinear system. In 1948, the author of [2] established the Kantorovich theorem based on the assumption that the Jacobian matrix of the nonlinear

system is Lipschitz continuous on some domain. The Kantorovich theorem first gives the condition to ensure that a simple solution exists close to the initial point and the Newton iteration sequence quadratically converges to this simple solution. Using the technique of point estimates, Smale et al. [15–17] developed the alpha theory to locate and approximate simple solutions. The alpha theory requires only information concerning the nonlinear system at the initial point of the Newton iteration. By introducing the dominating sequence technique, Wang and Han [18] improved both the condition and conclusion of the alpha theory. With the aid of Schröder operator, Giusti et al. [3] provided a criterion for locating clusters of solutions of univariate nonlinear functions. Later on, Giusti et al. [4] generalized their results to locate breadth-one singular solutions of multivariate nonlinear systems. For the performance of the alpha theory, Hauenstein and Sottile [6] described the program alphaCertified to certify solutions of polynomial systems. Recently, Hauenstein and Levandovskyy [5] extended the program alphaCertified to verify solutions to polynomial-exponential systems.

Interval arithmetic is another important tool of verification methods. In 1960s, Krawczyk [9] first introduced an interval version of Newton's method for verifying the existence of simple solutions. Moore [10] proposed computationally verifiable sufficient condition for interval Newton's method given by Krawczyk. Rump [12] made interval Newton's method perform better in practice, which is called Rump's interval theorem and included in **verifynlss** function in INT-LAB toolbox [13] in Matlab. Based on the deflation technique using smoothing parameters, Rump and Graillat [14] described a verification algorithm to verify multiple solutions of univariate nonlinear functions and double solutions of multivariate nonlinear systems. Further, Li and Zhi [8] provided an algorithm to verify breadth-one singular solutions of polynomial systems, which had been generalized to deal with the verification of isolated singular solutions in [7]. By combining interval algorithms with some other methods, Yang et al. [19] investigated the verification for real solutions of positive-dimensional polynomial systems.

In 1980, Rall [11] exhibited the relationship between the Kantorovich theorem and Moore's interval theorem. By the quantities of the Kantorovich theorem, Rall provided the conditions under which Moore's verifiable sufficient condition holds. For an initial approximate $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ and a radius $\rho > 0$, let $\boldsymbol{X}_\rho$ denote the set $\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}^{(0)}\|_\infty < \rho\}$, and let η, B, κ be the constants such that

$$\|\boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}(\boldsymbol{x}^{(0)})\|_\infty \leq \eta,$$
$$\|\boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\|_\infty \leq B,$$
$$\|\boldsymbol{f}'(\boldsymbol{u}) - \boldsymbol{f}'(\boldsymbol{v})\|_\infty \leq \kappa\|\boldsymbol{u} - \boldsymbol{v}\|, \qquad \boldsymbol{u}, \boldsymbol{v} \in \Omega,$$

where Ω is a sufficiently large region containing $\boldsymbol{x}^{(0)}$. Rall's conclusion is that if

$$h = B\kappa\eta < \frac{1}{4},$$

then

$$\boldsymbol{x}^{(0)} - \boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}(\boldsymbol{x}^{(0)}) + (I - \boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}'(\boldsymbol{X}_\rho))(\boldsymbol{X}_\rho - \boldsymbol{x}^{(0)}) \subset \boldsymbol{X}_\rho$$

holds for any ρ satisfying the inequality

$$\frac{1-\sqrt{1-4h}}{2h}\eta \leq \rho \leq \frac{1+\sqrt{1-4h}}{2h}\eta.$$

Since the alpha theory and Rump's interval theorem are respectively the generalization of the Kantorovich theorem and Moore's interval theorem, we generalize Rall's conclusion to discuss the relationship between the alpha theory and Rump's interval theorem in this paper. By only the information of the given system at the initial approximate point, we provide the conditions to guarantee that we can obtain an approximate point after finite Newton's iterations, where this approximate iteration point corresponds to an interval solution satisfying the condition of Rump's interval theorem. The next section will give some notation and background results.

2   Notation and Preliminaries

First of all, we emphasize that the norm $\|\cdot\|$ of the vector and the matrix in this paper are both the infinite norm $\|\cdot\|_\infty$ since the metric for the interval vector is closely related to the infinite norm.

Henceforward, we use boldface letters to express tuples and denote their entries by the same letter with subscripts, for example $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_n)^T$. Denote the usual product order on $\mathbb{R}^n$ by $\leq$, that is, for arbitrary $\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{R}^n$, $\boldsymbol{\alpha} \leq \boldsymbol{\beta}$ if and only if $\alpha_i \leq \beta_i$ for $1 \leq i \leq n$.

For $\boldsymbol{x} \in \mathbb{R}^n$, if $\boldsymbol{f}'(\boldsymbol{x})$ is nonsingular, then define

$$\alpha(\boldsymbol{f},\boldsymbol{x}) = \beta(\boldsymbol{f},\boldsymbol{x})\gamma(\boldsymbol{f},\boldsymbol{x}),$$
$$\beta(\boldsymbol{f},\boldsymbol{x}) = \|\boldsymbol{f}'(\boldsymbol{x})^{-1}\boldsymbol{f}(\boldsymbol{x})\|,$$
$$\gamma(\boldsymbol{f},\boldsymbol{x}) = \sup_{k\geq 2} \|\boldsymbol{f}'(\boldsymbol{x})^{-1}\frac{\boldsymbol{f}^{(k)}(\boldsymbol{x})}{k!}\|^{1/(k-1)}.$$

Given the initial approximate $\boldsymbol{x}^{(0)}$ with the associated simple root $\boldsymbol{x}^*$ of $\boldsymbol{f}$, we let α, β and γ to stand for $\alpha(\boldsymbol{f},\boldsymbol{x}^{(0)})$, $\beta(\boldsymbol{f},\boldsymbol{x}^{(0)})$ and $\gamma(\boldsymbol{f},\boldsymbol{x}^{(0)})$, respectively. Applying Newton's method for $\boldsymbol{f}$ can get the Newton iteration sequence $\{\boldsymbol{x}^{(k)}\}$, that is,

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \boldsymbol{f}'(\boldsymbol{x}^{(k)})^{-1}\boldsymbol{f}(\boldsymbol{x}^{(k)}), \qquad k \in \mathbb{N}.$$

The alpha theory provides the convergence conditions which ensure the sequence $\{\boldsymbol{x}^{(k)}\}$ converges to $\boldsymbol{x}^*$ only with the values α, β and γ. The dominating sequence technique is a powerful tool for improving the alpha theory. The dominating sequence $\{t^{(k)}\}$ is produced by the Newton iteration with the initial approximate $t^{(0)} = 0$ for the univariate function

$$h(t) = \beta - t + \frac{\gamma t^2}{1-\gamma t},$$

where the equation $h(t) = 0$ has the following two solutions

$$t^* = \frac{2\beta}{1 + \alpha + \sqrt{1 - 6\alpha + \alpha^2}}, \quad t^{**} = \frac{1 + \alpha + \sqrt{1 - 6\alpha + \alpha^2}}{4\gamma}. \tag{1}$$

The following theorem is a version of the alpha theory given by Wang and Han, where the condition on the quantity α is best possible.

Theorem 1. [4,18] *If $0 < \alpha < 3 - 2\sqrt{2}$, then for any $t^* \le \rho < t^{**}$, the system $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ has exactly one simple root $\boldsymbol{x}^*$ in $\overline{B}(\boldsymbol{x}^{(0)}, \rho)$. In addition, the Newton iteration sequence $\{\boldsymbol{x}^{(k)}\}$ converges quadratically to $\boldsymbol{x}^*$, and the dominating sequence $\{t^{(k)}\}$ increases and converges quadratically to t^*. Furthermore, for all $k \in \mathbb{N}$,*

$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\| \le t^{(k+1)} - t^{(k)},$$

$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\| \le q(\alpha)^{2^k - 1}\beta$$

with

$$q(\alpha) = \frac{4\alpha}{\left(1 - \alpha + \sqrt{1 - 6\alpha + \alpha^2}\right)^2}. \tag{2}$$

Denote the set of intervals by $\mathbb{IR}$. An interval vector $\mathbf{X} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}] \in \mathbb{IR}^n$ with $\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}} \in \mathbb{R}^n$ and $\underline{\boldsymbol{x}} \le \overline{\boldsymbol{x}}$ is defined by

$$\mathbf{X} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}] = \{\boldsymbol{x} \in \mathbb{R}^n : \underline{\boldsymbol{x}} \le \boldsymbol{x} \le \overline{\boldsymbol{x}}\}.$$

For $\boldsymbol{x} \in \mathbb{R}^n$, $\mathbf{X} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}] \in \mathbb{IR}^n$, $\boldsymbol{x} + \mathbf{X} = [\boldsymbol{x} + \underline{\boldsymbol{x}}, \boldsymbol{x} + \overline{\boldsymbol{x}}]$. Let $\mathbf{Y}_\rho = \{\boldsymbol{y} \in \mathbb{R}^n : \|\boldsymbol{y}\| \le \rho\}$, then $\overline{B}(\boldsymbol{x}, \rho) = \boldsymbol{x} + \mathbf{Y}_\rho$.

The norm of the interval vector $\mathbf{X} = [\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}]$ is defined by

$$\|\mathbf{X}\| = \|[\underline{\boldsymbol{x}}, \overline{\boldsymbol{x}}]\| = \max\{\|\boldsymbol{x}\| : \boldsymbol{x} \in \mathbf{X}\}.$$

Besides $\mathrm{int}(\mathbf{X})$ designates the interior of the interval vector $\mathbf{X}$. Given a set $\mathbf{Z} \subset \mathbb{R}^n$, the interval hull of $\mathbf{Z}$ is the narrowest interval vector containing $\mathbf{Z}$, namely,

$$\mathrm{hull}(\mathbf{Z}) = \bigcap\{\mathbf{X} \in \mathbb{IR}^n : \mathbf{X} \supseteq \mathbf{Z}\}.$$

Given a continuous mapping $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^m$ and an interval vector $\mathbf{X}$, the interval vector $\boldsymbol{g}(\mathbf{X}) \in \mathbb{IR}^m$ is defined as

$$\boldsymbol{g}(\mathbf{X}) = \mathrm{hull}\{\boldsymbol{g}(\boldsymbol{x}) : \boldsymbol{x} \in \mathbf{X}\}.$$

Given an interval matrix $\mathbf{A} \in \mathbb{IR}^{m \times n}$ and an interval vector $\mathbf{X} \in \mathbb{IR}^n$, the interval vector $\mathbf{AX}$ is defined by

$$\mathbf{AX} = \mathrm{hull}\{A\boldsymbol{x} : A \in \mathbf{A}, \boldsymbol{x} \in \mathbf{X}\}.$$

Specially, the norm of the interval matrix $\mathbf{A}$ is defined as

$$\|\mathbf{A}\| = \max\{\|A\| : A \in \mathbf{A}\}.$$

The following theorem is a version of the interval Newton's method given by Rump.

Theorem 2. [12] *Given $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$, $\mathbf{Y} \in \mathbb{IR}^n$ with $\mathbf{0} \in \mathbf{Y}$, $R \in \mathbb{R}^{n \times n}$, if*

$$S(\mathbf{Y}, \boldsymbol{x}^{(0)}) := -R\boldsymbol{f}(\boldsymbol{x}^{(0)}) + (I - R\boldsymbol{f}'(\boldsymbol{x}^{(0)} + \mathbf{Y}))\mathbf{Y} \subseteq \mathrm{int}(\mathbf{Y}),$$

then there exists a unique $\boldsymbol{x}^* \in \boldsymbol{x}^{(0)} + \mathbf{Y}$ such that $\boldsymbol{f}(\boldsymbol{x}^*) = \mathbf{0}$.

3   Main Results

To give the main results of this paper, we need the following functions,

$$\psi(u) = 2u^2 - 4u + 1,$$
$$g_1(\alpha) = \alpha^2 - 4\alpha + 10,$$
$$g_2(\alpha) = \alpha^3 - 6\alpha^2 + 21\alpha + 28,$$
$$g_3(\alpha) = 4\alpha^3 - 25\alpha^2 + 88\alpha - 8,$$
$$\theta(\alpha) = \frac{1}{3} \arccos(\frac{g_2(\alpha)}{\sqrt{g_1(\alpha)^3}}),$$
$$\omega_1(\alpha) = \sqrt{g_1(\alpha)} \cos(\theta(\alpha) + \frac{2\pi}{3}) + 2 + \frac{\alpha}{2},$$
$$\omega_2(\alpha) = \sqrt{g_1(\alpha)} \cos(\theta(\alpha) + \frac{4\pi}{3}) + 2 + \frac{\alpha}{2},$$
$$\omega_3(\alpha) = \sqrt{g_1(\alpha)} \cos(\theta(\alpha)) + 2 + \frac{\alpha}{2}.$$

Here we give some lemmas and one proposition from which the main theorem will easily follow.

Lemma 1. [15] *Given $\widetilde{\boldsymbol{x}} \in \mathbb{R}^n$, if $\gamma\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}^{(0)}\| < 1 - \sqrt{2}/2$, then*

$$\gamma(\boldsymbol{f}, \widetilde{\boldsymbol{x}}) \leq \frac{\gamma}{\psi(\gamma\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}^{(0)}\|)(1 - \gamma\|\widetilde{\boldsymbol{x}} - \boldsymbol{x}^{(0)}\|)}.$$

Lemma 2. *If $0 < \alpha < 3 - 2\sqrt{2}$, then for all $k \geq 1$,*

$$\gamma\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\| < 1 - \frac{\sqrt{2}}{2}, \tag{3}$$

$$\gamma(f, \boldsymbol{x}^{(k)}) \leq \frac{\gamma}{\psi(\gamma\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\|)(1 - \gamma\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\|)}. \tag{4}$$

Proof. If $0 < \alpha < 3 - 2\sqrt{2}$, then by Theorem 1,

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\| \leq \sum_{j=1}^{k} \|\boldsymbol{x}^{(j)} - \boldsymbol{x}^{(j-1)}\| \leq \sum_{j=1}^{k} (t^{(j)} - t^{(j-1)})$$
$$\leq t^{(k)} \leq t^*.$$

Therefore
$$\gamma\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\| \leq \frac{2\alpha}{1 + \alpha + \sqrt{1 - 6\alpha + \alpha^2}}.$$

Since the right hand side of the above inequality monotonously increases from 0 to $1 - \sqrt{2}/2$ as α goes from 0 to $3 - 2\sqrt{2}$, it follows that (3) holds. By means of Lemma 1, (4) follows.   □

Lemma 3. *If*

$$0 < \rho < \frac{1 - \sqrt{2}/2}{\gamma}, \tag{5}$$

$$\beta + \rho(\frac{1}{(1 - \gamma\rho)^2} - 1) < \rho, \tag{6}$$

then

$$S(\boldsymbol{Y}_\rho, \boldsymbol{x}^{(0)}) \subseteq \operatorname{int}(\boldsymbol{Y}_\rho). \tag{7}$$

Proof.   Given an arbitrary real vector $\boldsymbol{y} \in \boldsymbol{Y}_\rho$, we expand the Jacobian matrix $\boldsymbol{f}'(\boldsymbol{x}^{(0)} + \boldsymbol{y})$ into power series and get

$$\boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}'(\boldsymbol{x}^{(0)} + \boldsymbol{y}) = \boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}(\boldsymbol{f}'(\boldsymbol{x}^{(0)}) + \sum_{k=2}^{\infty}\boldsymbol{f}^{(k)}(\boldsymbol{x}^{(0)})\frac{\boldsymbol{y}^{k-1}}{(k-1)!})$$

$$= I + \sum_{k=2}^{\infty}k\boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}^{(k)}(\boldsymbol{x}^{(0)})\frac{\boldsymbol{y}^{k-1}}{k!}.$$

If (5) holds, then

$$\|I - \boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}'(\boldsymbol{x}^{(0)} + \boldsymbol{y})\| \leq \sum_{k=2}^{\infty}k\|\boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\frac{\boldsymbol{f}^{(k)}(\boldsymbol{x}^{(0)})}{k!}\|\|\boldsymbol{y}\|^{k-1}$$

$$\leq \sum_{k=2}^{\infty}k(\gamma\rho)^{k-1}$$

$$= \frac{1}{(1 - \gamma\rho)^2} - 1.$$

Suppose that (5) (6) hold, then for an arbitrary real vector $\boldsymbol{y} \in \boldsymbol{X}_\rho$, we can infer that

$$\| - \boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}'(\boldsymbol{x}^{(0)}) + (I - \boldsymbol{f}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{f}'(\boldsymbol{x}^{(0)} + \boldsymbol{y}))\boldsymbol{y}\| < \rho,$$

which implies that (7) holds.   □

Lemma 4. *Let*

$$\alpha^* = -\frac{1}{12}(22247 + 1320\sqrt{330})^{1/3} + \frac{431}{12(22247 + 1320\sqrt{330})^{1/3}} + \frac{25}{12} \tag{8}$$

$$\approx 0.093347623,$$

then for an arbitrary $0 < \alpha < \alpha^*$, we have

$$g_1(\alpha^*) < g_1(\alpha) < g_1(0),$$
$$g_2(0) < g_2(\alpha) < g_2(\alpha^*),$$
$$g_3(0) < g_3(\alpha) < g_3(\alpha^*) = 0,$$
$$\theta(\alpha^*) < \theta(\alpha) < \theta(0),$$
$$\omega_1(0) < \omega_1(\alpha) < \omega_1(\alpha^*),$$
$$\omega_2(\alpha^*) < \omega_2(\alpha) < \omega_2(0) < 3 - \frac{3\sqrt{2}}{2},$$
$$\omega_3(\alpha) > 3 - \frac{3\sqrt{2}}{2}.$$

Proof. According to Cartan's root-finding formula, the equation $g_3(\alpha) = 0$ has only a positive real root α^* as in (8). Obviously, $g_1'(\alpha) < 0$ and $g_2'(\alpha) > 0$ for all $0 < \alpha < \alpha^*$. Thus both $\theta(\alpha)$ and $\omega_2(\alpha)$ monotonously decrease on the interval $(0, 3 - 2\sqrt{2})$. A routine computation gives rise to

$$\omega_1'(\alpha) = \frac{(4 - 2\alpha)\sqrt{-3g_3(\alpha)}\sin(\theta(\alpha) + \frac{\pi}{6}) + (84 - 36\alpha)\cos(\theta(\alpha) + \frac{\pi}{6})}{2\sqrt{-3g_1(\alpha)g_3(\alpha)}} + \frac{1}{2},$$

then for all $0 < \alpha < \alpha^*$, $\omega_1'(\alpha) > 0$. This lemma follows immediately from what we have proved. $\square$

Proposition 1. *Under the condition (5), the inequality (6) holds if and only if*

$$0 < \alpha < \alpha^*, \tag{9}$$

$$\frac{\omega_1(\alpha)}{3\gamma} < \rho < \frac{\omega_2(\alpha)}{3\gamma}. \tag{10}$$

Proof. Define $\Delta = (q/2)^2 + (p/3)^3$ with

$$p = -\frac{1}{3}\left(\frac{4 + \beta\gamma}{-2\gamma}\right)^2 + \frac{1 + 2\beta\gamma}{2\gamma^2},$$
$$q = 2\left(\frac{4 + \beta\gamma}{-6\gamma}\right)^3 + \frac{\beta}{2\gamma^2} - \frac{(4 + \beta\gamma)(1 + 2\beta\gamma)}{-12\gamma^3},$$

then

$$p = -\frac{g_1(\alpha)}{12\gamma^2}, \quad q = -\frac{g_2(\alpha)}{108\gamma^3}, \quad \Delta = \frac{g_3(\alpha)}{1728\gamma^6}.$$

It follows by Lemma 4 that for all $\alpha > 0$, $p < 0$ and $q < 0$, then we have three cases to consider.

Case I: $\alpha > \alpha^*$. In this case, $\Delta > 0$ and the equation

$$\rho^3 - \frac{4 + \beta\gamma}{2\gamma}\rho^2 + \frac{1 + 2\beta\gamma}{2\gamma^2}\rho - \frac{\beta}{2\gamma^2} = 0 \tag{11}$$

has only one real solution

$$\rho^* = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}.$$

An easy computation yields that for all $\alpha > \alpha^*$, $\rho^* > (1 - \sqrt{2}/2)/\gamma$. Therefore, in this case, (6) holds if and only if $\rho > \rho^*$, which contradicts the condition (5).

Case II: $\alpha = \alpha^*$. In this case, $\Delta = 0$ and (11) has only two unequal real solutions

$$-\sqrt[3]{\frac{-q}{2}}, \quad 2\sqrt[3]{\frac{-q}{2}}.$$

Clearly,

$$-\sqrt[3]{\frac{-q}{2}} < 0, \quad 2\sqrt[3]{\frac{-q}{2}} > \frac{1 - \sqrt{2}/2}{\gamma}.$$

Hence in this case, (6) can not hold under the condition (5).

Case III: $0 < \alpha < \alpha^*$. In this case, $\Delta < 0$ and (11) has three unequal real solutions

$$\frac{\omega_1(\alpha)}{3\gamma}, \quad \frac{\omega_2(\alpha)}{3\gamma}, \quad \frac{\omega_3(\alpha)}{3\gamma}.$$

Recalling Lemma 4, we know that for all $0 < \alpha < \alpha^*$,

$$\frac{\omega_3(\alpha)}{3\gamma} > \frac{1 - \sqrt{2}/2}{\gamma},$$

$$\frac{\omega_2(\alpha)}{3\gamma} < \frac{1 - \sqrt{2}/2}{\gamma},$$

$$0 < \frac{\omega_1(\alpha)}{3\gamma} < \frac{\omega_2(\alpha)}{3\gamma}.$$

Thus in this case, (6) holds if and only if (10) holds.

As a whole, under the condition (5), the inequality (6) holds if and only if (9) and (10) hold.   □

On the basis of α, β and γ, the following theorem provides the conditions such that (7) holds, which is an immediate conclusion of Proposition 1.

Theorem 3. *If $0 < \alpha < \alpha^*$, then for any ρ satisfying the inequality*

$$\frac{\omega_1(\alpha)}{3\gamma} < \rho < \frac{\omega_2(\alpha)}{3\gamma},$$

the condition (7) holds.

The following corollary indicates that the alpha theory is of greater precision than Rump's interval theorem, which can be directly deduced by an easy computation.

Corollary 1. *If $0 < \alpha < \alpha^*$, then*

$$\frac{\omega_1(\alpha)}{3\gamma} > t^*.$$

For the Newton iteration sequence $\{\boldsymbol{x}^{(k)}\}$, define

$$\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(k)}) = \frac{\omega_1(\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}))}{3\gamma(\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}))}, \quad \rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(k)}) = \frac{\omega_2(\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}))}{3\gamma(\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}))}. \tag{12}$$

In view of the quantity α of the alpha theory proposed by Wang and Han [18], we give the following convergence condition of interval Newton's algorithm proposed by Rump.

Proposition 2. *Let $\lceil \cdot \rceil$ be the integer ceiling function, $p(\alpha)$ be defined by*

$$p(\alpha) = \frac{2\alpha}{1 + \alpha + \sqrt{1 - 6\alpha + \alpha^2}},$$

and $q(\alpha)$ be defined in (2). If $0 < \alpha < 3 - 2\sqrt{2}$, then for any

$$k \geq \lceil \log_2(\frac{\ln \alpha^* + \ln(1 - p(\alpha)) + \ln \psi(p(\alpha)) - \ln \alpha}{\ln q(\alpha)} + 1)\rceil, \tag{13}$$

the condition

$$S(\boldsymbol{Y}_{\rho^{(k)}}, \boldsymbol{x}^{(k)}) \subseteq \text{int}(\boldsymbol{Y}_{\rho^{(k)}}) \tag{14}$$

holds for any $\rho^{(k)}$ satisfying the inequality

$$\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(k)}) < \rho^{(k)} < \rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(k)}). \tag{15}$$

Proof. By Theorem 1 and Lemma 2, we know that if $0 < \alpha < 3 - 2\sqrt{2}$, then for any $k \in \mathbb{N}$,

$$\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}) \leq \frac{\alpha q(\alpha)^{2^k - 1}}{\psi(\gamma \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\|)(1 - \gamma \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\|)}.$$

If $0 < \alpha < 3 - 2\sqrt{2}$, then for all $k \in \mathbb{N}$,

$$\gamma \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\| \leq p(\alpha) < 1 - \frac{\sqrt{2}}{2},$$

which implies that for all $k \in \mathbb{N}$,

$$\psi(\gamma \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\|)(1 - \gamma \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(0)}\|) \geq \psi(p(\alpha))(1 - p(\alpha)).$$

Hence if $0 < \alpha < 3 - 2\sqrt{2}$, then for all $k \in \mathbb{N}$,

$$\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}) \leq \frac{\alpha q(\alpha)^{2^k - 1}}{\psi(p(\alpha))(1 - p(\alpha))},$$

which implies that $0 < \alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}) < \alpha^*$ holds if the iteration number k satisfies the inequality (13). Our conclusion will follow from Theorem 3.   □

With the aid of the quantity α of the alpha theory given in [1], the conclusion of the above proposition can be improved as follows.

Proposition 3. *If $0 < \alpha \leq (13 - 3\sqrt{17})/4$, then for any $k \geq 3$, the condition (14) holds for any $\rho^{(k)}$ satisfying the inequality (15).*

Proof. Since the function $\alpha/\psi(\alpha)^2$ monotonously increases on the interval $[0, 1 - \sqrt{2}/2)$, it follows that $\alpha/\psi(\alpha)^2 < 1$ for any $0 < \alpha \leq (13 - 3\sqrt{17})/4$. Recalling Proposition 1 in [15], we can deduce that if $0 < \alpha \leq (13 - 3\sqrt{17})/4$, then for any $k \in \mathbb{N}$,

$$\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}) \leq \left(\frac{\alpha}{\psi(\alpha)^2}\right)^{2^k - 1} \alpha.$$

As a result, if $0 < \alpha \leq (13 - 3\sqrt{17})/4$, then $0 < \alpha(\boldsymbol{f}, \boldsymbol{x}^{(3)}) < \alpha^*$.   □

4   Example

In this section, we propose some examples to illustrate our conclusion, which are done in Matlab R2012a with INTLAB V6 under Windows 7. In these examples, the true interval of the display as $-0.0059_____$ is obtained by subtracting and adding 1 to the last displayed digit, namely,

$$-0.0059_____ = [-0.00600000000000, -0.00580000000000].$$

Example 1. Let

$$f_1 = x_1^2 + x_2 - 2 = 0,$$
$$f_2 = x_1 - x_2 = 0,$$

and $\boldsymbol{x}^{(0)} = (1.8, 1.8)^T$. The program of the alpha theory computes

$$\alpha^* < \alpha = 0.1436672967 < \frac{13 - 3\sqrt{17}}{4},$$

then by Proposition 3, we can immediately deduce that $0 < \alpha(\boldsymbol{f}, \boldsymbol{x}^{(3)}) < \alpha^*$. The values of $\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)})$, $\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(k)})$, $\rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(k)})$, $k = 1, 2, 3$, are shown in Table 1, and the values of $\boldsymbol{x}^{(k)}$, $\rho^{(k)}$, $S(\mathbf{Y}_{\rho^{(k)}}, \boldsymbol{x}^{(k)})$, $k = 1, 2, 3$, are shown in Table 2.

Example 2. [11] Let

$$f_i(\boldsymbol{x}) = x_i - 0.7 x_i \sum_{j=1}^{9} a_{i,j} x_j - 1 = 0, \qquad i = 1, 2 \ldots, 9,$$

with

$$a_{i,j} = \frac{3}{4} \frac{t_i (1 - t_j^2)^2 w_j}{t_i + t_j},$$

Table 1. The values of $\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)})$, $\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(k)})$ and $\rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(k)})$ about Example 1

k	$\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)})$	$\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(k)})$	$\rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(k)})$
1	0.04063913776	0.1474393618	0.8654682610
2	0.001956685388	0.005916476543	0.8785582303
3	0.000003858607108	0.00001157539288	0.8785582303

Table 2. The values of $\boldsymbol{x}^{(k)}$, $\rho^{(k)}$ and $S(\mathbf{Y}_{\rho^{(k)}}, \boldsymbol{x}^{(k)})$ about Example 1

k	$\boldsymbol{x}^{(k)}$	$\rho^{(k)}$	$S(\mathbf{Y}_{\rho^{(k)}}, \boldsymbol{x}^{(k)})$
1	$\begin{pmatrix} 1.13913043490000 \\ 1.13913043490000 \end{pmatrix}$	0.147440361800000	$\begin{pmatrix} [-0.14648800759416, -0.11996338231131] \\ [-0.14648800759416, -0.11996338231131] \end{pmatrix}$
2	$\begin{pmatrix} 1.00590473980000 \\ 1.00590473980000 \end{pmatrix}$	0.005917476543000	$\begin{pmatrix} -0.0059\text{_____} \\ -0.0059\text{_____} \end{pmatrix}$
3	$\begin{pmatrix} 1.00001157619900 \\ 1.00001157619900 \end{pmatrix}$	1.257539288000000e-05	$1.0\text{e-}005 * \begin{pmatrix} -0.11576\text{_____} \\ -0.11576\text{_____} \end{pmatrix}$

where $t_i, w_i, i = 1, 2, \ldots, 9$, are respectively the nodes and weights of Gaussian integration rule of order 17 on the interval $[0, 1]$ (See Table 3).

For the initial approximate

$$\boldsymbol{x}^{(0)} = [1.36, 1.36, 1.36, 1.36, 1.36, 1.36, 1.36, 1.36, 1.36]^T,$$

we have

$$\alpha^* < \alpha = 0.144327213206543 < \frac{13 - 3\sqrt{17}}{4}.$$

It follows by Proposition 3 that $0 < \alpha(\boldsymbol{f}, \boldsymbol{x}^{(3)}) < \alpha^*$. To be precise,

$$\alpha(\boldsymbol{f}, \boldsymbol{x}^{(3)}) = 9.771724278220212\text{e-}10,$$

$$\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(3)}) = 2.106476714805001\text{e-}09,$$

$$\rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(3)}) = 0.721018979786819.$$

Choose $\rho^{(3)} = 2.206476714805001\text{e-}09$, then $S(\mathbf{Y}_{\rho^{(3)}}, \boldsymbol{x}^{(3)}) \subseteq \text{int}(\mathbf{Y}_{\rho^{(3)}})$. The values of $\boldsymbol{x}^{(3)}$ and $S(\mathbf{Y}_{\rho^{(3)}}, \boldsymbol{x}^{(3)})$ are shown in Table 4, where i stands for the coordinate index.

Example 3. Let

$$f_i = x_i^2 + x_{i+1} - 2 = 0, \qquad i = 1, 2, \ldots, 99,$$
$$f_{100} = x_{99} - x_{100} = 0.$$

Given $\boldsymbol{x}^{(0)}$ with $x_i^{(0)} = 1.28$, $i = 1, 2, \ldots, 100$, it follows that

$$\frac{13 - 3\sqrt{17}}{4} < \alpha = 0.165370210314031 < 3 - 2\sqrt{2}.$$

Table 3. The values of t_i, w_i of Example 2

i	t_i	w_i
1	0.015919880000000	0.040637193262940
2	0.081984445000000	0.090324080584283
3	0.193314285000000	0.130305351576427
4	0.337873290000000	0.156173536255424
5	0.500000000000000	0.165119676661738
6	0.662126710000000	0.156173536254610
7	0.806685715000000	0.130305351576465
8	0.918015555000000	0.090324080583571
9	0.984080120000000	0.040637193262741

Table 4. The values of $\boldsymbol{x}^{(3)}$ and $S(\mathbf{Y}_{\rho^{(3)}}, \boldsymbol{x}^{(3)})$ about Example 2

i	$\boldsymbol{x}^{(3)}$	$S(\mathbf{Y}_{\rho^{(3)}}, \boldsymbol{x}^{(3)})$
1	1.03266743259998	1.0e-008 * -0.2175469960____
2	1.10583043469589	1.0e-008 * 0.058000438_____
3	1.17693975483393	1.0e-008 * 0.074880771_____
4	1.23474234890867	1.0e-008 * -0.046156608_____
5	1.27801354526920	1.0e-008 * 0.033335094_____
6	1.30888887603008	1.0e-008 * -0.052842004_____
7	1.32995480221212	1.0e-008 * 0.015004650_____
8	1.34328756776955	1.0e-008 * -0.029283159_____
9	1.35027189327776	1.0e-008 * 0.057999652_____

Recalling Proposition 2, we know that there exists $k \in \mathbb{N}$ such that $\alpha(\boldsymbol{f}, \boldsymbol{x}^{(k)}) < \alpha^*$. Indeed, for the first-step iteration point $\boldsymbol{x}^{(1)}$, we have

$$\alpha(\boldsymbol{f}, \boldsymbol{x}^{(1)}) = 0.0209408111113277 < \alpha^*,$$
$$\rho^*(\boldsymbol{f}, \boldsymbol{x}^{(1)}) = 0.0229019461130693,$$
$$\rho^{**}(\boldsymbol{f}, \boldsymbol{x}^{(1)}) = 0.291551749088838.$$

Choose $\rho^{(1)} = 0.0229019461130694$, then $S(\mathbf{Y}_{\rho^{(1)}}, \boldsymbol{x}^{(1)}) \subseteq \text{int}(\mathbf{Y}_{\rho^{(1)}})$.

References

1. Blum, L., Cucker, F., Shub, M., Smale, S.: Complexity and Real Computation. Springer, New York (1998)
2. Kantorovich, L.V.: Functional analysis and applied mathematics. Uspehi. Mat. Nauk. **3**(6), 89–185 (1948)

3. Giusti, M., Lecerf, G., Salvy, B., Yakoubsohn, J.C.: On location and approximation of clusters of zeros of analytic functions. Found. Comput. Math. **5**(3), 257–311 (2005)
4. Giusti, M., Lecerf, G., Salvy, B., Yakoubsohn, J.C.: On location and approximation of clusters of zeros: case of embedding dimension one. Found. Comput. Math. **7**(1), 1–58 (2007)
5. Hauenstein, J.D., Levandovskyy, V.: Certifying solutions to square systems of polynomial-exponential equations. J. Symb. Comput. **79**(3), 575–593 (2015)
6. Hauenstein, J.D., Sottile, F.: Algorithm 921: alphacertified: certifying solutions to polynomial systems. ACM Trans. Math. Softw. **38**(4), 1–20 (2011)
7. Li, N., Zhi, L.: Verified error bounds for isolated singular solutions of polynomial systems: case of breadth one. Theor. Comput. Sci. **479**, 163–173 (2013)
8. Li, N., Zhi, L.: Verified error bounds for isolated singular solutions of polynomial systems. SIAM J. Numer. Anal. **52**(4), 1623–1640 (2014)
9. Krawczyk, R.: Newton-algorithmen zur bestimmung von nullstellen mit fehlerschranken. Computing **4**, 187–201 (1969)
10. Moore, R.E.: A test for existence of solutions to nonlinear system. SIAM J. Numer. Anal. **14**(4), 611–615 (1977)
11. Rall, L.B.: A comparison of the existence theorems of Kantorvich and Moore. SIAM J. Numer. Anal. **17**(1), 148–161 (1980)
12. Rump, S.M.: Solving algebraic problems with high accuracy. In: Kulisch, W.L., Miranker, W.L. (eds.) A New Approach to Scientific Computation, pp. 51–120. Academic Press, San Diego (1983)
13. Rump, S.M.: INTLAB-Interval Laboratory. Springer, Netherlands, Berlin (1999)
14. Rump, S.M., Graillat, S.: Verified error bounds for multiple roots of systems of nonlinear equations. Numer. Algorithm **54**, 359–377 (2009)
15. Smale, S.: Newton method estimates from data at one point. The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, pp. 185C–196C. Springer, Berlin (1986)
16. Shub, M., Smale, S.: Computational complexity: on the geometry of polynomials and a theory of cost. I. Ann. Sci. Ècole Norm. Sup. **18**(1), 107–142 (1985)
17. Shub, M., Smale, S.: Computational complexity: on the geometry of polynomials and a theory of cost. II. SIAM J. Comput. **15**(1), 145–161 (1986)
18. Wang, X.H., Han, D.F.: On dominating sequence method in the point estimate and Smale theorem. Sci. China Ser. A **33**(2), 135–144 (1990)
19. Yang, Z., Zhi, L., Zhu, Y.: Verfied error bounds for real solutions of positive-dimensional polynomial systems. In: The 2013 International Symposium on Symbolic and Algebraic Computation, pp. 371–378. ACM press, San Jose (2013)