MATHEMATICS OF COMPUTATION
Volume 89, Number 322, March 2020, Pages 879–909
https://doi.org/10.1090/mcom/3479
Article electronically published on October 1, 2019

ON ISOLATION OF SIMPLE MULTIPLE ZEROS AND CLUSTERS OF ZEROS OF POLYNOMIAL SYSTEMS

ZHIWEI HAO, WENRONG JIANG, NAN LI, AND LIHONG ZHI

Abstract. Given a well-constrained polynomial system f associated with a simple multiple zero x of multiplicity μ, we give a computable separation bound for isolating x from the other zeros of f. When x is only given with a limited accuracy, we give a numerical criterion for isolating a nearby cluster of μ zeros of f (counting multiplicities) in a ball around x.

1. Introduction

There are two challenging problems in solving polynomial systems with singular zeros: computing the local separation bound of an exactly given singular zero and isolating a cluster of zeros near an approximately given singular zero.

Definition 1.1. Given a well-constrained[1] polynomial system $f = \{f_1, \ldots, f_n\}$, where $f_i \in \mathbb{C}[X_1, \ldots, X_n]$ for $i = 1, \ldots, n$, then $x \in \mathbb{C}^n$ is an isolated zero of f of multiplicity μ if it satisfies the following three conditions:

(1) $f(x) = 0$,
(2) there exists a ball $B(x, r)$ of radius $r > 0$ such that $B(x, r) \cap f^{-1}(0) = \{x\}$,
(3) a generic analytic function g sufficiently close to f possesses μ simple zeros in $B(x, r)$.

The last condition is related to Rouché's Theorem [2, Theorem 2.12] which shows that any analytic g satisfying

$$\|f(y) - g(y)\| < \|f(y)\| \quad \forall\, y \in \partial B(x, r) := \{y \in \mathbb{C}^n \mid \|y - x\| = r\}, \tag{1.1}$$

has a finite number of zeros in $B(x, r)$ and the sum of multiplicities of zeros is μ.

According to Definition 1.1, the problem of computing the local separation bound of an exactly given singular zero x is to determine an upper bound of r such that $B(x, r)$ isolates x from the other zeros of f. The problem of isolating a cluster of zeros of f near an approximately given singular zero x is to construct a nearby polynomial system g such that it possesses x as its isolated zero of multiplicity μ and satisfies (1.1), which implies that f has a cluster of μ zeros in $B(x, r)$ according to Rouché's Theorem.

Received by the editor August 20, 2018, and, in revised form, February 21, 2019, and May 6, 2019.

2010 *Mathematics Subject Classification.* Primary 65H10, 74G35, 68W30, 32-04, 32S99.

This work was carried out when the third author was with Tianjin University. This work was supported in part by the National Key Research Project of China 2018YFA0306702 (Zhi) and the National Natural Science Foundation of China 11571350 (Zhi), 11601378 (Li).

The third author is the corresponding author.

[1]We always assume f is a well-constrained polynomial system throughout the paper.

Our contributions. Given a well-constrained polynomial system $f = \{f_1, \ldots, f_n\}$, where $f_i \in \mathbb{C}[X_1, \ldots, X_n]$ for $i = 1, \ldots, n$, we generalize Dedieu and Shub's quantitative results for simple double zeros [10] to simple multiple zeros of arbitrary high multiplicities.

Let $Df(x)$ denote the Jacobian matrix of f at a point $x \in \mathbb{C}^n$. Given an isolated zero x of f of multiplicity μ, if $\operatorname{corank}(Df(x)) = 1$, then its multiplicity structure can be described by a closed basis $\{\Lambda_0, \Lambda_1, \ldots, \Lambda_{\mu-1}\}$ of the local dual space of f at x; see Section 2.1. Let Δ_k be the nonlinear component of the kth order differential functional Λ_k.

Definition 1.2. A point $x \in \mathbb{C}^n$ is a simple multiple zero of a polynomial system $f : \mathbb{C}^n \to \mathbb{C}^n$ of multiplicity μ, if

(A) $f(x) = 0$,
(B) $\dim \ker Df(x) = 1$,
(C) $\Delta_k(f) \in \operatorname{im} Df(x)$, for $k = 2, \ldots, \mu - 1$,
(D) $\Delta_\mu(f) \notin \operatorname{im} Df(x)$.

Definition 1.3. Suppose x is a simple multiple zero of f. Then $Df(x)$ is of normal form if

$$Df(x) = \begin{pmatrix} 0 & D\hat{f}(x) \\ 0 & 0 \end{pmatrix}, \tag{1.2}$$

where $D\hat{f}(x)$ is the invertible Jacobian matrix at x of the polynomials $\hat{f} = \{f_1, \ldots, f_{n-1}\}$ with respect to the variables $\hat{X} = \{X_2, \ldots, X_n\}$.

If x is a simple multiple zero of f and $Df(x)$ is of normal form, then

$$\Delta_k(f) \in \operatorname{im} Df(x) = \operatorname{im} \begin{pmatrix} D\hat{f}(x) \\ 0 \end{pmatrix} \Leftrightarrow \Delta_k(f_n) = 0.$$

Therefore, the conditions (C), (D) in Definition 1.2 are equivalent to

(C) $\Delta_k(f_n) = 0$, for $k = 2, \ldots, \mu - 1$,
(D) $\Delta_\mu(f_n) \neq 0$.

We show in Section 2.2 that it is always possible to perform a unitary transformation to obtain an equivalent polynomial system such that its Jacobian matrix at the simple multiple zero is of normal form. Since the unitary transformation does not change the distance between two zeros of f, without loss of generality, we can always assume condition (1.2) is satisfied.

Our major results are summarized as follows:

- We propose a local separation bound for isolating a simple multiple zero x of multiplicity μ from the other zeros of f. Let y be another zero of f. Then

$$\|y - x\| \geq \frac{d}{2\gamma_\mu(f, x)^\mu},$$

where d is the smallest positive real root of a univariate equation that relies only on μ (see (3.11) and (3.23)) and the constants

$$\begin{aligned}
\gamma_\mu(f,x) &= \max(\hat{\gamma}_\mu(f,x), \gamma_{\mu,n}(f,x)),\\
\hat{\gamma}_\mu(f,x) &= \max\left(1,\ \sup_{k\geq 2}\left\|D\hat{f}(x)^{-1}\cdot\frac{D^k\hat{f}(x)}{k!}\right\|^{\frac{1}{k-1}}\right),\\
\gamma_{\mu,n}(f,x) &= \max\left(1,\ \sup_{k\geq 2}\left\|\frac{1}{\Delta_\mu(f_n)}\cdot\frac{D^k f_n(x)}{k!}\right\|^{\frac{1}{k-1}}\right);
\end{aligned}$$

see Theorem 3.4 in Section 3.1 and Theorem 3.17 in Section 3.2.

- We propose a numerical criterion for isolating a cluster of μ zeros in the neighborhood of an approximately given simple multiple zero x from the other zeros of f. If

$$\|f(x)\| + \sum_{1\leq k\leq \mu-1}\|H_k\|\left(\frac{d}{4\gamma_\mu(g,x)^\mu}\right)^k < \frac{d^{\mu+1}}{2\left(4\gamma_\mu(g,x)^\mu\right)^\mu \|A^{-1}\|}$$

holds, where

$$\begin{aligned}
A^{-1} &= \begin{pmatrix} \frac{1}{\sqrt{2}}D\hat{f}(x)^{-1} & 0\\ 0 & \frac{\sqrt{2}}{\Delta_\mu(f_n)}\end{pmatrix}\in\mathbb{C}^{n\times n},\\
g(X) &= f(X) - f(x) - \sum_{1\leq k\leq\mu-1} H_k(X-x)^k,\\
H_1 &= \begin{pmatrix} \frac{\partial\hat{f}(x)}{\partial X_1} & 0\\ \frac{\partial f_n(x)}{\partial X_1} & \frac{\partial f_n(x)}{\partial\hat{X}}\end{pmatrix}\in\mathbb{C}^{n\times n},\\
H_k &= \left(\begin{pmatrix} 0 & 0\\ 0 & \Delta_k(f_n)\end{pmatrix}\ \mathbf{0}_{\underbrace{n\times\cdots\times n}_{k}\times(n-1)}\right)\in\mathbb{C}^{\overbrace{n\times\cdots\times n}^{k+1}},\ 2\leq k\leq\mu-1,
\end{aligned}$$

then f has μ zeros (counting multiplicities) inside the ball of radius $\frac{d}{4\gamma_\mu(g,x)^\mu}$ around x; see Theorem 3.8 in Section 3.1 and Theorem 3.20 in Section 3.2.

Related works. Our work is closely related to Dedieu and Shub's quantitative results [10] for simple double zeros, which generalized Smale's α-theory [3, 46–50] for simple zeros. Dedieu and Shub [10] explicitly gave an upper bound for separating simple double zeros of analytic functions, and a numeric criterion for separating a cluster of two zeros (counting multiplicities). Yakoubsohn [52] extended α-theory to clusters of zeros of univariate polynomials and provided an algorithm to compute them [53]. Giusti, Lecerf, Salvy, and Yakoubsohn [14] studied criteria on point estimates for locating clusters of zeros of analytic functions in the univariate case and provided bounds on the diameter of the cluster of μ zeros (counting multiplicities). They proposed an algorithm based on Schröder's iteration for approximating the cluster and a stopping criterion which guarantees that the algorithm quadratically converges to the cluster. In [15], they further generalized their results to locate and approximate clusters of zeros of analytic maps of embedding dimension one via the implicit function theorem and the symbolic deflation technique. We are inspired by their idea of reduction to one variable, but we perform a unitary transformation to obtain an equivalent polynomial system such that its Jacobian matrix at the simple multiple zero is of normal form. Then we efficiently compute a closed basis of the local dual space instead of computing the power series expansion of an implicitly known univariate analytic function, which might need a very high precision to guarantee the correctness of their algorithm, especially for some ill-posed systems.

There are other different numeric and symbolic approaches to compute multiple zeros of polynomial systems. In [40], Rall studied the convergence property of Newton's method for singular solutions, and many modifications of Newton's method to restore the quadratic convergence for singular solutions have been proposed in [7–9, 18, 40–42].

In [17], Griewank constructed a bordered system from the initial system f and the singular value decomposition of the Jacobian matrix $Df(x)$ to restore the quadratic convergence of Newton's method when $Df(x)$ has corank one. The method was extended by Shen and Ypma [44, 45] to the case where $Df(x)$ has arbitrary high rank deficiency.

In [37, 39, 54], Ojika et al. proposed a deflation method to construct a regular system to refine an approximate isolated singular solution to high accuracy. The deflation method has been further developed and generalized by Leykin, Verschelde and Zhao [28, 29] for singular solutions whose Jacobian matrix has arbitrary high rank deficiency and for overdetermined polynomial systems. Furthermore, they proved that the number of deflations needed to derive a regular solution of an augmented system is strictly less than the multiplicity. Dayton and Zeng [4, 5] proved that the depth of the local dual space is a tighter bound for the number of deflations. In [27], Lecerf gave a deflation algorithm which outputs a regular triangular system at the singular solution. In [34], Mantzaflaris and Mourrain proposed a one-step deflation method and verified a multiple root of a nearby system with a given multiplicity structure, which depends on the accuracy of the given approximate multiple root. Hauenstein, Mourrain, and Szanto [20, 21] proposed a novel deflation method which extends their early works [1, 34] to verify the existence of an isolated singular zero with a given multiplicity structure up to a given order. More recently, in [16], Giusti and Yakoubsohn proposed a new deflation sequence using the kerneling operator defined by the Schur complement of the Jacobian matrix and proved a new γ-theorem for analytic regular systems.

Since arbitrary perturbations of coefficients may transform an isolated singular solution into a cluster of simple roots, it is more difficult to verify that a polynomial system has a multiple root. However, one can always certify that a perturbed system has an isolated multiple zero or certify that the polynomial system has a cluster of μ zeros in a small ball centered at x.

In [24], based on the deflated square systems proposed by Yamamoto in [54], Kanzawa and Oishi presented a numerical method for proving the existence of "imperfect singular solutions" of nonlinear equations with guaranteed accuracy. Rump and Graillat [43] described a numeric algorithm for computing verified and narrow error bounds with the property that a perturbed system is certified to have a simple double zero within the computed error bounds. In [32], Li and Zhi generalized the algorithm in [43] to compute guaranteed error bounds such that a perturbed system is proved to have a breadth-one isolated singular solution within the computed error bounds. In [33], they further generalized their results to treat isolated singular zeros in arbitrary cases.

In [11, 25, 26], Kearfott et al. presented completely different methods based on verifying a nonzero topological degree to certify the existence of singular zeros of nonlinear systems.

Structure of the paper. In Section 2, we first recall some notation and show how to incrementally compute a closed basis of the local dual space of f at a simple multiple zero x of multiplicity μ. Then we demonstrate how to compute an equivalent polynomial system such that its Jacobian matrix at the simple multiple zero is of normal form. In Section 3, we begin by showing how to extend the main results in [10] to simple triple zeros. We present an upper bound for separating a simple triple zero x from the other zeros of f and an explicit criterion that guarantees the existence of a cluster of three zeros of f around an approximately given x. Then we generalize these results to simple multiple zeros with arbitrary high multiplicities. In Section 4, we demonstrate the performance of our algorithm on isolating simple multiple zeros and clusters of zeros for a list of benchmark examples. We also compare our algorithm with [10] and [15] with two detailed examples.

2. Definition of simple multiple zeros

2.1. **Local dual space.** Let $\mathbf{d}_x^{\boldsymbol{\alpha}} : \mathbb{C}[X] \to \mathbb{C}$ denote the differential functional defined by

$$\mathbf{d}_x^{\boldsymbol{\alpha}}(g) = \frac{1}{\alpha_1! \cdots \alpha_n!} \cdot \frac{\partial^{|\boldsymbol{\alpha}|} g}{\partial X_1^{\alpha_1} \cdots \partial X_n^{\alpha_n}}(x) \quad \forall g \in \mathbb{C}[X], \tag{2.1}$$

where $x \in \mathbb{C}^n$ and $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_n] \in \mathbb{N}^n$. Clearly,

$$\mathbf{d}_x^{\boldsymbol{\alpha}}\left((X - x)^{\boldsymbol{\beta}}\right) = \begin{cases} 1, & \text{if } \boldsymbol{\alpha} = \boldsymbol{\beta}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

Let I_f denote the ideal generated by $f = \{f_1, \ldots, f_n\}$, where $f_i \in \mathbb{C}[X_1, \ldots, X_n]$. The local dual space of I_f at an isolated zero x is a subspace of $\mathfrak{D}_x = \mathrm{span}_{\mathbb{C}}\{\mathbf{d}_x^{\alpha}\}$,

$$\mathcal{D}_{f,x} = \{\Lambda \in \mathfrak{D}_x \mid \Lambda(g) = 0 \ \forall g \in I_f\}. \tag{2.3}$$

When the zero x is clear from the context, we write $d_1^{\alpha_1} \cdots d_n^{\alpha_n}$ instead of $\mathbf{d}_x^{\boldsymbol{\alpha}}$, where

$$d_i^{\alpha_i} = \frac{1}{\alpha_i!} \cdot \frac{\partial^{\alpha_i}}{\partial X_i^{\alpha_i}}, \text{ for } i = 1, \ldots, n.$$

Let $\mathcal{D}_{f,x}^{(k)}$ denote the subspace of $\mathcal{D}_{f,x}$ of differential functionals of order bounded by k. We define

(1) breadth $\kappa = \dim\left(\mathcal{D}_{f,x}^{(1)}\right) - \dim\left(\mathcal{D}_{f,x}^{(0)}\right)$,

(2) depth $\rho = \min\left(\left\{k \mid \dim\left(\mathcal{D}_{f,x}^{(k+1)}\right) = \dim\left(\mathcal{D}_{f,x}^{(k)}\right)\right\}\right)$,

(3) multiplicity $\mu = \dim\left(\mathcal{D}_{f,x}^{(\rho)}\right)$.

If x is an isolated zero of f, then $1 \le \kappa \le n$ and $\rho < \mu < \infty$.

Let $\Phi_\sigma : \mathfrak{D}_x \to \mathfrak{D}_x$ denote the morphism defined by

$$\Phi_\sigma(d_1^{\alpha_1} \cdots d_n^{\alpha_n}) = \begin{cases} d_1^{\alpha_1} \cdots d_\sigma^{\alpha_\sigma - 1} \cdots d_n^{\alpha_n}, & \text{if } \alpha_\sigma > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then computing a closed basis of $\mathcal{D}_{f,x}^{(k)}$ is done essentially by matrix-kernel computations based on the stability property of $\mathcal{D}_{f,x}$ [5,35,36,51]:

$$\forall \Lambda \in \mathcal{D}_{f,x}^{(k)}\ \Phi_\sigma(\Lambda) \in \mathcal{D}_{f,x}^{(k-1)}, \quad \sigma = 1, \ldots, n. \tag{2.4}$$

In this paper, we deal with the simple multiple zeros satisfying $f(x) = 0$ and $\dim \ker Df(x) = 1$. They are also called breadth-one multiple zeros in [5] because $\dim(\mathcal{D}_{f,x}^{(k)}) - \dim(\mathcal{D}_{f,x}^{(k-1)}) = 1$, $k = 1, \ldots, \rho$, and $\rho = \mu - 1$. Therefore, for the breadth-one case,

$$\mathcal{D}_{f,x} = \operatorname{span}_{\mathbb{C}}\{\Lambda_0 = 1, \Lambda_1, \ldots, \Lambda_{\mu-1}\}, \tag{2.5}$$

where $\deg(\Lambda_k) = k$. Let $\Psi_\sigma : \mathfrak{D}_x \to \mathfrak{D}_x$ denote the morphism defined by

$$\Psi_\sigma(d_1^{\alpha_1} \cdots d_n^{\alpha_n}) = \begin{cases} d_\sigma^{\alpha_\sigma+1} \cdots d_n^{\alpha_n}, & \text{if } \alpha_1 = \cdots = \alpha_{\sigma-1} = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 2.1 ([30, Theorem 3.1])**.** *Suppose we are given a simple multiple zero x satisfying $f(x) = 0$, $\dim \ker Df(x) = 1$. Let $(a_{1,1}, \ldots, a_{1,n})^T \in \ker Df(x)$. Without loss of generality, we assume $a_{1,1} = 1$. Then*

$$\Lambda_1 = d_1 + a_{1,2}d_2 + \cdots + a_{1,n}d_n \in \mathcal{D}_{f,x}^{(1)},$$

and Λ_k can be incrementally constructed by $\Lambda_k = \Delta_k + a_{k,2}d_2 + \cdots + a_{k,n}d_n$, where

$$\Delta_k = \sum_{\sigma=1}^{n} \Psi_\sigma(a_{1,\sigma}\Lambda_{k-1} + \cdots + a_{k-1,\sigma}\Lambda_1), \ 2 \le k \le \mu. \tag{2.6}$$

The parameters $a_{k,2}, \ldots, a_{k,n}$ are determined by solving the linear system

$$\begin{pmatrix} d_2(f_1) & \cdots & d_n(f_1) \\ \vdots & \ddots & \vdots \\ d_2(f_n) & \cdots & d_n(f_n) \end{pmatrix} \begin{pmatrix} a_{k,2} \\ \vdots \\ a_{k,n} \end{pmatrix} = - \begin{pmatrix} \Delta_k(f_1) \\ \vdots \\ \Delta_k(f_n) \end{pmatrix}, \ 2 \le k < \mu. \tag{2.7}$$

When $\mu = 2$, suppose $\ker Df(x) = \operatorname{span}_{\mathbb{C}}\{v\}$ and $\|v\| = 1$. Then $\Lambda_1(f) = Df(x) \cdot v = v_1 d_1(f) + \cdots + v_n d_n(f)$ and

$$\begin{aligned} \Delta_2(f) &= \sum_{\sigma=1}^{n} \Psi_\sigma(v_\sigma \Lambda_1)(f) = \sum_{\sigma=1}^{n} \Psi_\sigma(v_\sigma(v_1 d_1 + \cdots + v_n d_n))(f) \\ &= \sum_{i>j} v_i v_j d_i d_j(f) + \Sigma v_i^2 d_i^2(f) = \frac{1}{2} D^2 f(x)(v,v). \end{aligned}$$

Therefore, the condition $\Delta_2(f) \notin \operatorname{im} Df(x)$ in Definition 1.2 is equivalent to the condition $D^2 f(x)(v,v) \notin \operatorname{im} Df(x)$ for simple double zeros; see formula (B) in [10].

In the following example, we show how to incrementally compute Λ_k and check the conditions satisfied by the simple multiple zero listed in Definition 1.2.

Example 2.2 ([39])**.** Given an isolated zero $x = (1, 2)$ of a polynomial system,

$$f = \begin{cases} X_1^2 + X_2 - 3, \\ X_1 + \frac{1}{8} X_2^2 - \frac{3}{2}. \end{cases} \tag{2.8}$$

- For $k = 1$, since $Df(x) = \left[\begin{smallmatrix} 2 & 1 \\ 1 & \frac{1}{2} \end{smallmatrix}\right]$, we have $\Lambda_1 = d_1 - 2d_2$.

- For $k = 2$, according to (2.6), we construct $\Delta_2 = d_1^2 - 2d_1d_2 + 4d_2^2$. Then $\Delta_2(f) = \begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} \in \operatorname{im} Df(x)$, and we solve (2.7) to obtain $\Lambda_2 = \Delta_2 - d_2$.
- For $k = 3$, according to (2.6), $\Delta_3 = d_1^3 - 2d_1^2d_2 + 4d_1d_2^2 - d_1d_2 - 8d_2^3 + 2d_2^2$, and $\Delta_3(f) = \begin{bmatrix} 0 \\ \frac{1}{4} \end{bmatrix} \notin \operatorname{im} Df(x)$. Therefore, x is a simple triple zero of f.

2.2. **Unitary transformations.** We show below that it is always possible to perform a unitary transformation to obtain an equivalent polynomial system such that its Jacobian matrix at the simple multiple zero is of normal form. We also show that the unitary transformation does not change the distance between two zeros.

Let x be a simple multiple zero of f. By Definition 1.2, we have $\dim\ker Df(x) = 1$. Suppose the Jacobian matrix $Df(x)$ is not of normal form. We compute two unit vectors $v \in \ker Df(x)$ and $u \in \ker Df(x)^T$ and then apply the Gram–Schmidt process to obtain the unitary vectors $v_1, \ldots, v_{n-1}$ and $u_1, \ldots, u_{n-1}$ such that

$$U^T \cdot Df(x) \cdot W = \begin{pmatrix} 0 & M \\ 0 & 0 \end{pmatrix}$$

where $W = (v, v_1, \ldots, v_{n-1})$, $U = (u_1, \ldots, u_{n-1}, u)$ are two unitary matrices and $M \in \mathbb{C}^{(n-1)\times(n-1)}$ is invertible. Let $g = U^T \cdot f(W \cdot X)$. Suppose x is a simple multiple zero of f of multiplicity μ. Then W^*x is a simple multiple zero of g of multiplicity μ and the Jacobian matrix $Dg(W^*x)$ is of normal form, since

$$Dg(W^*x) = U^T \cdot Df(x) \cdot W = \begin{pmatrix} 0 & M \\ 0 & 0 \end{pmatrix}.$$

Furthermore, suppose y is another zero of f. Then W^*y is another zero of g, and

$$\|W^*x - W^*y\| = \|W^*(x - y)\| = \|x - y\|. \tag{2.9}$$

Remark 2.3. It is clear that if x is given as an exact simple multiple zero of f, then as shown above, we can always construct an exact unitary transformation

$$W^*W = WW^* = I_{n\times n} \tag{2.10}$$

such that W^*x is a simple multiple zero of $g = U^T \cdot f(W \cdot X)$ of multiplicity μ and $Dg(W^*x)$ is of normal form. According to (2.9), an exact unitary transformation does not change the distance between two zeros. Therefore, in the following sections, if x is exactly given, then we always assume that $Df(x)$ is of normal form.

Example 2.2 (continued). We compute two unit vectors $v = u = \begin{bmatrix} \frac{\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} \end{bmatrix}$ such that $Df(x)v = Df(x)^T u = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Then use the Gram–Schmidt process to get $v_1 = u_1 = \begin{bmatrix} \frac{2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \end{bmatrix}$ such that

$$\begin{bmatrix} \frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \end{bmatrix} \cdot Df(x) \cdot \begin{bmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{bmatrix} = \begin{bmatrix} 0 & \frac{5}{2} \\ 0 & 0 \end{bmatrix},$$

which is of normal form. By performing a unitary transformation introduced above, we transform $x = (1, 2)$ to $z = (-\frac{3\sqrt{5}}{5}, \frac{4\sqrt{5}}{5})$ and the polynomial system f to

$$g = \begin{cases} \frac{\sqrt{5}}{10}X_1^2 + \frac{3\sqrt{5}}{10}X_1X_2 + \frac{13\sqrt{5}}{40}X_2^2 - \frac{3}{5}X_1 + \frac{4}{5}X_2 - \frac{3\sqrt{5}}{2}, \\ \frac{\sqrt{5}}{5}X_1X_2 + \frac{3\sqrt{5}}{20}X_2^2 - \frac{4}{5}X_1 - \frac{3}{5}X_2. \end{cases} \tag{2.11}$$

Another zero $y = (-3, -6)$ of f is transformed to $w = (\frac{9\sqrt{5}}{5}, -\frac{12\sqrt{5}}{5})$. The distance between x and y is equal to the distance between z and w, i.e., $\|y - x\| = \|z - w\| = 4\sqrt{5}$. It is clear that $Dg(z) = \begin{bmatrix} 0 & \frac{5}{2} \\ 0 & 0 \end{bmatrix}$. Moreover, according to (2.6) and (2.7), we have

$$\Delta_2(g_2) = d_1^2(g_2) = 0, \Delta_3(g_2) = d_1^3(g_2) - \frac{\sqrt{5}}{25}d_1d_2(g_2) = -\frac{1}{25}.$$

Therefore, z is a simple triple zero of g and $Dg(z)$ is of normal form.

Remark 2.4. If x is only given with limited accuracy, we can compute two approximate null vectors v and u of $\ker Df(x)$ and $\ker Df(x)^T$, and then use the Gram–Schmidt process to get two unitary matrices $W = (v, v_1, \ldots, v_{n-1})$, $U = (u_1, \ldots, u_{n-1}, u)$. At this moment, $g = U^T \cdot f(W \cdot X)$ has a cluster of zeros near W^*x and $Dg(W^*x)$ is of normal form approximately, i.e., the entries of its first column and its last row are all approximately zero. According to (2.9), the unitary transformation does not change the distance between two zeros. Therefore, if we can certify that $B(W^*x, r)$ contains μ zeros of g, then $B(x, r)$ is certified to contain μ zeros of f.

Remark 2.5. It is also possible to perform the unitary transformation via the singular value decomposition. Let $Df(x) = U \cdot \left(\begin{smallmatrix} \Sigma_{n-1} & 0 \\ 0 & 0 \end{smallmatrix}\right) \cdot V^*$, where $U = (u_1, \ldots, u_n)$, $V = (v_1, \ldots, v_n)$ are unitary matrices, V^* is the Hermitian transpose of V, and Σ_{n-1} is an invertible diagonal matrix. Let $g = U^* \cdot f(W \cdot X)$, where $W = (v_n, v_1, \ldots, v_{n-1})$ is also a unitary matrix. Suppose x is a simple multiple zero of f of multiplicity μ. Then W^*x is a simple multiple zero of g of multiplicity μ and $Dg(W^*x)$ is of normal form, since

$$\begin{aligned} Dg(W^*x) &= U^* \cdot Df(x) \cdot W = U^* \cdot U \cdot \Sigma \cdot V^* \cdot W \\ &= \begin{pmatrix} \Sigma_{n-1} & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & I_{n-1} \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \Sigma_{n-1} \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

It should be noted that if we perform the numerical singular value decomposition of $Df(x)$, then the matrix W^* only satisfies condition (2.10) approximately, i.e., $\|W^*W - I_{n\times n}\| \le \epsilon$ and $\|WW^* - I_{n\times n}\| \le \epsilon$ for a given tolerance ϵ. The change in the Euclidian distance $\|W^*x - W^*y\|$ and $\|x - y\|$ is negligible; see the results obtained for Example 2.2 in Section 3.1.

3. Local separation bound and cluster location

We start by showing how to extend the main results in [10] to simple triple zeros. Then we generalize these quantitative results to simple multiple zeros with arbitrary high multiplicities.

3.1. **Simple triple zeros.** Let x be a simple triple zero of f and suppose $Df(x)$ is of normal form, i.e., $\frac{\partial f_i(x)}{\partial X_1} = 0$, $\frac{\partial f_n(x)}{\partial X_i} = 0$ for $1 \le i \le n$, and

$$\Delta_2(f_n) = 0,\ \Delta_3(f_n) \neq 0. \tag{3.1}$$

Recall that $D\hat{f}(x)$ is the Jacobian matrix of $\hat{f} = \{f_1, \dots, f_{n-1}\}$ with respect to $\hat{X} = \{X_2, \dots, X_n\}$, which is invertible since $Df(x)$ is of normal form. Let $\Lambda_0 = 1$ and $\Lambda_1 = d_1$. According to (2.6), we have $\Delta_2 = d_1^2$,

$$\Lambda_2 = d_1^2 + a_{2,2}d_2 + \cdots + a_{2,n}d_n,$$

where $a_{2,2}, \dots, a_{2,n}$ satisfy

$$\begin{pmatrix} a_{2,2} \\ \vdots \\ a_{2,n} \end{pmatrix} = -D\hat{f}(x)^{-1} \begin{pmatrix} \Delta_2(f_1) \\ \vdots \\ \Delta_2(f_{n-1}) \end{pmatrix} = -D\hat{f}(x)^{-1} \begin{pmatrix} d_1^2(f_1) \\ \vdots \\ d_1^2(f_{n-1}) \end{pmatrix}.$$

Moreover, since $a_{1,1} = 1$, $a_{2,1} = 0$, we have

$$\begin{aligned} \Delta_3 &= \sum_{\sigma=1}^{n} \Psi_\sigma(a_{1,\sigma}\Lambda_2 + a_{2,\sigma}\Lambda_1) = \Psi_1(\Lambda_2) + \sum_{\sigma=1}^{n} \Psi_\sigma(a_{2,\sigma}d_1) \\ &= d_1^3 + a_{2,2}d_1d_2 + \cdots + a_{2,n}d_1d_n \\ &= d_1^3 + (d_1d_2, \dots, d_1d_n) \cdot \left(-D\hat{f}(x)^{-1}\right) \cdot \begin{pmatrix} d_1^2(f_1) \\ \vdots \\ d_1^2(f_{n-1}) \end{pmatrix}. \end{aligned}$$

Since $d_i^{\alpha_i} = \frac{1}{\alpha_i!} \cdot \frac{\partial^{\alpha_i}}{\partial X_i^{\alpha_i}}$, $i = 1, \dots, n$, the condition (3.1) can be written explicitly as

$$\begin{aligned} \Delta_2(f_n) &= \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial X_1^2} = 0, \\ \Delta_3(f_n) &= \frac{1}{6}\frac{\partial^3 f_n(x)}{\partial X_1^3} - \frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}} \cdot D\hat{f}(x)^{-1} \frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial X_1^2} \neq 0. \end{aligned}$$

For two nonzero vectors $a,\ b \in \mathbb{C}^n$, we denote

$$d_P(a, b) = \arccos \frac{|\langle a, b\rangle|}{\|a\| \cdot \|b\|} \tag{3.2}$$

as the angle between them. Let y be another point in $\mathbb{C}^n$ with $y \neq x$ and define

$$y - x = (\zeta, \eta_2, \dots, \eta_n)^T,\ \eta = (\eta_2, \dots, \eta_n). \tag{3.3}$$

Let $\varphi = d_P(v, y - x)$, $v = (1, 0, \dots, 0)^T \in \ker Df(x)$; then we derive

$$|\zeta| = \|y - x\| \cos\varphi,\ \ \|\eta\| = \|y - x\| \sin\varphi. \tag{3.4}$$

For $k \geq 2$, we denote $D^k \hat{f}(x)$ as the partial derivatives of $\hat{f}$ of order k. We generalize the main results in [10] to simple triple zeros.

The following lemma has been given in [10, Lemma 1] that is devoted to bound the value of $\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\|$ from below when $0 < \varphi \leq \frac{\pi}{2}$. We present its short proof for completeness.

Lemma 3.1. *If $\hat{\gamma}_3(f, x)\|y - x\| \leq \frac{1}{2}$, then*

$$\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| \geq \|y - x\| \sin\varphi - 2\hat{\gamma}_3(f, x)\|y - x\|^2. \tag{3.5}$$

Proof. By Taylor's expansion of $\hat{f}(y)$ at x, and $\frac{\partial \hat{f}(x)}{\partial X_1} = 0$, we derive

$$\hat{f}(y) = \hat{f}(x) + D\hat{f}(x)\eta + \sum_{k\geq 2} \frac{D^k \hat{f}(x)(y-x)^k}{k!}.$$

Since $\hat{f}(x) = 0$ and $D\hat{f}(x)$ is invertible, it implies that

$$\eta = D\hat{f}(x)^{-1}\hat{f}(y) - \sum_{k\geq 2} D\hat{f}(x)^{-1}\frac{D^k \hat{f}(x)(y-x)^k}{k!}. \tag{3.6}$$

By the triangle inequality, we conclude that

$$\begin{aligned}
\|y-x\|\sin\varphi = \|\eta\| &\leq \left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| + \sum_{k\geq 2}\left\|D\hat{f}(x)^{-1}\frac{D^k\hat{f}(x)}{k!}\right\| \|y-x\|^k \\
&\leq \left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| + \sum_{k\geq 2}\hat{\gamma}_3(f,x)^{k-1}\|y-x\|^k \\
&\leq \left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| + 2\hat{\gamma}_3(f,x)\|y-x\|^2,
\end{aligned}$$

where the last inequality comes from the assumption that $\hat{\gamma}_3(f,x)\|y-x\| \leq \frac{1}{2}$. $\square$

For simplicity of symbols, we denote $\gamma_3(f,x) = \max(\hat{\gamma}_3(f,x), \gamma_{3,n}(f,x))$ by γ_3 in the subsection. Let

$$A = \begin{pmatrix} \sqrt{2}D\hat{f}(x) & 0 \\ 0 & \frac{1}{\sqrt{2}}\Delta_3(f_n) \end{pmatrix} \in \mathbb{C}^{n\times n}. \tag{3.7}$$

Since $D\hat{f}(x)$ is invertible and $\Delta_3(f_n) \neq 0$, we have

$$A^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}}D\hat{f}(x)^{-1} & 0 \\ 0 & \frac{\sqrt{2}}{\Delta_3(f_n)} \end{pmatrix}. \tag{3.8}$$

The following lemma extends the result [10, Lemma 3] for simple double zeros to simple triple zeros. It gives the lower bound of $\|A^{-1}f(y)\|$ when $0 \leq \varphi \leq \arctan\frac{1}{\sqrt{2}}$. The proof is based on Taylor's expansion of f_n at x and the conditions $\frac{\partial f_n(x)}{\partial X_1} = \cdots = \frac{\partial f_n(x)}{\partial X_n} = \frac{\partial^2 f_n(x)}{\partial X_1^2} = 0$, $\Delta_3(f_n) \neq 0$. The proof of Lemma 3.2 is quite technical, so we move it to the Appendix.

Lemma 3.2. *If $\gamma_3\|y-x\| \leq \frac{1}{2}$ and $0 \leq \varphi \leq \arctan\frac{1}{\sqrt{2}}$, we have*

$$\|A^{-1}f(y)\| \geq 2\gamma_3^3\|y-x\|^3(h(\varphi) - \|y-x\|) \tag{3.9}$$

where

$$h(\varphi) = \frac{\cos^3\varphi - 8\gamma_3^2\cos^2\varphi\sin\varphi - 7\gamma_3^2\cos\varphi\sin^2\varphi - 2\gamma_3^2\sin^3\varphi}{2\gamma_3^3(1+2\cos\varphi+\sin\varphi)}. \tag{3.10}$$

The following lemma extends the result of [10, Lemma 4] for simple double zeros to simple triple zeros by finding a suitable $\theta \in (0, \arctan\frac{1}{\sqrt{2}}]$ such that Lemmas 3.1 and 3.2 can be used to bound $\|y-x\|$ from below by a universal formula.

Lemma 3.3. *Let $d \approx 0.08507$ be the positive root of the equation*

$$(1 - 2d - 8d^2)\sqrt{1 - d^2} - 9d - d^2 + 6d^3 = 0, \tag{3.11}$$

and let θ be defined by $\sin\theta = \frac{d}{\gamma_3^2}$. If $\gamma_3\|y - x\| \le \frac{1}{2}$ and $y \in \mathbb{C}^n$, then either

$$\theta \le \varphi \le \frac{\pi}{2} \text{ and } \|A^{-1}f(y)\| \ge \sqrt{2}\gamma_3\|y - x\|\left(\frac{\sin\theta}{2\gamma_3} - \|y - x\|\right)$$

or

$$0 \le \varphi \le \theta \text{ and } \|A^{-1}f(y)\| \ge 2\gamma_3^3\|y - x\|^3\left(\frac{\sin\theta}{2\gamma_3} - \|y - x\|\right).$$

Proof. If $\theta \le \varphi \le \frac{\pi}{2}$, by Lemma 3.1, we derive

$$\begin{aligned}\sqrt{2}\|A^{-1}f(y)\| &= \left\|\begin{pmatrix} D\hat{f}(x)^{-1}\hat{f}(y) \\ \frac{2}{\Delta_3(f_n)}f_n(y)\end{pmatrix}\right\| \ge \left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| \\ &\ge 2\gamma_3\|y - x\|\left(\frac{\sin\theta}{2\gamma_3} - \|y - x\|\right).\end{aligned}$$

If $0 \le \varphi \le \arctan\frac{1}{\sqrt{2}}$, by Lemma 3.2, we derive

$$\|A^{-1}f(y)\| \ge 2\gamma_3^3\|y - x\|^3\left(h(\varphi) - \|y - x\|\right),$$

where $h(\varphi)$ is defined in (3.10). We claim that

$$h(\theta) \ge \frac{\sin\theta}{2\gamma_3}. \tag{3.12}$$

Because $\sin\theta = \frac{d}{\gamma_3^2}$, it is sufficient to show that

$$\left(1 - \frac{d^2}{\gamma_3^4}\right)^{\frac{3}{2}} - 8d\left(1 - \frac{d^2}{\gamma_3^4}\right) - \frac{7d^2}{\gamma_3^2}\sqrt{1 - \frac{d^2}{\gamma_3^4}} - \frac{2d^3}{\gamma_3^4} - d - 2d\sqrt{1 - \frac{d^2}{\gamma_3^4}} - \frac{d^2}{\gamma_3^2} \ge 0.$$

Since the left function for $\gamma_3 \ge 1$ is increasing for $\forall d \in \left[0,\ \frac{1}{6}\right]$, similar to the proof of [10, Lemma 4], it is sufficient to check this inequality for $\gamma_3 = 1$, i.e.,

$$(1 - 2d - 8d^2)\sqrt{1 - d^2} - 9d - d^2 + 6d^3 \ge 0.$$

The smallest positive root of equation (3.11), $d \approx 0.08507 \in \left[0,\ \frac{1}{6}\right]$, is actually a valid value. Therefore, the claim $h(\theta) \ge \frac{\sin\theta}{2\gamma_3}$ follows.

Furthermore, for $0 \le \varphi \le \theta \le \arctan\frac{1}{\sqrt{2}}$, the function $h(\varphi)$ is nonnegative and decreasing since its numerator is decreasing, its denominator is increasing, and both of them are nonnegative. It implies that

$$h(\varphi) \ge h(\theta) \ge \frac{\sin\theta}{2\gamma_3}.$$

Since $\theta \le \arctan\frac{1}{\sqrt{2}}$ for $d \approx 0.08507$ and $\gamma_3 \ge 1$, by Lemma 3.2, we conclude that

$$\|A^{-1}f(y)\| \ge 2\gamma_3^3\|y - x\|^3\left(h(\varphi) - \|y - x\|\right) \ge 2\gamma_3^3\|y - x\|^3\left(\frac{\sin\theta}{2\gamma_3} - \|y - x\|\right),$$

for $0 \le \varphi \le \theta$. $\square$

Let $d \approx 0.08507$ be the smallest positive root of the equation (3.11). The following four theorems generalize the main results in [10] to simple triple zeros.

Theorem 3.4. *Let x be a simple triple zero of f. If y is another zero of f, then*

$$\|y-x\| \geq \frac{d}{2\gamma_3^3}. \tag{3.13}$$

Proof. When $\gamma_3\|y-x\| \leq \frac{1}{2}$, by Lemma 3.3 and $\sin\theta = \frac{d}{\gamma_3^2}$, we conclude that

$$\|y-x\| \geq \frac{\sin\theta}{2\gamma_3} = \frac{d}{2\gamma_3^3},$$

since $f(y)=0$. When $\gamma_3\|y-x\| \geq \frac{1}{2}$, the conclusion holds as $\gamma_3 \geq 1$ and $d<1$. $\square$

According to Theorem 3.4, for $r < \frac{d}{2\gamma_3^3}$, x is the only zero in the ball $B(x,r)$. The local separation bound $\frac{d}{2\gamma_3^3}$ can be explicitly computed as shown by the following example.

Example 2.2 (continued). We have shown in Section 2.2, after performing the unitary transformation, that we have a polynomial system g (2.11), which has a simple triple zero $z = (-\frac{3\sqrt{5}}{5}, \frac{4\sqrt{5}}{5})$. Because $Dg(z)$ is of normal form and $\mu=3$, we calculate

$$\hat{\gamma}_3 = \hat{\gamma}_3(g,z) = \max\left(1, \left\|\frac{2}{5}\cdot\frac{D^2 g_1(z)}{2}\right\|\right) = 1,$$
$$\gamma_{3,2} = \gamma_{3,2}(g,z) = \max\left(1, \left\|25\cdot\frac{D^2 g_2(z)}{2}\right\|\right) = 5\sqrt{5},$$

and thus $\gamma_3 = \max(\hat{\gamma}_3, \gamma_{3,2}) = 5\sqrt{5}$. According to Theorem 3.4, $\frac{d}{2\gamma_3^3} \approx 0.00003044$ is an upper bound of r such that $B(z,r)$ isolates z from the other zeros of g. Since the unitary transformation does not change the distance between two zeros, we guarantee that x is the only zero of f in the ball $B(x,r)$ for $r<0.00003044$.

Remark 3.5. The separation bound plays an important role in the subdivision-based algorithms (like [12]) for isolating all zeros of a polynomial system. For Example 2.2, although our local separation bound 0.00003044 is still smaller than the actual distance $\|x-y\| = 4\sqrt{5}$, it is much better than the global separation bound $\ll 10^{-10}$ computed by the method in [12].

The separation bound can be used to obtain a numerical criterion for isolating a cluster of three zeros in the neighborhood of an approximately given simple triple zero x from the other zeros of f. The following theorem provides a lower bound of the value $\|f(y)\|$ for any y in the ball $B(x, \frac{d}{4\gamma_3^3})$.

Theorem 3.6. *Let x be a simple triple zero of f. If $\|y-x\| \leq \frac{d}{4\gamma_3^3}$, then*

$$\|f(y)\| \geq \frac{d\|y-x\|^3}{2\|A^{-1}\|}.$$

Proof. When $\|y-x\| = \|y-x\| \leq \frac{d}{4\gamma_3^3} = \frac{\sin\theta}{4\gamma_3}$, by Lemma 3.3, we have

$$\|A^{-1}f(y)\| \geq 2\gamma_3^3\|y-x\|^3\left(\frac{\sin\theta}{2\gamma_3} - \|y-x\|\right) \geq 2\gamma_3^3\|y-x\|^3\frac{\sin\theta}{4\gamma_3} = \frac{d}{2}\|y-x\|^3. \ \square$$

For $R>0$, we define $d_R(f,g) = \max_{\|y-x\|\leq R}\|f(y)-g(y)\|$. Theorem 3.7 follows straight from Rouché's Theorem.

Theorem 3.7. *Let x be a simple triple zero of f, and let $0 < R \leq \frac{d}{4\gamma_3^3}$. If*

$$d_R(f,g) < \frac{dR^3}{2\|A^{-1}\|},$$

then the sum of multiplicities of zeros of g in $B(x,R)$ is three.

Proof. By Theorem 3.6, for any y such that $\|y-x\| = R$, we derive

$$\|f(y) - g(y)\| \leq d_R(f,g) < \frac{dR^3}{2\|A^{-1}\|} = \frac{d\|y-x\|^3}{2\|A^{-1}\|} \leq \|f(y)\|.$$

According to Rouché's Theorem: if f possesses μ zeros (counting multiplicities) in $B(x,R)$, then any analytic g satisfying $\|f(y) - g(y)\| < \|f(y)\|\ \forall\, y \in \partial B(x,R)$ has μ zeros (counting multiplicities) in $B(x,R)$ (see [2, Theorem 2.12]). According to Theorem 3.4, if $R \leq \frac{d}{4\gamma_3^3}$, x is the only triple zero of f in $B(x,R)$. Therefore, g has three zeros (counting multiplicities) in $B(x,R)$. □

Next we consider a more difficult but useful case. Suppose x is an approximately given simple triple zero of f and $Df(x)$ is of approximately normal form, i.e., $f(x)$, the entries of $Df(x)$'s first column and last row, $|\Delta_2(f_n)|$, are small with respect to a given tolerance but $|\Delta_3(f_n)|$ is not, then we propose a numerical criterion for isolating a cluster of three zeros in the neighborhood of x from the other zeros of f.

Let us explicitly write the formulas of H_1, H_2 and $g(X)$:

$$\begin{aligned}
H_1 &= \begin{pmatrix} \frac{\partial \hat{f}(x)}{\partial X_1} & 0 \\ \frac{\partial f_n(x)}{\partial X_1} & \frac{\partial f_n(x)}{\partial \hat{X}} \end{pmatrix}, \\
H_2 &= \left(\begin{pmatrix} 0 & 0 \\ \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial X_1^2} & 0 \end{pmatrix} \mathbf{0}_{n\times n\times(n-1)} \right), \\
g(X) &= f(X) - f(x) - H_1(X-x) - H_2(X-x)^2.
\end{aligned}$$

Theorem 3.8. *Let $\gamma_3 = \gamma_3(g,x)$. If*

$$\|f(x)\| + \|H_1\|\frac{d}{4\gamma_3^3} + \|H_2\|\frac{d^2}{16\gamma_3^6} < \frac{d^4}{128\gamma_3^9\|A^{-1}\|}, \tag{3.14}$$

then f has three zeros (counting multiplicities) in the ball of radius $\frac{d}{4{\gamma_3}^3}$ around x.

Proof. Clearly, x is a zero of g and $Dg(x) = Df(x) - H_1 = \begin{pmatrix} 0 & D\hat{f}(x) \\ 0 & 0 \end{pmatrix}$,

$$\begin{aligned}
\Delta_2(g_n) &= \frac{1}{2}\frac{\partial^2 g_n(x)}{\partial X_1^2} = \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial X_1^2} - \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial X_1^2} = 0, \\
\Delta_3(g_n) &= \frac{1}{6}\frac{\partial^3 g_n(x)}{\partial X_1^3} - \frac{\partial^2 g_n(x)}{\partial X_1 \partial \hat{X}} \cdot D\hat{g}(x)^{-1} \cdot \frac{1}{2}\frac{\partial^2 \hat{g}(x)}{\partial X_1^2} \\
&= \frac{1}{6}\frac{\partial^3 f_n(x)}{\partial X_1^3} - \frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}} \cdot D\hat{f}(x)^{-1} \cdot \frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial X_1^2} \neq 0.
\end{aligned}$$

Hence x is a simple triple zero of g and $Dg(x)$ is of normal form. Let $R = \frac{d}{4{\gamma_3}^3}$; we derive

$$d_R(g,f) = \max_{\|y-x\|\le R} \|g(y) - f(y)\| = \max_{\|y-x\|\le R} \|f(x) + H_1(y-x) + H_2(y-x)^2\|$$
$$\le \|f(x)\| + \|H_1\|R + \|H_2\|R^2 = \|f(x)\| + \|H_1\|\frac{d}{4\gamma_3^3} + \|H_2\|\frac{d^2}{16\gamma_3^6}.$$

If $\|f(x)\| + \|H_1\|\frac{d}{4\gamma_3^3} + \|H_2\|\frac{d^2}{16\gamma_3^6} < \frac{d^4}{128\gamma_3^9\|A^{-1}\|}$, then

$$d_R(g,f) < \frac{d^4}{128\gamma_3^9\|A^{-1}\|} = \frac{dR^3}{2\|A^{-1}\|}.$$

By Theorem 3.7, the sum of multiplicities of zeros of f in $B(x,R)$ is three. $\square$

Remark 3.9. When x is only given with limited accuracy, the radius $r = \frac{d}{4\gamma_3^3}$ of the ball $B(x,r)$ which isolates a cluster of three zeros from the other zeros of f is only half of the local separation bound $r = \frac{d}{2\gamma_3^3}$ obtained in Theorem 3.4 for isolating an exactly given simple triple zero x of f.

Given an approximate simple triple zero of a polynomial system, the validation of the criterion (3.14) can be examined via numerical computations as shown by the following example.

Example 2.2 (continued). Suppose we are given $x = (1 + 10^{-18}, 2 - 10^{-18})$; we set digits= 20 for computing $Df(x) = U \cdot \begin{pmatrix} \Sigma_{n-1} & 0 \\ 0 & 0 \end{pmatrix} \cdot V^*$ via the singular value decomposition. For simplicity, below we only show the results to four decimal places. We have

$$U \approx \begin{bmatrix} -0.8944 & -0.4472 \\ -0.4472 & 0.8944 \end{bmatrix}, V^* \approx \begin{bmatrix} -0.8944 & -0.4472 \\ -0.4472 & 0.8944 \end{bmatrix}.$$

After performing the unitary transformation defined in Remark 2.5, we obtain an approximate zero $z \approx (-1.3416, 1.7889)$ of

$$g \approx \begin{cases} 0.2236X_1^2 + 0.6708X_1X_2 + 0.7267X_2^2 - 0.6X_1 + 0.8X_2 - 3.3541, \\ -2.5\times 10^{-19}X_1^2 + 0.4472X_1X_2 + 0.3354X_2^2 - 0.8X_1 - 0.6X_2 + 1.2\times 10^{-18}. \end{cases}$$

The unitary matrix $W \approx \left[\begin{smallmatrix} 0.4472 & 0.8944 \\ -0.8944 & 0.4472 \end{smallmatrix}\right]$ satisfies

$$W\cdot W^* - I_3 = W^*\cdot W - I_3 = \begin{bmatrix} -1\times 10^{-20} & 0 \\ 0 & -1\times 10^{-20} \end{bmatrix}. \tag{3.15}$$

It is easy to check that $\|g(z)\| \approx 1.1045\times 10^{-18}$, and $Dg(z)$ satisfies the condition of normal form (1.2) approximately if we set the tolerance to be 10^{-18}:

$$Dg(z) = \begin{bmatrix} -1.3\times 10^{-19} & 2.5 + 1.2\times 10^{-18} \\ 1.6\times 10^{-19} & -1.0\times 10^{-20} \end{bmatrix}.$$

Moreover, we can also check that $|\Delta_2(g_2)| = |d_1^2(g_2)| \approx 2.53\times 10^{-19}$, but $|\Delta_3(g_2)| \approx |d_1^3(g_2) - 0.08944 d_1 d_2(g_2)| \approx 0.04$. Therefore, we construct a new system

$$\tilde{g} = g - g(z) - \begin{pmatrix} \frac{\partial g_1(z)}{\partial X_1}(X_1 - z_1) \\ \frac{\partial g_2(z)}{\partial X_1}(X_1 - z_1) \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{\partial g_2(z)}{\partial X_2}(X_2 - z_2) \end{pmatrix} - \begin{pmatrix} 0 \\ \Delta_2(g_2)(X_1 - z_1)^2 \end{pmatrix}.$$

It is straightforward to check that z is a simple triple zero of $\tilde{g}$ and $D\tilde{g}(z)$ is of normal form. We calculate

$$\hat{\gamma}_3 = \hat{\gamma}_3(\tilde{g}, z) = \max\left(1, \left\|0.4 \cdot \frac{D^2\tilde{g}_1(z)}{2}\right\|\right) = 1,$$

$$\gamma_{3,2} = \gamma_{3,2}(\tilde{g}, z) = \max\left(1, \left\|25 \cdot \frac{D^2\tilde{g}_2(z)}{2}\right\|\right) \approx 11.1803,$$

and thus $\gamma_3 = \max(\hat{\gamma}_3, \gamma_{3,2}) \approx 11.1803$. Then we calculate the invertible matrix $A \approx \left(\begin{smallmatrix} 3.5355 & 0 \\ 0 & 0.02828 \end{smallmatrix}\right)$. Finally, the criterion (3.14)

$$1.1045 \times 10^{-18} \approx \|g(x)\| + \|H_1\|\frac{d}{4\gamma_3^3} + \|H_2\|\frac{d^2}{16\gamma_3^6} < \frac{d^4}{128\gamma_3^9\|A^{-1}\|} \approx 4.2397 \times 10^{-18}$$

is satisfied. According to Theorem 3.8, it implies that g has three zeros in the ball of radius $r = \frac{d}{4\gamma_3^3} \approx 0.00001522$. According to (3.15), the unitary matrix W satisfies condition (2.10) approximately, hence we conclude that f has a cluster of three zeros in the ball $B(x, r)$.

Remark 3.10. For Example 2.2, the value of $r = \frac{d}{4\gamma_3^3} \approx 0.00001522$ is half of the local separation bound $\frac{d}{2\gamma_3^3} \approx 0.00003044$, which is obtained from an exactly given simple triple zero and an exact unitary transformation. It shows that the computation of the value of r is numerically stable even if an approximate unitary transformation is performed.

Remark 3.11. The equality of $\gamma_\mu(g, x) = \gamma_\mu(f, x)$ is true for $\mu = 2$ [10, Theorem 4]. For [19, Example 2], we show that $\left\|\frac{1}{\Delta_3(f_2)} \cdot \frac{D^2 f_2(x)}{2}\right\| \neq \left\|\frac{1}{\Delta_3(g_2)} \cdot \frac{D^2 g_2(x)}{2}\right\|$. Hence, $\gamma_{3,n}(g, x)$ may not be equal to $\gamma_{3,n}(f, x)$ if they are not equal to 1.

3.2. **Simple multiple zeros.** We generalize the results in Section 3.1 to simple multiple zeros of arbitrary high multiplicities.

Let $f : \mathbb{C}^n \to \mathbb{C}^n$, let x be a simple multiple zero of f of multiplicity μ; and suppose $Df(x)$ is of normal form, i.e., $\frac{\partial f_i(x)}{\partial X_1} = 0$ and $\frac{\partial f_n(x)}{\partial X_i} = 0$ for $1 \leq i \leq n$, $\Delta_k(f_n) = 0$ for $k = 2, \ldots, \mu - 1$, $\Delta_\mu(f_n) \neq 0$.

For simplicity of symbols, we denote $\gamma_\mu(f, x) = \max(\hat{\gamma}_\mu(f, x), \gamma_{\mu,n}(f, x))$ by γ_μ in the subsection. Let

$$A = \begin{pmatrix} \sqrt{2}D\hat{f}(x) & 0 \\ 0 & \frac{1}{\sqrt{2}}\Delta_\mu(f_n) \end{pmatrix}. \tag{3.16}$$

Since $D\hat{f}(x)$ is invertible and $\Delta_\mu(f_n) \neq 0$, we have

$$A^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}}D\hat{f}(x)^{-1} & 0 \\ 0 & \frac{\sqrt{2}}{\Delta_\mu(f_n)} \end{pmatrix}. \tag{3.17}$$

It is clear that the proof of Lemma 3.1 does not depend on the multiplicity μ, therefore it is straightforward to be generalized to the following lemma.

Lemma 3.12. *If $\gamma_\mu\|y - x\| \leq \frac{1}{2}$, then*

$$\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| \geq 2\gamma_\mu\|y - x\|\left(\frac{\sin\varphi}{2\gamma_\mu} - \|y - x\|\right).$$

Lemma 3.14 extends Lemma 3.2 to simple multiple zeros of arbitrary multiplicity μ. The proof is based on a reformulation of Taylor's expansion of f_n at x in Lemma 3.13 and the conditions $\frac{\partial f_n(x)}{\partial X_1} = \cdots = \frac{\partial f_n(x)}{\partial X_n} = 0$, $\Delta_k(f_n) = 0$ for $k = 2, \ldots, \mu - 1$, $\Delta_\mu(f_n) \neq 0$. The proofs of Lemmas 3.13 and 3.14 are quite technical, so we move them to the Appendix.

Lemma 3.13. *We reformulate Taylor's expansion of $f_n(y)$ at x as*

(3.18)

$$\begin{aligned} f_n(y) &= C_2\zeta^2 + \cdots + C_\mu\zeta^\mu + \sum_{i+j=\mu, j>0} C_{i,j}\zeta^i\eta^j + \sum_{k\geq\mu+1} \frac{D^k f_n(x)(y-x)^k}{k!} \\ &\quad + \sum_{1\leq i+j-1\leq\mu-2} T_{i,j-1}\cdot\left(\sum_{k+i+j-1\geq\mu+1} D\hat{f}(x)^{-1}\frac{D^k\hat{f}(x)(y-x)^k}{k!}\zeta^i\eta^{j-1}\right) \\ &\quad - \sum_{1\leq i+j-1\leq\mu-2} T_{i,j-1}D\hat{f}(x)^{-1}\hat{f}(y)\zeta^i\eta^{j-1}. \end{aligned}$$

The coefficients of ζ^t in (3.18) satisfy

$$C_t = \Delta_t(f_n), \text{ for } 2 \leq t \leq \mu.$$

$C_{i,j}$ and $T_{i,j-1}$ in (3.18) satisfy the inequalities

$$\left\|\frac{1}{\Delta_\mu(f_n)}C_{i,j}\right\| \leq c_{i,j}\gamma_\mu^{i+j-1}, \quad \left\|\frac{1}{\Delta_\mu(f_n)}T_{i,j-1}\right\| \leq t_{i,j-1}\gamma_\mu^{i+j-1}, \tag{3.19}$$

where $c_{i,j}$ $(i+j=\mu, j>0)$ and $t_{i,j-1}$ $(2 \leq i+j \leq \mu-1)$ are constants, which can be computed inductively by

$$\begin{aligned} c_{i,j} &= c_{i,j}^{(\mu)} \ (i+j=\mu, j>0), \\ t_{i,j-1} &= c_{i,j}^{(i+j)} \ (2 \leq i+j \leq \mu-1), \\ c_{i,j}^{(i+j)} &= \frac{(i+j)!}{i!j!} + \sum_{\substack{2\leq p+q\leq i+j-1, q\geq 1 \\ p+k=i, q+l-1=j}} c_{p,q}^{(p+q)}\cdot\frac{(k+l)!}{k!l!} \ (2 < i+j \leq \mu, j \geq 1), \end{aligned} \tag{3.20}$$

for initializing $c_{i,j}^{(2)} = \frac{(i+j)!}{i!j!}$ $(i+j=2, j\geq 1)$.

Lemma 3.14. *If $\gamma_\mu\|y-x\| \leq \frac{1}{2}$, when $0 \leq \varphi \leq \arctan\frac{1}{\sqrt{\mu-1}}$, we have*

$$\|A^{-1}f(y)\| \geq 2\gamma_\mu^\mu\|y-x\|^\mu\left(h(\varphi) - \|y-x\|\right), \tag{3.21}$$

where

$$h(\varphi) = \frac{\cos^\mu\varphi - \sum_{i+j=\mu,j>0} c_{i,j}\gamma_\mu^{\mu-1}\cos^i\varphi\sin^j\varphi}{\sum_{1\leq i\leq\mu-2} 2t_{i,0}\gamma_\mu^\mu + \sum_{1\leq i+j\leq\mu-2,j>0} 2t_{i,j}\gamma_\mu^\mu\cos^i\varphi\sin^j\varphi + 2\gamma_\mu^\mu}, \tag{3.22}$$

where $c_{i,j}, t_{i,j-1} \in \mathbb{R}$ are constants computed inductively according to (3.20).

In order to generalize the results in Lemma 3.3 to simple multiple zeros of arbitrary multiplicity μ, we intend to find a suitable $\theta \in (0, \arctan\frac{1}{\sqrt{\mu-1}}]$ such that $h(\varphi) \geq h(\theta) \geq \frac{\sin\theta}{2\gamma_\mu}$ for $0 \leq \varphi \leq \theta$. Let d be defined by $d = \min(d_1, d_2, d_3)$, where $d_1 = \sqrt{\frac{1}{c_{\mu-1,1}^2+1}}$, $d_2 = \sqrt{\frac{1}{\mu-1}}$, and d_3 is the smallest positive real root of the function

$$p(d) = (1-d^2)^{\frac{\mu}{2}} - \sum_{i+j=\mu, j>0} c_{i,j} d (1-d^2)^{\frac{i}{2}} d^{j-1} - d\left(\sum_{1\le i\le \mu-2} t_{i,0} + \sum_{1\le i+j\le \mu-2, j>0} t_{i,j}(1-d^2)^{\frac{i}{2}} d^j + 1\right). \tag{3.23}$$

Lemmas 3.15 and 3.16 prove that $\theta = \arcsin \frac{d}{\gamma_\mu^{\mu-1}}$ is valid for generalizing the results in Lemma 3.3 to simple multiple zeros of arbitrary multiplicity μ.

Lemma 3.15. *Let θ be defined by $\sin\theta = \frac{d}{\gamma_\mu^{\mu-1}}$. Then $h(\theta) \ge \frac{\sin\theta}{2\gamma_\mu}$, where $h(\varphi)$ is defined in* (3.22).

Proof. By substituting $\sin\theta = \frac{d}{\gamma_\mu^{\mu-1}}$ and $\cos\theta = \left(1 - \frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{1/2}$ into (3.22), we need to show that

$$\left(1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{\frac{\mu}{2}} - c_{\mu-1,1} d \left(1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{\frac{\mu-1}{2}} - \sum_{i+j=\mu-1, j>0} c_{i,j+1} d \left(1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{\frac{i}{2}} \frac{d^j}{\gamma_\mu^{j(\mu-1)}} - d\left(\sum_{1\le i\le\mu-2} t_{i,0} + \sum_{1\le i+j\le\mu-2, j>0} t_{i,j}\left(1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{\frac{i}{2}} \frac{d^j}{\gamma_\mu^{j(\mu-1)}} + 1\right) \ge 0. \tag{3.24}$$

- The sum of the first two terms in (3.24) is increasing in γ_μ and nonnegative for $\gamma_\mu \ge 1$ and $d \le d_1 = \sqrt{\frac{1}{c_{\mu-1,1}^2+1}}$, since it equals
$$\left(1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{\frac{\mu-1}{2}} \left(\sqrt{1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}} - c_{\mu-1,1} d\right).$$
- The terms $\cos^i\varphi \sin^j\varphi$ $(j>0)$ are increasing for $\varphi \in \left[0,\ \arctan\sqrt{\frac{1}{i}}\right]$, since $(\cos^i\varphi\sin\varphi)' = \cos^{i-1}\varphi(\cos^2\varphi - i\sin^2\varphi) \ge 0$. Hence, for $1 \le i+j \le \mu-1$ and $j>0$, $\left(1-\frac{d^2}{\gamma_\mu^{2(\mu-1)}}\right)^{\frac{i}{2}} \frac{d^j}{\gamma_\mu^{j(\mu-1)}}$ is decreasing in γ_μ for $d \in \left[0,\ \sqrt{\frac{1}{\mu-1}}\right]$.
- The left-hand side of (3.24) is increasing in γ_μ, therefore it is sufficient to prove $p(d) \ge 0$ when $\gamma_\mu = 1$. The conclusion holds as $p(d)$ is decreasing in d for $0 \le d \le d_3$ and $p(0) = 1$.

□

Lemma 3.16. *For $0 \le \varphi \le \theta$, $h(\varphi)$ in* (3.22) *is nonnegative and decreasing.*

Proof. For $\varphi \in \left[0,\ \arctan\sqrt{\frac{1}{\mu-1}}\ \right]$, because $\cos^i\varphi\sin^j\varphi$ is increasing for $i+j=\mu$ and $j>0$, the numerator of $h(\varphi)$ is nonnegative and decreasing, and the denominator of $h(\varphi)$ is positive and increasing. Hence, $h(\varphi)$ is nonnegative and decreasing for $0 \le \varphi \le \theta$. □

According to Lemma 3.13, the coefficients $c_{i,j}$ and $t_{i,j}$ can be computed iteratively for arbitrary high multiplicities. It is worth noting that the value d is independent of the polynomial system f. Table 1 shows the values of d for $2 \leq \mu \leq 8$.

TABLE 1. Values of d for $2 \leq \mu \leq 8$

μ	2	3	4	5	6	7	8
d	0.2865	0.08507	0.02171	0.005043	0.001105	0.0002336	0.00004809

The following four theorems generalize the results in Section 3.1 to simple multiple zeros of arbitrary high multiplicities.

Theorem 3.17. *Let x be a simple multiple zero of f of multiplicity μ. If y is another zero of f, then*

$$\|y-x\| \geq \frac{d}{2\gamma_\mu^\mu}.$$

Proof. For $\theta \leq \varphi \leq \frac{\pi}{2}$, by Lemma 3.12, we conclude that

$$\|y-x\| \geq \frac{\sin\varphi}{2\gamma_\mu} \geq \frac{\sin\theta}{2\gamma_\mu} = \frac{d}{2\gamma_\mu^\mu}.$$

For $0 \leq \varphi \leq \theta$, by Lemmas 3.14, 3.15, and 3.16, we have

$$\begin{aligned}\|A^{-1}f(y)\| &\geq 2\gamma_\mu^\mu\|y-x\|^\mu\left(h(\varphi)-\|y-x\|\right) \geq 2\gamma_\mu^\mu\|y-x\|^\mu\left(h(\theta)-\|y-x\|\right)\\ &\geq 2\gamma_\mu^\mu\|y-x\|^\mu\left(\frac{\sin\theta}{2\gamma_\mu}-\|y-x\|\right).\end{aligned}$$

□

Theorem 3.18. *Let x be a simple multiple zero of f of multiplicity μ. If $\|y-x\| \leq \frac{d}{4\gamma_\mu^\mu}$, then we have*

$$\|f(y)\| \geq \frac{d\|y-x\|^\mu}{2\|A^{-1}\|}.$$

Proof. For $\theta \leq \varphi \leq \frac{\pi}{2}$, by Lemma 3.12, we derive

$$\|A^{-1}f(y)\| \geq \frac{1}{\sqrt{2}}\|D\hat{f}(x)^{-1}\hat{f}(y)\| \geq \sqrt{2}\gamma_\mu\|y-x\|\left(\frac{\sin\theta}{2\gamma_\mu}-\|y-x\|\right).$$

For $0 \leq \varphi \leq \theta$, by Lemma 3.14, 3.15, and 3.16, we derive

$$\|A^{-1}f(y)\| \geq 2\gamma_\mu^\mu\|y-x\|^\mu\left(\frac{\sin\theta}{2\gamma_\mu}-\|y-x\|\right).$$

When $\|y-x\| \leq \frac{d}{4\gamma_\mu^\mu} = \frac{\sin\theta}{4\gamma_\mu}$, we conclude that

$$\|A^{-1}f(y)\| \geq 2\gamma_\mu^\mu\|y-x\|^\mu\frac{\sin\theta}{4\gamma_\mu} = \frac{d\|y-x\|^\mu}{2}.$$

□

Theorem 3.19. *Let x be a simple multiple zero of f of multiplicity μ, and let $0 < R \leq \frac{d}{4\gamma_\mu^\mu}$. If*

$$d_R(f,g) < \frac{dR^\mu}{2\|A^{-1}\|},$$

then the sum of multiplicities of zeros of g in $B(x,R)$ is μ.

Proof. By Theorem 3.18, for any y such that $\|y-x\| = R < \frac{d}{4\gamma_\mu^\mu}$, we derive

$$\|f(y) - g(y)\| \leq d_R(f,g) < \frac{dR^\mu}{2\|A^{-1}\|} = \frac{d\|y-x\|^\mu}{2\|A^{-1}\|} \leq \|f(y)\|.$$

Then by Rouché's Theorem, f and g have the same number of zeros inside $B(x,R)$. By Theorem 3.17, when $R \leq \frac{d}{4\gamma_\mu^3}$, the only zero of f in $B(x,R)$ is x, which is of multiplicity μ. Therefore, g has μ zeros in $B(x,R)$. $\square$

Suppose x is an approximately given simple multiple zero of f of multiplicity μ and $Df(x)$ is of normal form approximately, i.e., $f(x)$, the entries of $Df(x)$'s first column and last row, $|\Delta_k(f_n)|$ $(2 \leq k \leq \mu-1)$ are small with respect to a given tolerance but $|\Delta_\mu(f_n)|$ is not, then we propose a numerical criterion for isolating a cluster of μ zeros in the neighborhood of x from the other zeros of f.

Let $H_1 = \begin{pmatrix} \frac{\partial \hat{f}(x)}{\partial X_1} & 0 \\ \frac{\partial f_n(x)}{\partial X_1} & \frac{\partial f_n(x)}{\partial \hat{X}} \end{pmatrix}$, and let

$$H_k = \left(\begin{pmatrix} 0 & 0 \\ 0 & \Delta_k(f_n) \end{pmatrix} \ \mathbf{0}_{\underbrace{n \times \cdots \times n}_{k} \times (n-1)} \right) \in \mathbb{C}^{\overbrace{n\times\cdots\times n}^{k+1}}, 2 \leq k \leq \mu - 1. \tag{3.25}$$

We construct a new system, $g(X) = f(X) - f(x) - \sum_{1\leq k\leq\mu-1} H_k(X-x)^k$.

Theorem 3.20. *Let $\gamma_\mu = \gamma_\mu(g,x)$. If*

$$\|f(x)\| + \sum_{1\leq k\leq\mu-1} \|H_k\| \left(\frac{d}{4\gamma_\mu^\mu}\right)^k < \frac{d^{\mu+1}}{2\,(4\gamma_\mu^\mu)^\mu \,\|A^{-1}\|}, \tag{3.26}$$

then f has μ zeros (counting multiplicities) in the ball of radius $\frac{d}{4\gamma_\mu^\mu}$ around x.

Proof. It is clear that x is a zero of g and $Dg(x)$ is of normal form, since $Dg(x) = Df(x) - H_1 = \begin{pmatrix} 0 & D\hat{f}(x) \\ 0 & 0 \end{pmatrix}$, $\Delta_k(g_n) = 0$, $2 \leq k \leq \mu-1$, $\Delta_\mu(g_n) = \Delta_\mu(f_n) \neq 0$. Hence, x is a simple multiple zero of g of multiplicity μ.

Let $R = \frac{d}{4\gamma_\mu^\mu(g,x)}$; we derive

$$\begin{aligned} d_R(g,f) &= \max_{\|y-x\|\leq R} \|g(y) - f(y)\| = \max_{\|y-x\|\leq R} \|f(x) + \sum_{1\leq k\leq\mu-1} H_k(X-x)^k\| \\ &\leq \|f(x)\| + \sum_{1\leq k\leq\mu-1} \|H_k\| R^k = \|f(x)\| + \sum_{1\leq k\leq\mu-1} \|H_k\| \left(\frac{d}{4\gamma_\mu^\mu}\right)^k. \end{aligned}$$

If $\|f(x)\| + \sum_{1\leq k\leq\mu-1} \|H_k\| \left(\frac{d}{4\gamma_\mu^\mu}\right)^k < \frac{d^{\mu+1}}{2(4\gamma_\mu{}^\mu)^\mu\|A^{-1}\|}$, then

$$d_R(g,f) < \frac{d^{\mu+1}}{2\,(4\gamma_\mu{}^\mu)^\mu\,\|A^{-1}\|} = \frac{dR^\mu}{2\|A^{-1}\|}.$$

By Theorem 3.19, the sum of multiplicities of zeros of f in $B(x,R)$ is μ. $\square$

4. Experiments

In this section, we present the numerical experiments for isolating simple multiple zeros or clusters of zeros of polynomial systems. All experiments are done in Maple 17 on a desktop computer with Intel (R) Core (TM) i5-3470S CPU @ 2.90GHz and RAM of 8GB running 64-bite Windows 8. The Maple environment variable is set by the statement "`UseHardwareFloats:=false`". All timings are measured as elapsed time in seconds. The Maple codes of our algorithms and all test results are available at `http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/SimpleMultipleZeros/`.

We first present some computational details. The local separation bound $r = \frac{d}{2\gamma_\mu^\mu}$ given in Theorem 3.17 depends on d and γ_μ. For a given multiplicity μ, the value d is determined by finding the smallest positive real root of the function $p(d)$ defined by (3.23), which is independent of the given polynomial system f (see Table 1 for $2 \le \mu \le 8$). The value γ_μ involves the calculation of values $\left\| D\hat{f}(x)^{-1} \cdot \frac{D^k \hat{f}(x)}{k!} \right\|$ and $\left\| \frac{1}{\Delta_\mu(f_n)} \cdot \frac{D^k f_n(x)}{k!} \right\|$, where $\Delta_\mu(f_n)$ is computed according to Proposition 2.1. It is well known that computing the operator norm of tensors of order larger than two is NP-hard [22]. Therefore, in our implementation, we use $\|\cdot\|_\infty$ to calculate an upper bound of γ_μ according to [15, Lemma B.2]. It is also possible to get an easily computable bound of the operator norm of tensors according to [13, Lemma 9.1]. Since the numerator d is monotonically decreasing on μ and the denominator contains the μth power of γ_μ, the separation bound $r = \frac{d}{2\gamma_\mu^\mu}$ becomes less tight as μ grows. Hence, in order to guarantee the criterion (3.26) to be satisfied, we may need to run the algorithm in a large number of digits to obtain an approximate zero x of high accuracy.

We test some benchmark examples in the literature and list their results in Table 2. All polynomial systems f are given with n equations and n unknowns. We use `deg` to denote the largest total degree of the polynomials in f, $\|f(x)\|$ and $\|Df(x)v\|$ show the quality of the approximate zero x, r denotes the radius of the ball $B(x,r)$ containing a cluster of μ zeros, `Digits` denotes the Maple environment variable used in the software-defined floating point arithmetic, and `time` measures the elapsed time for outputting a successful certification of $B(x,r)$ from inputs f, x, and μ. As shown in Table 2, in order to satisfy the criterion (3.26), the required accuracy of x might be around r^μ.

When x is only given with low accuracy such that it fails to validate the criterion (3.26), then we need to run the algorithm in [31] for refining x to a higher accuracy and try the certification procedure again. In Table 3, we show the number of steps of iterations required for a valid certification when the accuracy of x is around 10^{-3} and the total elapsed time for refining and certifying. For most examples listed in Table 3, it only requires a few seconds to successfully certify the existence of a cluster of μ zeros of f near x, even if the multiplicity is high (e.g., LiZhi2) or the size of the system is large (e.g., LiZhi1).

We also test two examples to compare our algorithm with [10] and [15].

TABLE 2. Isolating simple multiple zeros or clusters of zeros.

Systems	n	deg	Digits	$\|f(x)\|$	$\|Df(x)v\|$	μ	r	time
Ojika1 [37]	2	2	20	2.26 e-020	4.04 e-021	3	5.68 e-06	0.047
Ojika2 [37]	3	2	16	9.60 e-011	5.57 e-011	2	1.10 e-03	0.031
Ojika3 [39]	3	3	16	7.96 e-013	1.32 e-014	2	2.23 e-04	0.016
Ojika4 [38]	3	6	16	1.15 e-031	8.52 e-033	3	1.93 e-09	0.203
Decker2 [6]	2	4	16	8.95 e-013	8.01 e-025	4	1.09 e-02	0.031
DZ4 [5]	3	4	20	9.83 e-061	3.69 e-121	5	3.50 e-11	0.062
DZ3 [5]	2	3	30	4.56 e-029	1.77 e-028	5	4.51 e-05	0.047
RuGr09 [43]	2	3	16	6.47 e-021	1.48 e-041	4	2.71 e-03	0.031
LiZhi1 [30]	10	3	16	2.70 e-050	5.72 e-101	3	2.29 e-13	1.172
LiZhi2 [30]	3	3	100	8.28 e-161	4.98 e-360	8	5.59 e-14	0.125
GLSY1 [15]	3	5	30	6.26 e-050	2.17 e-051	4	8.15 e-11	0.125

TABLE 3. Refining and isolating simple multiple zeros or clusters of zeros.

Systems	$\|f(x)\|$	$\|Df(x)v\|$	r	step	time
Ojika1 [37]	2.26 e-03	4.03 e-04	5.68 e-06	3	0.157
Ojika2 [37]	9.60 e-04	5.57 e-04	1.10 e-03	2	0.047
Ojika3 [39]	7.92 e-04	1.31 e-05	2.23 e-04	2	0.078
Ojika4 [38]	1.18 e-02	8.44 e-04	1.93 e-09	4	0.437
Decker2 [6]	4.72 e-04	2.23 e-07	1.09 e-02	2	0.141
DZ4 [5]	1.49 e-03	1.02 e-06	3.50 e-11	5	0.485
DZ3 [5]	2.64 e-03	5.36 e-06	4.51 e-05	3	0.266
RuGr09 [43]	2.46 e-04	2.71 e-08	2.71 e-03	3	0.172
LiZhi1 [30]	1.69 e-03	4.56 e-07	2.29 e-13	5	2.094
LiZhi2 [30]	8.29 e-04	2.57 e-10	5.59 e-14	6	0.813
GLSY1 [15]	2.39 e-03	3.24 e-04	8.15 e-11	4	0.453

Example 4.1. Given a simple double zero $x = (0, 0)$ of a polynomial system,

$$f = \begin{cases} X_1^2 - \frac{1}{4}X_1 - \frac{1}{2}X_2, \\ \frac{1}{2}X_1X_2. \end{cases} \tag{4.1}$$

It is clear that f has another zero $y = (1/4, 0)$, and the actual minimal distance of two zeros of f is $\|y - x\| = 0.25$.

For the method in [10], it is easy to check $v = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{pmatrix} \in \ker Df(x)$ and

$$A = Df(x) + \frac{1}{2}D^2f(x)(v, \Pi_v) = \begin{pmatrix} -\frac{1}{4} + \frac{8}{5\sqrt{5}} & -\frac{1}{2} - \frac{4}{5\sqrt{5}} \\ -\frac{2}{5\sqrt{5}} & \frac{1}{5\sqrt{5}} \end{pmatrix}$$

is invertible. Then

$$A^{-1}\frac{D^2f(x)}{2} = \left(\begin{pmatrix} -\frac{4}{5} & -\frac{2(25+8\sqrt{5})}{10\sqrt{5}} \\ -\frac{8}{5} & -\frac{-25+32\sqrt{5}}{20\sqrt{5}} \end{pmatrix} \begin{pmatrix} -\frac{2(25+8\sqrt{5})}{10\sqrt{5}} & 0 \\ -\frac{-25+32\sqrt{5}}{20\sqrt{5}} & 0 \end{pmatrix} \right).$$

The computation of the norm of the order 3 tensor $A^{-1}\frac{D^2f(x)}{2}$ is challenging since it is NP-hard [22]. Instead of bounding it by the infinity norm, for this example, using our SOS certificates for global optima of polynomials and rational functions [23], we can verify that

$$\gamma_2(f,x) = \max\left(1, \left\|A^{-1}\frac{D^2f(x)}{2}\right\|\right) \geq 3.1121.$$

Therefore, the local separation bound computed by the method in [10] satisfies

$$\frac{d}{2\gamma_2(f,x)^2} \leq 0.01546,$$

for $d \approx 0.2976$.

Remark 4.2. We found two typos in [10]. The coefficient of the second $\sqrt{1-d^2}$ is $-2d$ instead of $-d$ in [10, Lemma 4]. Therefore, we have $d \approx 0.2976$. Furthermore, the degree of $\gamma_2(f,x)$ is 2 instead of 1 in [10, Theorem 1].

We compute two unit vectors $v = \begin{bmatrix} -\frac{2\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} \end{bmatrix}$ and $u = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ such that $Df(x)v = Df(x)^Tu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Then use the Gram–Schmidt process to get $v_1 = \begin{bmatrix} -\frac{\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} \end{bmatrix}$ and $u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ such that

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot Df(x) \cdot \begin{bmatrix} -\frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \\ \frac{\sqrt{5}}{5} & -\frac{2\sqrt{5}}{5} \end{bmatrix} = \begin{bmatrix} 0 & \frac{\sqrt{5}}{4} \\ 0 & 0 \end{bmatrix},$$

which is of normal form. We perform the unitary transformation to get a simple double zero $z = (0,0)$ of

$$g = \begin{cases} \frac{4}{5}X_1^2 + \frac{4}{5}X_1X_2 + \frac{2}{5}X_2^2 + \frac{\sqrt{5}}{4}X_3, \\ -\frac{1}{5}X_1^2 + \frac{3}{10}X_1X_2 + \frac{1}{5}X_2^2. \end{cases}$$

By (2.6) and (2.7), $\Delta_2(g_2) = d_1^2(g_2) = -\frac{1}{5}$. As $Dg(z)$ is of normal form, we calculate that

$$\hat{\gamma}_2 = \hat{\gamma}_2(g,z) = \max\left(1, \left\|\frac{4}{\sqrt{5}} \cdot \frac{D^2g_1(z)}{2}\right\|\right) = \frac{4}{\sqrt{5}},$$
$$\gamma_{2,2} = \gamma_{2,2}(g,z) = \max\left(1, \left\|5 \cdot \frac{D^2g_2(z)}{2}\right\|\right) = \frac{5}{4},$$

thus $\gamma_2 = \gamma_2(g,z) = \max(\hat{\gamma}_2, \gamma_{2,2}) = \frac{4}{\sqrt{5}}$.

Our local separation bound $\frac{d}{2\gamma_2^2} \approx 0.04478$ is tighter than $\frac{d}{2\gamma_2(f,x)^2} \leq 0.01546$ obtained using the method in [10]. Although our local separation bound is smaller than the actual distance 0.25 of two zeros of f, it is much larger than the global separation bound $\ll 10^{-10}$ obtained by using the method in [12].

Example 4.3. Given a simple double zero $x = (0, 0)$ and a nearby simple multiple zero, $y = (10^{-3k}, -10^{-2k})$ of multiplicity l of a polynomial system

$$f = \begin{cases} X_1^2 + X_2^3 + X_1 + 10^{-k} X_2, \\ (X_1 + 10^{-k} X_2) \cdot (X_1 - 10^{-3k})^{l-1}, \end{cases} \tag{4.2}$$

where k controls the distance between two zeros and l controls the multiplicity of y.

We compare our algorithm with the algorithm in [15] for computing a certified ball near the simple double zero x such that it contains two zeros (counting multiplicities). We list the minimal digits needed (`Digits`) for an approximate zero of limited accuracy $\|x\|$ for outputting a successful certification r in Table 4.

TABLE 4. Isolating a cluster of two zeros of (4.2)

	algorithm in [15]			our algorithm		
k, l	Digits	$\|x\|$	r	Digits	$\|x\|$	r
$k = 2$, $l = 2$	16	7.49 e-15	4.31 e-09	16	1.08 e-44	1.40 e-21
$k = 2$, $l = 5$	36	8.08 e-34	2.02 e-18	16	8.96 e-47	8.78 e-23
$k = 2$, $l = 10$	66	7.07 e-66	7.72 e-35	16	9.84 e-49	1.73 e-23
$k = 3$, $l = 2$	16	1.10 e-21	4.16 e-11	16	8.16 e-65	1.43 e-31
$k = 3$, $l = 5$	46	4.15 e-48	1.75 e-25	16	1.01 e-66	8.94 e-33
$k = 3$, $l = 10$	92	1.29 e-94	1.72 e-48	16	8.13 e-69	1.77 e-33
$k = 4$, $l = 2$	16	6.93 e-25	9.00 e-14	16	7.68 e-85	1.43 e-41
$k = 4$, $l = 5$	60	8.51 e-62	4.90 e-32	16	6.08 e-87	8.95 e-43
$k = 4$, $l = 10$	120	3.80 e-125	8.18 e-63	16	9.58 e-89	1.77 e-43

As shown in Table 4, when we fix the multiplicity l of y and let k grow, then the distance between two zeros x and y will decrease. Therefore, the value of $\|x\|$ and the radius r of the certified ball containing a cluster of two zeros will get smaller for both algorithms. However, when we fix k and let l grow to increase the multiplicity of y, the value of $\|x\|$ and r do not vary too much for our algorithm, while these values decrease fast for the algorithm in [15]. Furthermore, when k or l grows, the algorithm in [15] has to increase the necessary number of digits, while our algorithm succeeds with the fixed precision. Since our method is based on computing the local dual space at x, this makes our algorithm less sensitive to the singularity of nearby zeros. The algorithm in [15] reduces looking for simple multiple zeros of f to looking for zeros of a univariate analytic function via the implicit function theorem. Therefore, they might need more digits to guarantee the success of the reduction. On the other hand, since $y = (10^{-3k}, -10^{-2k})$ is another zero of f near to $x = (0, 0)$, the value of Δ_μ is smaller than 10^{-10} for $k = 2$. Hence, the radius r we computed is much smaller than the one computed by the algorithm in [15] for $l = 2$ and $l = 5$, and the value of $\|x\|$ is also much smaller. Similarly, the values of $\|x\|$ and r computed by the algorithm in [15] are larger than ours for $k = 3, 4$ and $l = 2, 5$. We would like to seek better ways to compute the local separation bound when there are nearby clusters around the simple multiple zero x.

5. Appendix

We give the proof of Lemma 3.2.

Proof. Since $\frac{\partial f_n(x)}{\partial X_1} = \cdots = \frac{\partial f_n(x)}{\partial X_n} = \frac{\partial^2 f_n(x)}{\partial X_1^2} = 0$, the Taylor series yields

$$f_n(y) = \left(\frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}}\zeta\eta + \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}\eta^2\right) + \frac{1}{6}\frac{\partial^3 f_n(x)}{\partial X_1^3}\zeta^3 + \frac{1}{2}\frac{\partial^3 f_n(x)}{\partial X_1^2 \partial \hat{X}}\zeta^2\eta + \frac{1}{2}\frac{\partial^3 f_n(x)}{\partial X_1 \partial \hat{X}^2}\zeta\eta^2 + \frac{1}{6}\frac{\partial^3 f_n(x)}{\partial \hat{X}^3}\eta^3 + \sum_{k\geq 4}\frac{D^k f_n(x)(y-x)^k}{k!}.$$

Substituting the rightmost η in $\left(\frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}}\zeta + \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}\eta\right)\eta$ by the expansion of η (3.6), as $\Delta_3(f_n) \neq 0$, it implies that

$$\begin{aligned}
\frac{1}{\Delta_3(f_n)}f_n(y) &= \frac{1}{\Delta_3(f_n)}\frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}}D\hat{f}(x)^{-1}\hat{f}(y)\zeta + \frac{1}{\Delta_3(f_n)}\frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}D\hat{f}(x)^{-1}\hat{f}(y)\eta \\
&+ \zeta^3 + \frac{1}{\Delta_3(f_n)}C_{2,1}\zeta^2\eta + \frac{1}{\Delta_3(f_n)}C_{1,2}\zeta\eta^2 + \frac{1}{\Delta_3(f_n)}C_{0,3}\eta^3 \\
&+ \sum_{k\geq 4}\frac{1}{\Delta_3(f_n)}\frac{D^k f_n(x)(y-x)^k}{k!} + \frac{1}{\Delta_3(f_n)}T_{1,0}\sum_{k\geq 3}D\hat{f}(x)^{-1}\frac{D^k\hat{f}(x)(y-x)^k}{k!}\zeta \\
&+ \frac{1}{\Delta_3(f_n)}T_{0,1}\sum_{k\geq 3}D\hat{f}(x)^{-1}\frac{D^k\hat{f}(x)(y-x)^k}{k!}\eta,
\end{aligned}$$

where

$$\begin{aligned}
C_{2,1} &= \frac{1}{2}\frac{\partial^3 f_n(x)}{\partial X_1^2 \partial \hat{X}} - \frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}}\cdot D\hat{f}(x)^{-1}\frac{\partial^2 \hat{f}(x)}{\partial X_1 \partial \hat{X}} - \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}\cdot D\hat{f}(x)^{-1}\frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial X_1^2}, \\
C_{1,2} &= \frac{1}{2}\frac{\partial^3 f_n(x)}{\partial X_1 \partial \hat{X}^2} - \frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}}\cdot D\hat{f}(x)^{-1}\frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial \hat{X}^2} - \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}\cdot D\hat{f}(x)^{-1}\frac{\partial^2 \hat{f}(x)}{\partial X_1 \partial \hat{X}}, \\
C_{0,3} &= \frac{1}{6}\frac{\partial^3 f_n(x)}{\partial \hat{X}^3} - \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}\cdot D\hat{f}(x)^{-1}\frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial \hat{X}^2}, \\
T_{1,0} &= -\frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}},\ T_{0,1} = -\frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}.
\end{aligned}$$

For the classical operator norm, the following inequalities hold for $i + j = k$:

$$\left\|\frac{\partial^k \hat{f}(x)}{\partial X_1^i \partial \hat{X}^j}\right\| \leq \|D^k \hat{f}(x)\|, \quad \left\|\frac{\partial^k f_n(x)}{\partial X_1^i \partial \hat{X}^j}\right\| \leq \|D^k f_n(x)\|.$$

Therefore, after moving ζ^3 to the left side and moving $\frac{1}{\Delta_3(f_n)}f_n(y)$ to the right side of the expansion, by the triangle inequalities and $\gamma_3(f,x)\|y-x\| \leq \frac{1}{2}$, we have

$$\begin{aligned}
|\zeta|^3 \leq &\left|\frac{1}{\Delta_3(f_n)}f_n(y)\right| + (2\gamma_{3,n}|\zeta| + \gamma_{3,n}\|\eta\|)\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| + 8\gamma_3^2|\zeta|^2\|\eta\| \\
&+ 7\gamma_3^2|\zeta|\|\eta\|^2 + 2\gamma_3^2\|\eta\|^3 + 2\gamma_3^3\|y-x\|^4 + 4\gamma_3^3\|y-x\|^3|\zeta| + 2\gamma_3^3\|y-x\|^3\|\eta\|.
\end{aligned}$$

Substituting $|\zeta| = \|y-x\|\cos\varphi,\ \|\eta\| = \|y-x\|\sin\varphi$ in the inequality, we derive

$$\begin{aligned}\|y-x\|^3\cos^3\varphi \le &\left|\frac{1}{\Delta_3(f_n)}f_n(y)\right| + \gamma_{3,n}\|y-x\|(2\cos\varphi+\sin\varphi)\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\|\\ &+8\gamma_3^2\|y-x\|^3\cos^2\varphi\sin\varphi + 7\gamma_3^2\|y-x\|^3\cos\varphi\sin^2\varphi + 2\gamma_3^2\|y-x\|^3\sin^3\varphi\\ &+2\gamma_3^3\|y-x\|^4(1+2\cos\varphi+\sin\varphi).\end{aligned}$$

Since $0 \le \varphi \le \arctan\frac{1}{\sqrt{2}}$, we have

$$1 \le 2\cos\varphi + \sin\varphi \le \sqrt{5}.$$

Then above inequality implies

$$\begin{aligned}&\left(\frac{\cos^3\varphi - 8\gamma_3^2\cos^2\varphi\sin\varphi - 7\gamma_3^2\cos\varphi\sin^2\varphi - 2\gamma_3^2\sin^3\varphi}{1+2\cos\varphi+\sin\varphi}\right)\|y-x\|^3 - 2\gamma_3^3\|y-x\|^4\\ &\le \frac{1}{1+2\cos\varphi+\sin\varphi}\left|\frac{1}{\Delta_3(f_n)}f_n(y)\right| + \frac{\gamma_{3,n}\|y-x\|(2\cos\varphi+\sin\varphi)}{1+2\cos\varphi+\sin\varphi}\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\|\\ &\le \left|\frac{1}{\Delta_3(f_n)}f_n(y)\right| + \frac{\sqrt{5}}{2+2\sqrt{5}}\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\|\\ &\le \left\|\begin{pmatrix}\frac{1}{\sqrt{2}}D\hat{f}(x)^{-1} & 0\\ 0 & \frac{\sqrt{2}}{\Delta_3(f_n)}\end{pmatrix}\begin{pmatrix}\hat{f}(y)\\ f_n(y)\end{pmatrix}\right\| = \|A^{-1}f(y)\|.\end{aligned}$$

This finishes the proof of Lemma 3.2. □

The proof of Lemma 3.13 is given below.

Proof. Let us apply the differential functional Δ_t (2.6) to both sides of (3.18):

$$\begin{aligned}\Delta_t(f_n) = &C_2\Delta_t(\zeta^2) + \cdots + C_\mu\Delta_t(\zeta^\mu) + \sum_{i+j=\mu, j>0} C_{i,j}\Delta_t(\zeta^i\eta^j)\\ &+\sum_{1\le i+j\le\mu-2} T_{i,j}\cdot\left(\sum_{k\ge\mu+1-i-j} D\hat{f}(x)^{-1}\Delta_t\left(\frac{D^k\hat{f}(x)(y-x)^k}{k!}\zeta^i\eta^j\right)\right)\\ &-\sum_{1\le i+j\le\mu-2} T_{i,j}D\hat{f}(x)^{-1}\Delta_t\left(\hat{f}(y)\zeta^i\eta^j\right) + \sum_{k\ge\mu+1}\Delta_t\left(\frac{D^kf_n(x)(y-x)^k}{k!}\right).\end{aligned}$$

According to (2.2), (2.4), and the fact that d_1^t is the only differential monomial of the highest order t in Δ_t and no other d_1^s with $s < t$ in Δ_t, for $2 \le t \le \mu$, we have:

(1) $\Delta_t(\zeta^s) = 1$ if $s = t$ and 0 otherwise;
(2) $\Delta_t(\zeta^i\eta^j) = 0$ for $t \le i+j = \mu$ and $j > 0$;
(3) $\Delta_t\left(\frac{D^k\hat{f}(x)(y-x)^k}{k!}\zeta^i\eta^j\right) = 0$ for $t \le \mu < i+j+k$;
(4) $\Delta_t\left(\hat{f}(y)\zeta^i\eta^j\right) = 0$ for $1 \le i+j \le \mu-2$;
(5) $\Delta_t\left(\frac{D^kf_n(x)(y-x)^k}{k!}\right) = 0$ for $t \le \mu < k$.

Hence, we conclude that $C_t = \Delta_t(f_n)$ for $t = 2, \ldots, \mu$.

Now let us show (3.19). By Taylor's expansion of $f_n(y)$ at x, we derive

$$f_n(y) = \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial X_1^2}\zeta^2 + \frac{\partial^2 f_n(x)}{\partial X_1 \partial \hat{X}}\zeta\eta + \frac{1}{2}\frac{\partial^2 f_n(x)}{\partial \hat{X}^2}\eta^2 + \cdots + \frac{1}{\mu!}\frac{\partial^\mu f_n(x)}{\partial X_1^\mu}\zeta^\mu + \frac{1}{(\mu-1)!}\frac{\partial^\mu f_n(x)}{\partial X_1^{\mu-1}\partial \hat{X}}\zeta^{\mu-1}\eta + \cdots + \frac{1}{\mu!}\frac{\partial^\mu f_n(x)}{\partial \hat{X}^\mu}\eta^\mu + \sum_{k\geq \mu+1}\frac{D^k f_n(x)(y-x)^k}{k!}.$$

Note that the coefficient of the term $\zeta^i\eta^j$ is $\frac{1}{i!j!}\frac{\partial^{i+j} f_n(x)}{\partial X_1^i \partial \hat{X}^j}$, which satisfies

$$\left\| \frac{1}{\Delta_\mu(f_n)} \cdot \frac{1}{i!j!}\frac{\partial^{i+j} f_n(x)}{\partial X_1^i \partial \hat{X}^j} \right\| \leq \frac{(i+j)!}{i!j!}\gamma_\mu^{i+j-1}. \tag{5.1}$$

For the monomial $\zeta^i\eta^j$, $i+j<\mu$, and $j>0$, after substituting one η in $\zeta^i\eta^j$ by

$$\eta = -D\hat{f}(x)^{-1}\left(\frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial X_1^2}\zeta^2 + \frac{\partial^2 \hat{f}(x)}{\partial X_1\partial \hat{X}}\zeta\eta + \frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial \hat{X}^2}\eta^2 + \cdots + \sum_{0\leq k\leq \mu+1-i-j}\frac{1}{(\mu+1-i-j-k)!k!}\frac{\partial^{\mu+1-i-j}\hat{f}(x)}{\partial X_1^{\mu+1-i-j-k}\partial \hat{X}^k}\zeta^{\mu+1-i-j-k}\eta^k + \sum_{k\geq \mu+2-i-j}\frac{D^k\hat{f}(x)(y-x)^k}{k!} - \hat{f}(y)\right),$$

we have

$$\zeta^i\eta^j = -D\hat{f}(x)^{-1}\left(\frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial X_1^2}\zeta^{i+2}\eta^{j-1} + \frac{\partial^2 \hat{f}(x)}{\partial X_1\partial \hat{X}}\zeta^{i+1}\eta^j + \frac{1}{2}\frac{\partial^2 \hat{f}(x)}{\partial \hat{X}^2}\zeta^i\eta^{j+1} + \cdots + \sum_{0\leq k\leq \mu+1-i-j}\frac{1}{(\mu+1-i-j-k)!k!}\frac{\partial^{\mu+1-i-j}\hat{f}(x)}{\partial X_1^{\mu+1-i-j-k}\partial \hat{X}^k}\zeta^{\mu+1-j-k}\eta^{k+j-1} + \sum_{k+i+j-1\geq \mu+1}\frac{D^k\hat{f}(x)(y-x)^k}{k!}\zeta^i\eta^{j-1} - \hat{f}(y)\zeta^i\eta^{j-1}\right), \tag{5.2}$$

where the total degree of each term in the above expression is at least $i+j+1$. Moreover, the norm of the coefficient of the new term $\zeta^{i+k}\eta^{j-1+l}$, $i+k+j-1+l\leq \mu$, obtained after the substitution and divided by $\Delta_\mu(f_n)$. is bounded by

$$\left(\frac{(i+j)!}{i!j!}\gamma_\mu^{i+j-1}\right)\left\| D\hat{f}(x)^{-1}\frac{1}{k!l!}\frac{\partial^{k+l}\hat{f}(x)}{\partial X_1^k\partial \hat{X}^l}\right\| \leq \frac{(i+j)!}{i!j!}\frac{(k+l)!}{k!l!}\gamma_\mu^{i+k+j+l-2}. \tag{5.3}$$

For simplicity, we replace $j-1$ by j in the last two terms of (3.18). The iterative formula for $c_{i,j}$ is obtained by (5.1) and (5.3). Let $C_{i,j}^{(i+j)}$ denote the coefficient of $\zeta^i\eta^j$ after $i+j-2$ substitutions (i.e., the coefficient of $\zeta^i\eta^j$ when x is a simple multiple zero of multiplicity $i+j$). Then we conclude that

$$\left\|\frac{1}{\Delta_{i+j}(f_n)}C_{i,j}^{(i+j)}\right\| \leq c_{i,j}^{(i+j)}\gamma_{i+j}^{i+j-1}.$$

On the other hand, by (5.2), we derive

$$T_{i,j-1} = -C_{i,j}^{(i+j)}$$

and

$$\left\| \frac{1}{\Delta_\mu(f_n)} T_{i,j-1} \right\| \leq c_{i,j}^{(i+j)} \gamma_{i+j}^{i+j-1}.$$

Hence, we have $t_{i,j-1} = c_{i,j}^{(i+j)}$ for $2 \leq i+j \leq \mu - 1$. $\square$

Now let us prove Lemma 3.14.

Proof. Since x is a simple multiple zero of f of multiplicity μ, by Lemma 3.13, we have $C_2 = \cdots = C_{\mu-1} = 0$ and $C_\mu = \Delta_\mu(f_n) \neq 0$. By (3.18), we derive that

$$\begin{aligned}
\zeta^\mu = &- \frac{1}{\Delta_\mu(f_n)} \sum_{i+j=\mu, j>0} C_{i,j} \zeta^i \eta^j - \frac{1}{\Delta_\mu(f_n)} \sum_{k \geq \mu+1} \frac{D^k f_n(x)(y-x)^k}{k!} \\
&- \frac{1}{\Delta_\mu(f_n)} \sum_{1 \leq i+j \leq \mu-2} T_{i,j} \cdot \left(\sum_{k \geq \mu+1-i-j} D\hat{f}(x)^{-1} \frac{D^k \hat{f}(x)(y-x)^k}{k!} \zeta^i \eta^j \right) \\
&+ \frac{1}{\Delta_\mu(f_n)} \sum_{1 \leq i+j \leq \mu-2} T_{i,j} D\hat{f}(x)^{-1} \hat{f}(y) \zeta^i \eta^j + \frac{1}{\Delta_\mu(f_n)} f_n(y).
\end{aligned}$$

By the triangle inequality and Lemma 3.13, we have

$$\begin{aligned}
|\zeta|^\mu \leq &\sum_{i+j=\mu, j>0} \left\| \frac{1}{\Delta_\mu(f_n)} C_{i,j} \right\| |\zeta|^i \|\eta\|^j + \sum_{k \geq \mu+1} \left\| \frac{1}{\Delta_\mu(f_n)} \frac{D^k f_n(x)}{k!} \right\| \|y-x\|^k \\
&+ \sum_{1 \leq i+j \leq \mu-2} \left\| \frac{1}{\Delta_\mu(f_n)} T_{i,j} \right\| \cdot \left(\sum_{k \geq \mu+1-i-j} \left\| D\hat{f}(x)^{-1} \frac{D^k \hat{f}(x)}{k!} \right\| \|y-x\|^k |\zeta|^i \|\eta\|^j \right) \\
&+ \sum_{1 \leq i+j \leq \mu-2} \left\| \frac{1}{\Delta_\mu(f_n)} T_{i,j} \right\| \cdot \left\| D\hat{f}(x)^{-1} \hat{f}(y) \right\| |\zeta|^i \|\eta\|^j + \left| \frac{1}{\Delta_\mu(f_n)} f_n(y) \right| \\
\leq &\sum_{i+j=\mu, j>0} c_{i,j} \gamma_\mu^{i+j-1} |\zeta|^i \|\eta\|^j + \sum_{k \geq \mu+1} \gamma_{\mu,n}^{k-1} \|y-x\|^k \\
&+ \sum_{1 \leq i+j \leq \mu-2} t_{i,j} \gamma_\mu^{i+j} \cdot 2\hat{\gamma}_\mu^{\mu-i-j} \|y-x\|^{\mu-i-j+1} |\zeta|^i \|\eta\|^j \\
&+ \sum_{1 \leq i+j \leq \mu-2} t_{i,j} \gamma_\mu^{i+j} \cdot \left\| D\hat{f}(x)^{-1} \hat{f}(y) \right\| |\zeta|^i \|\eta\|^j + \left| \frac{1}{\Delta_\mu(f_n)} f_n(y) \right| \\
\leq &\sum_{i+j=\mu, j>0} c_{i,j} \gamma_\mu^{\mu-1} |\zeta|^i \|\eta\|^j + \sum_{1 \leq i+j \leq \mu-2} 2t_{i,j} \gamma_\mu^\mu \|y-x\|^{\mu-i-j+1} |\zeta|^i \|\eta\|^j + 2\gamma_\mu^\mu \|y-x\|^{\mu+1} \\
&+ \sum_{1 \leq i+j \leq \mu-2} t_{i,j} \gamma_\mu^{i+j} \cdot \left\| D\hat{f}(x)^{-1} \hat{f}(y) \right\| |\zeta|^i \|\eta\|^j + \left| \frac{1}{\Delta_\mu(f_n)} f_n(y) \right|.
\end{aligned}$$

By $|\zeta| = \|y-x\|\sin\varphi,\ \|\eta\| = \|y-x\|\cos\varphi$, we obtain

$$\begin{aligned}
\|y-x\|^{\mu}\cos^{\mu}\varphi \leq & \sum_{i+j=\mu, j>0} c_{i,j}\gamma_{\mu}^{\mu-1}\|y-x\|^{\mu}\cos^{i}\varphi\sin^{j}\varphi \\
& + \sum_{1\leq i+j\leq \mu-2} 2t_{i,j}\gamma_{\mu}^{\mu}\|y-x\|^{\mu+1}\cos^{i}\varphi\sin^{j}\varphi + 2\gamma_{\mu}^{\mu}\|y-x\|^{\mu+1} \\
& + \sum_{1\leq i+j\leq \mu-2} t_{i,j}\gamma_{\mu}^{i+j}\|y-x\|^{i+j}\cos^{i}\varphi\sin^{j}\varphi \cdot \left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| + \left|\frac{1}{\Delta_{\mu}(f_n)}f_n(y)\right|.
\end{aligned}$$

Therefore, we conclude that

$$\begin{aligned}
& \frac{\cos^{\mu}\varphi - \sum_{i+j=\mu, j>0} c_{i,j}\gamma_{\mu}^{\mu-1}\cos^{i}\varphi\sin^{j}\varphi}{1+\sum_{1\leq i+j\leq\mu-2} t_{i,j}\cos^{i}\varphi\sin^{j}\varphi}\cdot\|y-x\|^{\mu} - 2\gamma_{\mu}^{\mu}\|y-x\|^{\mu+1} \\
\leq & \frac{1}{1+\sum_{1\leq i+j\leq\mu-2} t_{i,j}\cos^{i}\varphi\sin^{j}\varphi}\left|\frac{1}{\Delta_{\mu}(f_n)}f_n(y)\right| \\
& \quad + \frac{\sum_{1\leq i+j\leq\mu-2} t_{i,j}\gamma_{\mu}^{i+j}\|y-x\|^{i+j}\cos^{i}\varphi\sin^{j}\varphi}{1+\sum_{1\leq i+j\leq\mu-2} t_{i,j}\cos^{i}\varphi\sin^{j}\varphi}\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| \\
\leq & \left|\frac{1}{\Delta_{\mu}(f_n)}f_n(y)\right| + \frac{1}{2}\left\|D\hat{f}(x)^{-1}\hat{f}(y)\right\| \leq \|A^{-1}f(y)\|.
\end{aligned}$$

Now it is clear that

$$\|A^{-1}f(y)\| \geq 2\gamma_{\mu}^{\mu}\|y-x\|^{\mu}\left(h(\varphi) - \|y-x\|\right),$$

for $h(\varphi)$ defined by (3.22). $\square$

Acknowledgments

The authors are very grateful to Grégoire Lecerf for his Maple code and valuable comments. We also thank Elias Tsigaridas for fruitful discussions at the Fields Institute in Toronto during the Major Program in Computer Algebra.

References

[1] T. Ayyildiz Akoglu, J. D. Hauenstein, and A. Szanto, *Certifying solutions to overdetermined and singular polynomial systems over* $\mathbb{Q}$, J. Symbolic Comput. **84** (2018), 147–171, DOI 10.1016/j.jsc.2017.03.004. MR3673011

[2] C. A. Berenstein, R. Gay, A. Vidras, and A. Yger, *Residue Currents and Bezout identities*, Progress in Mathematics, vol. 114, Birkhäuser Verlag, Basel, 1993. MR1249478

[3] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*, Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp. MR1479636

[4] B. H. Dayton, T.-Y. Li, and Z. Zeng, *Multiple zeros of nonlinear systems*, Math. Comp. **80** (2011), no. 276, 2143–2168, DOI 10.1090/S0025-5718-2011-02462-2. MR2813352

[5] B. H. Dayton and Z. Zeng, *Computing the Multiplicity Structure in Solving Polynomial systems*, ISSAC'05, ACM, New York, 2005, pp. 116–123, DOI 10.1145/1073884.1073902. MR2280537

[6] D. W. Decker, H. B. Keller, and C. T. Kelley, *Convergence rates for Newton's method at singular points*, SIAM J. Numer. Anal. **20** (1983), no. 2, 296–314, DOI 10.1137/0720020. MR694520

[7] D. W. Decker and C. T. Kelley, *Newton's method at singular points. I*, SIAM J. Numer. Anal. **17** (1980), no. 1, 66–70, DOI 10.1137/0717009. MR559463

[8] D. W. Decker and C. T. Kelley, *Newton's method at singular points. II*, SIAM J. Numer. Anal. **17** (1980), no. 3, 465–471, DOI 10.1137/0717039. MR581492

[9] D. W. Decker and C. T. Kelley, *Convergence acceleration for Newton's method at singular points*, SIAM J. Numer. Anal. **19** (1982), no. 1, 219–229, DOI 10.1137/0719012. MR646604

[10] J.-P. Dedieu and M. Shub, *On simple double zeros and badly conditioned zeros of analytic functions of* n *variables*, Math. Comp. **70** (2001), no. 233, 319–327, DOI 10.1090/S0025-5718-00-01194-7. MR1680867

[11] J. Dian and R. B. Kearfott, *Existence verification for singular and nonsmooth zeros of real nonlinear systems*, Math. Comp. **72** (2003), no. 242, 757–766, DOI 10.1090/S0025-5718-02-01427-8. MR1954966

[12] I. Z. Emiris, B. Mourrain, and E. P. Tsigaridas, *The DMM bound: Multivariate (Aggregate) Separation Bounds*, ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2010, pp. 243–250, DOI 10.1145/1837934.1837981. MR2920560

[13] S. Friedland and L.-H. Lim, *Nuclear norm of higher-order tensors*, Math. Comp. **87** (2018), no. 311, 1255–1281, DOI 10.1090/mcom/3239. MR3766387

[14] M. Giusti, G. Lecerf, B. Salvy, and J.-C. Yakoubsohn, *On location and approximation of clusters of zeros of analytic functions*, Found. Comput. Math. **5** (2005), no. 3, 257–311, DOI 10.1007/s10208-004-0144-z. MR2168678

[15] M. Giusti, G. Lecerf, B. Salvy, and J.-C. Yakoubsohn, *On location and approximation of clusters of zeros: case of embedding dimension one*, Found. Comput. Math. **7** (2007), no. 1, 1–49, DOI 10.1007/s10208-004-0159-5. MR2283341

[16] Marc Giusti and Jean-Claude Yakoubsohn, *Numerical approximation of multiple isolated roots of analytic systems*. Manuscript available at `https://arxiv.org/abs/1707.06301`, 2017.

[17] A. Griewank, *On solving nonlinear equations with simple singularities or nearly singular solutions*, SIAM Rev. **27** (1985), no. 4, 537–563, DOI 10.1137/1027141. MR812453

[18] A. Griewank and M. R. Osborne, *Newton's method for singular problems when the dimension of the null space is* > 1, SIAM J. Numer. Anal. **18** (1981), no. 1, 145–149, DOI 10.1137/0718011. MR603436

[19] Zhiwei Hao, Wenlong Jiang, Nan Li, and Lihong Zhi, *Computing simple multiple zeros of polynomial systems*. Manuscript available at `https://www.arxiv.org/pdf/1703.03981.pdf`, 2017.

[20] J. D. Hauenstein, B. Mourrain, and A. Szanto, *Certifying Isolated Singular Points and Their Multiplicity Structure*, ISSAC'15—Proceedings of the 2015 ACM International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2015, pp. 213–220. MR3388302

[21] J. D. Hauenstein, B. Mourrain, and A. Szanto, *On deflation and multiplicity structure*, J. Symbolic Comput. **83** (2017), 228–253, DOI 10.1016/j.jsc.2016.11.013. MR3645652

[22] C. J. Hillar and L.-H. Lim, *Most tensor problems are NP-hard*, J. ACM **60** (2013), no. 6, Art. 45, 39, DOI 10.1145/2512329. MR3144915

[23] E. L. Kaltofen, B. Li, Z. Yang, and L. Zhi, *Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients*, J. Symbolic Comput. **47** (2012), no. 1, 1–15, DOI 10.1016/j.jsc.2011.08.002. MR2854844

[24] Yuchi Kanzawa and Shin'ichi Oishi, *Imperfect singular solutions of nonlinear equations and a numerical method of proving their existence*, IEICE Transactions on Fundamentals of Electronics, Communications and Computer sciences, 82(6):1062–1069, 1999.

[25] R. B. Kearfott and J. Dian, *Existence verification for higher degree singular zeros of nonlinear systems*, SIAM J. Numer. Anal. **41** (2003), no. 6, 2350–2373, DOI 10.1137/S0036142901386057. MR2034619

[26] R. B. Kearfott, J. Dian, and A. Neumaier, *Existence verification for singular zeros of complex nonlinear systems*, SIAM J. Numer. Anal. **38** (2000), no. 2, 360–379, DOI 10.1137/S0036142999361074. MR1770053

[27] G. Lecerf, *Quadratic Newton iteration for systems with multiplicity*, Found. Comput. Math. **2** (2002), no. 3, 247–293, DOI 10.1007/s102080010026. MR1907381

[28] A. Leykin, J. Verschelde, and A. Zhao, *Newton's method with deflation for isolated singularities of polynomial systems*, Theoret. Comput. Sci. **359** (2006), no. 1-3, 111–122, DOI 10.1016/j.tcs.2006.02.018. MR2251604

[29] A. Leykin, J. Verschelde, and A. Zhao, *Higher-Order Deflation for Polynomial Systems with Isolated Singular Solutions*, Algorithms in algebraic geometry, IMA Vol. Math. Appl., vol. 146, Springer, New York, 2008, pp. 79–97, DOI 10.1007/978-0-387-75155-9_5. MR2397938

[30] N. Li and L. Zhi, *Computing the multiplicity structure of an isolated singular solution: case of breadth one*, J. Symbolic Comput. **47** (2012), no. 6, 700–710, DOI 10.1016/j.jsc.2011.12.027. MR2908589

[31] N. Li and L. Zhi, *Computing isolated singular solutions of polynomial systems: case of breadth one*, SIAM J. Numer. Anal. **50** (2012), no. 1, 354–372, DOI 10.1137/110827247. MR2888317

[32] N. Li and L. Zhi, *Verified error bounds for isolated singular solutions of polynomial systems: case of breadth one*, Theoret. Comput. Sci. **479** (2013), 163–173, DOI 10.1016/j.tcs.2012.10.028. MR3034563

[33] N. Li and L. Zhi, *Verified error bounds for isolated singular solutions of polynomial systems*, SIAM J. Numer. Anal. **52** (2014), no. 4, 1623–1640, DOI 10.1137/120902914. MR3229659

[34] A. Mantzaflaris and B. Mourrain, *Deflation and Certified Isolation of Singular Zeros of Polynomial Systems*, ISSAC 2011—Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2011, pp. 249–256, DOI 10.1145/1993886.1993925. MR2895219

[35] Maria Grazia Marinari, Teo Mora, and Hans Michael Möller, *Gröbner duality and multiplicities in polynomial system solving*, in Proceedings of the 1995 International Symposium on Symbolic and Algebraic Computation, ISSAC '95, pages 167–179, New York, NY, USA, 1995. ACM.

[36] B. Mourrain, *Isolated points, duality and residues*, J. Pure Appl. Algebra **117/118** (1997), 469–493, DOI 10.1016/S0022-4049(97)00023-6. Algorithms for algebra (Eindhoven, 1996). MR1457851

[37] T. Ojika, *Modified deflation algorithm for the solution of singular problems. I. A system of nonlinear algebraic equations*, J. Math. Anal. Appl. **123** (1987), no. 1, 199–221, DOI 10.1016/0022-247X(87)90304-0. MR881541

[38] T. Ojika, *A numerical method for branch points of a system of nonlinear algebraic equations*, Appl. Numer. Math. **4** (1988), no. 5, 419–430, DOI 10.1016/0168-9274(88)90018-9. MR948507

[39] T. Ojika, S. Watanabe, and T. Mitsui, *Deflation algorithm for the multiple roots of a system of nonlinear equations*, J. Math. Anal. Appl. **96** (1983), no. 2, 463–479, DOI 10.1016/0022-247X(83)90055-0. MR719330

[40] L. B. Rall, *Convergence of the Newton process to multiple solutions*, Numer. Math. **9** (1966), 23–37, DOI 10.1007/BF02165226. MR0210316

[41] G. W. Reddien, *On Newton's method for singular problems*, SIAM J. Numer. Anal. **15** (1978), no. 5, 993–996, DOI 10.1137/0715064. MR507559

[42] G. W. Reddien, *Newton's method and high order singularities*, Comput. Math. Appl. **5** (1979), no. 2, 79–86, DOI 10.1016/0898-1221(79)90061-0. MR539566

[43] S. M. Rump and S. Graillat, *Verified error bounds for multiple roots of systems of nonlinear equations*, Numer. Algorithms **54** (2010), no. 3, 359–377, DOI 10.1007/s11075-009-9339-3. MR2661328

[44] Y.-Q. Shen and T. J. Ypma, *Newton's method for singular nonlinear equations using approximate left and right nullspaces of the Jacobian*, Appl. Numer. Math. **54** (2005), no. 2, 256–265, DOI 10.1016/j.apnum.2004.09.029. MR2148044

[45] Y.-Q. Shen and T. J. Ypma, *Solving rank-deficient separable nonlinear equations*, Appl. Numer. Math. **57** (2007), no. 5-7, 609–615, DOI 10.1016/j.apnum.2006.07.025. MR2322434

[46] M. Shub and S. Smale, *Complexity of Bezout's theorem. IV. Probability of success; extensions*, SIAM J. Numer. Anal. **33** (1996), no. 1, 128–148, DOI 10.1137/0733008. MR1377247

[47] M. Shub and S. Smale, *Computational complexity. On the geometry of polynomials and a theory of cost. I*, Ann. Sci. École Norm. Sup. (4) **18** (1985), no. 1, 107–142. MR803197

[48] M. Shub and S. Smale, *Computational complexity: on the geometry of polynomials and a theory of cost. II*, SIAM J. Comput. **15** (1986), no. 1, 145–161, DOI 10.1137/0215011. MR822199

[49] S. Smale, *The fundamental theorem of algebra and complexity theory*, Bull. Amer. Math. Soc. (N.S.) **4** (1981), no. 1, 1–36, DOI 10.1090/S0273-0979-1981-14858-8. MR590817

[50] S. Smale, *Newton's Method Estimates from Data at One Point*, The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985), Springer, New York, 1986, pp. 185–196. MR870648

[51] H. J. Stetter, *Numerical Polynomial Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2004. MR2048781

[52] J.-C. Yakoubsohn, *Finding a cluster of zeros of univariate polynomials*, Complexity theory, real machines, and homotopy (Oxford, 1999), J. Complexity **16** (2000), no. 3, 603–638, DOI 10.1006/jcom.2000.0555. MR1787887

[53] J.-C. Yakoubsohn, *Simultaneous Computation of all the Zero-clusters of a Univariate Polynomial*, Foundations of computational mathematics (Hong Kong, 2000), World Sci. Publ., River Edge, NJ, 2002, pp. 433–455. MR2021992

[54] N. Yamamoto, *Regularization of solutions of nonlinear equations with singular Jacobian matrices*, J. Inform. Process. **7** (1984), no. 1, 16–21. MR760383

Key Laboratory of Mathematics Mechanization, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; and University of Chinese Academy of Sciences, Beijing 100049, China

Email address: haozhiwei@mmrc.iss.ac.cn

Key Laboratory of Mathematics Mechanization, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, and China; University of Chinese Academy of Sciences, Beijing 100049, China

Email address: jiangwr@amss.ac.cn

College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, China

Email address: nan.li@szu.edu.cn

Key Laboratory of Mathematics Mechanization, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; and University of Chinese Academy of Sciences, Beijing 100049, China

Email address: lzhi@mmrc.iss.ac.cn