COMPUTING ISOLATED SINGULAR SOLUTIONS OF POLYNOMIAL SYSTEMS: CASE OF BREADTH ONE*

NAN LI AND LIHONG ZHI †

Abstract. We present a symbolic-numeric method to refine an approximate isolated singular solution $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_n)$ of a polynomial system $F = \{f_1, \ldots, f_n\}$, when the Jacobian matrix of F evaluated at $\hat{\mathbf{x}}$ has corank one approximately. Our new approach is based on the regularized Newton iteration and the computation of differential conditions satisfied at the approximate singular solution. The size of matrices involved in our algorithm is bounded by $n \times n$. The algorithm converges quadratically if $\hat{\mathbf{x}}$ is close to the isolated exact singular solution.

Key words. Root refinement, isolated singular solution, regularized Newton iteration, local dual space, quadratic convergence.

AMS subject classifications.

1. Introduction.

Motivation and problem statement. Consider an ideal I generated by a polynomial system $F = \{f_1, \ldots, f_n\}$, where $f_i \in \mathbb{C}[x_1, \ldots, x_n]$. Suppose $\hat{\mathbf{x}} = \hat{\mathbf{x}}_e + \hat{\mathbf{x}}_\epsilon$, where $\hat{\mathbf{x}}_e$ denotes the isolated exact singular solution of F and $\hat{\mathbf{x}}_\epsilon$ denotes the error in the solution. The *multiplicity* μ of $\hat{\mathbf{x}}_e$ is defined as $\mu = \dim(\mathbb{C}[\mathbf{x}]/Q)$, where Q is the isolated primary component whose associate prime ideal is $P = (x_1 - \hat{x}_{1,e}, \ldots, x_n - \hat{x}_{n,e})$, and the *index* ρ of $\hat{\mathbf{x}}_e$ is defined as the minimal nonnegative integer ρ such that $P^\rho \subseteq Q$ [37].

In [38, 39], they compute the truncated coefficient matrix of the involutive system to the order ρ, and generate multiplication matrices from its approximate null vectors. Then a basis of the approximate local dual space (Definition 2.1) of I at $\hat{\mathbf{x}}$ can be obtained from these vectors (Theorem 5.4 in [38]). Let $\hat{\mathbf{y}}$ be the vector whose i-th element is the average of the trace of the multiplication matrix with respect to x_i. In [39], it has been proved that if the given approximation $\hat{\mathbf{x}}$ satisfies $\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \varepsilon$, for a small positive number ε, and the index ρ and the multiplicity μ are computed correctly, then the refined solution obtained by adding $\hat{\mathbf{y}}$ to $\hat{\mathbf{x}}$ will satisfy $\|\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e\| = \mathrm{O}(\varepsilon^2)$. Here and hereafter, $\|\cdot\|$ is denoted as the l^2-norm. The size of these coefficient matrices in [39] is bounded by $n\binom{\rho+n}{n} \times \binom{\rho+n}{n}$, which will be very big when ρ is large. Especially, when the corank of the Jacobian $F'(\hat{\mathbf{x}}_e)$ is one, then $\rho = \mu$, which is also called the breadth one case in [4, 5]. As pointed out in [10], the breath one case is the least degenerate one and therefore most likely to be of practical significance. Moreover, it is also the worst case for the deflation method [5, 18, 27, 28] since the deflation always terminates at step $\mu - 1$, hence the size of the matrices grows extremely fast with the multiplicity.

In [20], we present a new algorithm which is based on Stetter's strategies [35] for computing a closed basis $L = \{L_0, \ldots, L_{\mu-1}\}$ of the approximate local dual space of $I = (f_1, \ldots, f_n)$ at $\hat{\mathbf{x}}$ incrementally in the breadth one case. The size of matrices we used in computing each order of differential conditions is bounded by $n \times n$, which does not depend on the multiplicity. Moreover, during the computation, we only need to store the input polynomial system F, the last $n-1$ columns of the Jacobian $F'(\hat{\mathbf{x}})$

* This research is supported by the Chinese National Natural Science Foundation under grant NSFC60821002/F02, 60911130369 and 10871194.

† Key Lab of Mathematics Mechanization, AMSS, Beijing 100190, China, linan08@amss.ac.cn, lzhi@mmrc.iss.ac.cn

and the computed differential conditions. Therefore, in the breadth one case, both storage space and execution time for computing a closed basis of the approximate local dual space are reduced significantly by the algorithm in [20]. This motivates us to consider whether we can get rid of large coefficient matrices in [38, 39] and refine approximate singular solutions more efficiently based on the computed differential conditions.

Main contribution. Suppose we are given an approximate singular solution $\hat{\mathbf{x}}$ of a polynomial system F satisfying $\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \varepsilon$, where the positive number ε is small enough such that there are no other solutions of F nearby. We also assume that the corank of the Jacobian matrix $F'(\hat{\mathbf{x}}_e)$ is one. In order to restore the quadratic convergence of the Newton method, we first apply one regularized Newton iteration (in Section 3.1) to obtain a new approximation $\hat{\mathbf{x}}+\hat{\mathbf{y}}$ which also satisfies the assumptions above, and then compute the approximate null vector $\mathbf{r}_1$ of the Jacobian $F'(\hat{\mathbf{x}}+\hat{\mathbf{y}})$ which gives a generalized Newton direction, and the step length δ is obtained by solving a linear system formulated by the computed differential operators using the algorithm in [20]. We show that $\|\hat{\mathbf{x}}+\hat{\mathbf{y}}+\delta\mathbf{r}_1-\hat{\mathbf{x}}_e\| = \mathrm{O}(\varepsilon^2)$. The size of matrices involved in our algorithm is bounded by $n \times n$. The method has been implemented in Maple. Moreover, we also prove the conjecture in [5] that the breadth one depth-deflation always terminates at step $\mu - 1$, where μ is the multiplicity.

Structure of the paper. Section 2 is devoted to recall some notations and well-known facts. In Section 3, we describe an algorithm for refining approximate isolated singular solutions of polynomial systems in the breadth one case. Moreover, we prove that the algorithm converges quadratically if the approximate solution is close to the isolated exact singular solution. Some experiment results are given in Section 4. We mention some ongoing research in Section 5.

2. Preliminaries. Let $D(\alpha) = D(\alpha_1, \ldots, \alpha_n) : \mathbb{C}[\mathbf{x}] \to \mathbb{C}[\mathbf{x}]$ denote the differential operator defined by:

$$D(\alpha_1, \ldots, \alpha_n) := \frac{1}{\alpha_1! \cdots \alpha_n!} \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}},$$

for nonnegative integer array $\alpha = [\alpha_1, \ldots, \alpha_n]$. We write $\mathfrak{D} = \{D(\alpha),\ |\alpha| \geq 0\}$ and denote by $\mathrm{Span}_{\mathbb{C}}(\mathfrak{D})$ the $\mathbb{C}$-vector space generated by $\mathfrak{D}$. Introducing a morphism on $\mathfrak{D}$ that acts as “integral”:

$$\Phi_j(D(\alpha)) := \begin{cases} D(\alpha_1, \ldots, \alpha_j - 1, \ldots, \alpha_n), & \text{if } \alpha_j > 0, \\ 0, & \text{otherwise.} \end{cases}$$

As a counterpart of the anti-differentiation operator Φ_j, we define the differential operator Ψ_j as

$$\Psi_j(D(\alpha)) := D(\alpha_1, \ldots, \alpha_j + 1, \ldots, \alpha_n).$$

Definition 2.1. *Given a zero $\hat{\mathbf{x}}_e$ of an ideal $I = (f_1, \ldots, f_n)$, we define the local dual space of I at $\hat{\mathbf{x}}_e$ as*

$$\triangle_{\hat{\mathbf{x}}_e}(I) := \{L \in \mathrm{Span}_{\mathbb{C}}(\mathfrak{D}) \,|\, L(f)_{\mathbf{x}=\hat{\mathbf{x}}_e} = 0,\ \forall f \in I\}. \tag{2.1}$$

The vector space $\triangle_{\hat{\mathbf{x}}_e}(I)$ and conditions equivalent to $L(f)_{\mathbf{x}=\hat{\mathbf{x}}_e} = 0,\ \forall L \in \triangle_{\hat{\mathbf{x}}_e}(I)$ are also called *Max Noether space* and *Max Noether conditions* in [25] respectively. Notice that the local dual space defined here consists of differential operators instead of differential functionals in [5, 21, 22, 24, 25, 26]. For a nonnegative integer k, $\triangle^{(k)}_{\hat{\mathbf{x}}_e}(I)$ is a subspace of $\triangle_{\hat{\mathbf{x}}_e}(I)$ which consists of differential operators with differential order bounded by k. Obviously, $\triangle^{(0)}_{\hat{\mathbf{x}}_e}(I) = D(0,\ldots,0)$. We have that

$$\dim_{\mathbb{C}}(\triangle_{\hat{\mathbf{x}}_e}(I)) = \mu, \tag{2.2}$$

where μ is the multiplicity of the zero $\hat{\mathbf{x}}_e$.

Definition 2.2. *[21, 22, 24] A subspace $\triangle$ of $\mathrm{Span}_{\mathbb{C}}(\mathfrak{D})$ is said to be closed if and only if its dimension is finite and*

$$L \in \triangle \Longrightarrow \Phi_j(L) \in \triangle,\ \ j = 1,\ldots,n.$$

Suppose $\mathrm{Span}(L_0, L_1, \ldots, L_{\mu-1})$ is closed and $L_0, \ldots, L_{\mu-1}$ are linearly independent differential operators which satisfy that $L_i(f_j)_{\mathbf{x}=\hat{\mathbf{x}}_e} = 0,\ j = 1,\ldots,n,\ i = 0,\ldots,\mu-1$, then due to the closedness, $L_i(q\cdot f_j)_{\mathbf{x}=\hat{\mathbf{x}}_e} = 0,\ \forall q \in \mathbb{C}[x_1,\ldots,x_n]$. Hence, $\triangle_{\hat{\mathbf{x}}_e}(I) = \mathrm{Span}(L_0, L_1, \ldots, L_{\mu-1})$.

Lemma 2.3. *Let $F'(\hat{\mathbf{x}}_e)$ be the Jacobian of a polynomial system $F = \{f_1,\ldots,f_n\}$ evaluated at $\hat{\mathbf{x}}_e$. Suppose the corank of $F'(\hat{\mathbf{x}}_e)$ is one, i.e., the dimension of its null space is one, then $\dim(\triangle^{(k)}_{\hat{\mathbf{x}}_e}(I)) = \dim(\triangle^{(k-1)}_{\hat{\mathbf{x}}_e}(I)) + 1$ for $1 \le k \le \mu-1$ and $\dim(\triangle^{(k)}_{\hat{\mathbf{x}}_e}(I)) = \dim(\triangle^{(\mu-1)}_{\hat{\mathbf{x}}_e}(I))$, for $k \ge \mu$. Hence we have $\mu = \rho$.*

Proof. Lemma 2.3 is an immediate consequence of [34, Theorem 2.2] and [5, Lemma 1]. □

Theorem 2.4. *[20] Suppose we are given an isolated multiple root $\hat{\mathbf{x}}_e$ of the polynomial system $F = \{f_1,\ldots,f_n\}$ with the multiplicity μ and the corank of the Jacobian $F'(\hat{\mathbf{x}}_e)$ is one, and $L_1 = D(1,0,\ldots,0) \in \triangle^{(1)}_{\hat{\mathbf{x}}_e}(I)$. We can construct the k-th order differential condition retaining the closedness incrementally for k from 2 to $\mu-1$ by the following formulas:*

$$L_k = P_k + a_{k,2}D(0,1,\ldots,0) + \cdots + a_{k,n}D(0,\ldots,1), \tag{2.3}$$

where P_k has no free parameters and is obtained from previous computed $L_1,\ldots,L_{k-1}$ by the following formula:

$$P_k = \Psi_1(L_{k-1}) + \Psi_2(Q_{k,2})_{\alpha_1=0} + \cdots + \Psi_n(Q_{k,n})_{\alpha_1=\alpha_2=\cdots=\alpha_{n-1}=0}, \tag{2.4}$$

where

$$\Phi_1(P_k) = L_{k-1},\ Q_{k,j} = \Phi_j(P_k) = a_{2,j}L_{k-2} + \cdots + a_{k-1,j}L_1,\ 2 \le j \le n. \tag{2.5}$$

Here $\Psi_j(Q_{k,j})_{\alpha_1=\cdots=\alpha_{j-1}=0}$ means that we only pick up differential operators $D(\alpha)$ in $Q_{k,j}$ where $\alpha_1 = \cdots = \alpha_{j-1} = 0$. The parameters $a_{k,j},\ j = 2,\ldots,n$ are determined by checking whether $[P_k(f_1)_{\mathbf{x}=\hat{\mathbf{x}}_e},\ldots,P_k(f_n)_{\mathbf{x}=\hat{\mathbf{x}}_e}]^T$ can be written as a linear combination of the last $n-1$ linearly independent columns of $F'(\hat{\mathbf{x}}_e)$.

Suppose $\hat{\mathbf{x}}$ is an approximation of $\hat{\mathbf{x}}_e$ and $\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \varepsilon \ll 1$, we can use the algorithm MultiplicityStructureBreadthOneNumeric in [20] to compute a closed basis $\{L_0,\ldots,L_{\mu-1}\}$ of the approximate local dual space of I at $\hat{\mathbf{x}}$. Since the errors in the matrix of the linear system

$$\left[P_k(F)_{\mathbf{x}=\hat{\mathbf{x}}}, \frac{\partial F(\hat{\mathbf{x}})}{\partial x_2}, \ldots, \frac{\partial F(\hat{\mathbf{x}})}{\partial x_n}\right] \cdot [1, a_{k,2}, \ldots, a_{k,n}]^T = 0,$$

used in Theorem 2.4 are bounded by $\mathrm{O}(\varepsilon)$ and

$$L_k = P_k + a_{k,2} D(0,1,0,\ldots,0) + \cdots + a_{k,n} D(0,\ldots,0,1),$$

is determined by its right singular vector $[1, a_{k,1}, \ldots, a_{k,n}]^T$ corresponding to its smallest singular value, we have

$$\|L_k(F)_{\mathbf{x}=\hat{\mathbf{x}}}\| = \mathrm{O}(\varepsilon), \tag{2.6}$$

according to [11, Corollary 8.6.2].

3. An Algorithm for Refining Approximate Singular Solutions. Suppose we are given an approximate solution

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}_e + \hat{\mathbf{x}}_\epsilon,$$

where $\hat{\mathbf{x}}_\epsilon$ denotes the error in the solution and $\hat{\mathbf{x}}_e$ denotes the exact solution of the polynomial system $F = \{f_1, \ldots, f_n\}$ with the multiplicity μ and the index ρ. In this section, we present a new method to refine $\hat{\mathbf{x}}$ in the breadth one case, i.e., $\mu = \rho$.

Let $A = F'(\hat{\mathbf{x}})$ be the Jacobian matrix of F evaluated at $\hat{\mathbf{x}}$ and $\mathbf{b} = -F(\hat{\mathbf{x}})$. Suppose the error in the solution is small enough, i.e., $\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \varepsilon \ll 1$, and A is invertible, then Newton's iteration computes

$$\hat{\mathbf{y}} = A^{-1}\mathbf{b}, \tag{3.1}$$

and $\|\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e\| = \mathrm{O}(\varepsilon^2)$ according to the well-known Kantorovich theorem [16]. However, if A is singular, as shown in [15], the convergence of Newton iterations for multivariate case is not guaranteed at irregular singularities.

Rall [29] studied the convergence properties of Newton's method at singular points. Some modifications of Newton's method to restore quadratic convergence have also been proposed in [1, 6, 7, 8, 12, 13, 14, 27, 28, 30, 31, 33]. In [13], a bordered system was introduced to restore the quadratic convergence of Newton's method when A has corank one approximately and $\hat{\mathbf{x}}$ is a simple singular solution. It is clear to see that the regularity condition in [13] can not be satisfied if the multiplicity is larger than 2.

For simplicity, we make an assumption throughout this section.

Assumption 1. *Suppose we are given an approximate singular solution $\hat{\mathbf{x}}$ of a polynomial system F satisfying $\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \varepsilon$, where the positive number ε is small enough such that there are no other solutions of F nearby. Moreover, we assume that the corank of the Jacobian matrix $F'(\hat{\mathbf{x}}_e)$ is one.*

Let $A = F'(\hat{\mathbf{x}})$ be the Jacobian matrix of F evaluated at $\hat{\mathbf{x}}$ and its singular values be $\sigma_1, \ldots, \sigma_s$. Under Assumption 1, we have $\|F(\hat{\mathbf{x}})\| = \mathrm{O}(\varepsilon)$, $\sigma_i = \Theta(1)$, $1 \le i \le n-1$ and $\sigma_n = \mathrm{O}(\varepsilon)$.

Remark 3.1. *Notice that the notation $\mathrm{O}(g)$ denotes that the value is bounded above by g up to a constant factor, while $\Theta(g)$ denotes that the value is bounded both above and below by g up to constant factors.*

3.1. Regularized Newton iteration. Under Assumption 1, $F'(\hat{\mathbf{x}})$ is approximately singular. Instead of using (3.1) to compute $\hat{\mathbf{y}}$, we solve the following damped least-squares problem

$$\min \|A\mathbf{y} - \mathbf{b}\|^2 + \lambda\|\mathbf{y}\|^2,$$

to obtain $\hat{\mathbf{y}}$, where $A = F'(\hat{\mathbf{x}})$ and $\mathbf{b} = -F(\hat{\mathbf{x}})$. The real number $\lambda > 0$ is called the regularization parameter [36].

THEOREM 3.2 (Regularized Newton Iteration). *Under Assumption 1, if we choose the smallest singular value σ_n of $F'(\hat{\mathbf{x}})$ as the regularization parameter, the solution $\hat{\mathbf{y}}$ of the following regularized least squares problem*

$$(A^*A + \sigma_n I_n)\hat{\mathbf{y}} = A^*\mathbf{b} \tag{3.2}$$

satisfies

$$\|\hat{\mathbf{y}}\| = \mathrm{O}(\varepsilon), \;\; \|F(\hat{\mathbf{x}} + \hat{\mathbf{y}})\| = \mathrm{O}(\varepsilon^2), \tag{3.3}$$

where A^* is the Hermitian (conjugate) transpose of $A = F'(\hat{\mathbf{x}})$, I_n is the $n \times n$ identity matrix and $\mathbf{b} = -F(\hat{\mathbf{x}})$.

Proof. Suppose $A = U \cdot \Sigma \cdot V^*$ is the singular value decomposition of A where $\Sigma = \mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$, then the solution of (3.2) is

$$\hat{\mathbf{y}} = V \cdot (\Sigma^2 + \sigma_n I_n)^{-1} \cdot \Sigma \cdot U^* \cdot \mathbf{b}. \tag{3.4}$$

Since $\sigma_i = \Theta(1), \; 1 \le i \le n-1$, $\sigma_n = \mathrm{O}(\varepsilon)$ and $\|\mathbf{b}\| = \mathrm{O}(\varepsilon)$, we have

$$\|\hat{\mathbf{y}}\|^2 = \sum_{i=1}^{n} \left(\frac{\sigma_i |\tilde{b}_i|}{\sigma_i^2 + \sigma_n} \right)^2 = \mathrm{O}(\varepsilon^2),$$

where $\tilde{\mathbf{b}} = [\tilde{b}_1, \ldots, \tilde{b}_n]^T = U^*\mathbf{b}$ and $\|\tilde{\mathbf{b}}\| = \|\mathbf{b}\| = \mathrm{O}(\varepsilon)$. Hence, $\|\hat{\mathbf{y}}\| = \mathrm{O}(\varepsilon)$.

From the Taylor expansion of F at $\hat{\mathbf{x}}$, we have

$$F(\hat{\mathbf{x}}_e) = -\mathbf{b} + A(\hat{\mathbf{x}}_e - \hat{\mathbf{x}}) + \mathrm{O}(\varepsilon^2).$$

Hence

$$\| -\mathbf{b} + A(\hat{\mathbf{x}}_e - \hat{\mathbf{x}})\| = \mathrm{O}(\varepsilon^2).$$

Furthermore, we have

$$\| -U^*\mathbf{b} + \Sigma \cdot V^*(\hat{\mathbf{x}}_e - \hat{\mathbf{x}})\| = \mathrm{O}(\varepsilon^2).$$

Since $\sigma_n = \mathrm{O}(\varepsilon)$ and $\|V^*(\hat{\mathbf{x}} - \hat{\mathbf{x}}_e)\| = \|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \varepsilon$, we derive that the last component of the vector $\tilde{\mathbf{b}}$ satisfies

$$|\tilde{b}_n| = \mathrm{O}(\varepsilon^2). \tag{3.5}$$

Since

$$A\hat{\mathbf{y}} - \mathbf{b} = U \cdot \mathrm{diag}\left\{ \frac{-\sigma_n}{\sigma_1^2 + \sigma_n}, \ldots, \frac{-\sigma_n}{\sigma_n^2 + \sigma_n} \right\} \cdot \tilde{\mathbf{b}},$$

we have

$$\|A\hat{\mathbf{y}} - \mathbf{b}\|^2 = \sum_{i=1}^{n} \left(\frac{\sigma_n |\tilde{b}_i|}{\sigma_i^2 + \sigma_n} \right)^2,$$

where

$$\frac{\sigma_n}{\sigma_i^2 + \sigma_n} = \mathrm{O}(\varepsilon), \text{ for } i = 1, \ldots, n-1,$$

and $\|\tilde{\mathbf{b}}\| = \mathrm{O}(\varepsilon)$. Although

$$\frac{\sigma_n}{\sigma_n^2 + \sigma_n} = \Theta(1),$$

we have from (3.5) that $|\tilde{b}_n| = \mathrm{O}(\varepsilon^2)$, hence

$$\|A\hat{\mathbf{y}} - \mathbf{b}\| = \mathrm{O}(\varepsilon^2). \tag{3.6}$$

Finally, from the Taylor expansion of F at $\hat{\mathbf{x}}$, we have

$$\|F(\hat{\mathbf{x}} + \hat{\mathbf{y}})\| \leq \| - \mathbf{b} + A\hat{\mathbf{y}}\| + \mathrm{O}(\varepsilon^2) = \mathrm{O}(\varepsilon^2).$$

□

According to Theorem 3.2, after applying one regularized Newton iteration to F and $\hat{\mathbf{x}}$, we get $\hat{\mathbf{y}}$ satisfies (3.3), and the new approximate singular solution $\hat{\mathbf{x}} + \hat{\mathbf{y}}$ satisfies

$$\|\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e\| \leq \|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| + \|\hat{\mathbf{y}}\| = \varepsilon + \mathrm{O}(\varepsilon).$$

If

$$\|\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e\| = \mathrm{O}(\varepsilon^2),$$

then we have already achieved the quadratic convergence. However, the convergence of the regularized Newton iteration is also not guaranteed when the Jacobian matrix is near singular. Hence, in most cases, we will have

$$\|\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e\| = \Theta(\varepsilon). \tag{3.7}$$

We show below how to restore the quadratic convergence when the computed approximate singular solution $\hat{\mathbf{x}} + \hat{\mathbf{y}}$ satisfies (3.3) and (3.7).

If $L_1 \in \triangle^{(1)}_{\hat{\mathbf{x}}+\hat{\mathbf{y}}}(I)$ is not $D(1, 0, \ldots, 0)$, as pointed out by Stetter in [35], we can compute the right singular vector of $F'(\hat{\mathbf{x}} + \hat{\mathbf{y}})$ corresponding to its smallest singular value σ'_n, denoted by $\mathbf{r}_1$ satisfying $\|\mathbf{r}_1\| = 1$ and

$$\|F'(\hat{\mathbf{x}} + \hat{\mathbf{y}})\, \mathbf{r}_1\| = \sigma'_n = \mathrm{O}(\varepsilon). \tag{3.8}$$

Let us form a unitary matrix $R = [\mathbf{r}_1, \ldots, \mathbf{r}_n]$ and perform the linear transformation

$$H(\mathbf{z}) = F(R\,\mathbf{z}). \tag{3.9}$$

It is clear that

$$\hat{\mathbf{z}}_e = R^{-1}\hat{\mathbf{x}}_e \tag{3.10}$$

is an exact root of $H(\mathbf{z})$ and

$$\hat{\mathbf{z}} = R^{-1}\,(\hat{\mathbf{x}} + \hat{\mathbf{y}}) \tag{3.11}$$

is an approximate root of $H(\mathbf{z})$. Moreover, we have

$$\|\hat{\mathbf{z}} - \hat{\mathbf{z}}_e\| = \|R^{-1}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e)\| = \|\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{x}}_e\| = \Theta(\varepsilon), \tag{3.12}$$

$$\|H(\hat{\mathbf{z}})\| = \|F(\hat{\mathbf{x}} + \hat{\mathbf{y}})\| = \mathrm{O}(\varepsilon^2), \tag{3.13}$$

and

$$\left\|\frac{\partial H(\hat{\mathbf{z}})}{\partial z_1}\right\| = \|F'(\hat{\mathbf{x}} + \hat{\mathbf{y}})\mathbf{r}_1\| = \sigma_n' = \mathrm{O}(\varepsilon). \tag{3.14}$$

Hence, the condition (3.7) is equivalent to (3.12). Here and hereafter, we always assume that $\hat{\mathbf{z}}$ satisfies

$$\|\hat{\mathbf{z}} - \hat{\mathbf{z}}_e\| = \Theta(\varepsilon). \tag{3.15}$$

THEOREM 3.3. *The root $\hat{\mathbf{z}}_e$ defined in (3.10) is an isolated singular solution of H with the multiplicity μ and the corank of $H'(\hat{\mathbf{z}}_e)$ is one.*

Proof. Since $H'(\hat{\mathbf{z}}_e) = F'(\hat{\mathbf{x}}_e)R$ and R is a unitary matrix, we derive that the corank of $H'(\hat{\mathbf{z}}_e)$ is one. Let μ' be the multiplicity of $\hat{\mathbf{z}}_e$, and $\{L_0, L_1, \ldots, L_{\mu'-1}\}$ be a closed basis of the local dual space of H at $\hat{\mathbf{z}}_e$. The operator $\Gamma_R : \mathrm{Span}_{\mathbb{C}}(\mathfrak{D}) \to \mathrm{Span}_{\mathbb{C}}(\mathfrak{D})$ is defined by:

$$\begin{aligned}
\Gamma_R(D(\alpha)) :&= \Gamma_R\left(\frac{1}{\alpha_1!\cdots\alpha_n!}\frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial z_1^{\alpha_1}\cdots\partial z_n^{\alpha_n}}\right)\\
&= \frac{1}{\alpha_1!\cdots\alpha_n!}\frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial(\mathbf{r}_1^*\cdot\mathbf{x})^{\alpha_1}\cdots\partial(\mathbf{r}_n^*\cdot\mathbf{x})^{\alpha_n}}\\
&= \frac{1}{\alpha_1!\cdots\alpha_n!}\sum_{|\beta|=|\alpha|} c_\beta\cdot\frac{\partial^{\beta_1+\cdots+\beta_n}}{\partial x_1^{\beta_1}\cdots\partial x_n^{\beta_n}},\\
&= \frac{1}{\alpha_1!\cdots\alpha_n!}\sum_{|\beta|=|\alpha|} c_\beta\cdot\beta_1!\cdots\beta_n!\cdot D(\beta),
\end{aligned}$$

where c_β is the coefficient of $\frac{\partial^{\beta_1+\cdots+\beta_n}}{\partial x_1^{\beta_1}\cdots\partial x_n^{\beta_n}}$ in the expansion of $\frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial(\mathbf{r}_1^*\cdot\mathbf{x})^{\alpha_1}\cdots\partial(\mathbf{r}_n^*\cdot\mathbf{x})^{\alpha_n}}$. Since $H(\mathbf{z}) = F(R\,\mathbf{z})$ and $\mathbf{x} = R\,\mathbf{z}$, according to multivariate chain rules, we have

$$\Gamma_R(L_k)(F)_{\mathbf{x}=\hat{\mathbf{x}}_e} = L_k(H)_{\mathbf{z}=\hat{\mathbf{z}}_e} = 0,$$

and for $1 \le j \le n$,

$$\begin{aligned}
\Phi_j(\Gamma_R(L_k)) =&\Gamma_R\left(\sum_{i=1}^{n} r_{i,j}\Phi_i(L_k)\right)\\
=&\Gamma_R\left(\sum_{i=1}^{k-2}(a_{k-i,2}r_{2,j} + \cdots + a_{k-i,n}r_{n,j})L_i \;+ r_{1,j}L_{k-1}\right)\\
=&\sum_{i=1}^{k-1}(a_{k-i,2}r_{2,j} + \cdots + a_{k-i,n}r_{n,j})\Gamma_R(L_i) \;+ r_{1,j}\Gamma_R(L_{k-1}),
\end{aligned}$$

where $0 \le k \le \mu' - 1$. Hence, $\{\Gamma_R(L_0), \Gamma_R(L_1), \ldots, \Gamma_R(L_{\mu'-1})\}$ is a closed basis of $\triangle_{\hat{\mathbf{x}}_e}^{(\mu'-1)}(I)$ and $\mu' \le \mu$. On the other hand, since $F(\mathbf{x}) = H(R^{-1}\mathbf{x})$, we derive that $\mu \le \mu'$. Hence, $\mu' = \mu$. □

Remark 3.4. *Since $H'(\hat{\mathbf{z}}) = F'(\hat{\mathbf{x}}+\hat{\mathbf{y}})R$ and R is a unitary matrix, we derive that the singular values of $H'(\hat{\mathbf{z}})$ are the same as those of $F'(\hat{\mathbf{x}}+\hat{\mathbf{y}})$ and the corank of $H'(\hat{\mathbf{z}})$ is one approximately. Suppose $\{L_0, L_1, \ldots, L_{\mu-1}\}$ is a closed basis of the approximate local dual space of H at $\hat{\mathbf{z}}$, where $L_0 = D(0, \ldots, 0)$ and $L_1 = D(1, 0, \ldots, 0)$. From the proof of Theorem 3.3 and (2.6), we have*

$$\Gamma_R(L_k)(F)_{\mathbf{x}=\hat{\mathbf{x}}+\hat{\mathbf{y}}} = L_k(H)_{\mathbf{z}=\hat{\mathbf{z}}} = \mathrm{O}(\varepsilon),$$

and

$$\Phi_j(\Gamma_R(L_k)) \in \mathrm{Span}\{\Gamma_R(L_1), \ldots, \Gamma_R(L_{k-1})\},$$

where $0 \le k \le \mu - 1$ and $1 \le j \le n$. Hence, $\{\Gamma_R(L_0), \Gamma_R(L_1), \ldots, \Gamma_R(L_{\mu-1})\}$ is a closed basis of $\triangle_{\hat{\mathbf{x}}+\hat{\mathbf{y}}}(I)$.

Remark 3.5. *It should be noticed that Theorem 3.3 holds as long as R is a regular matrix. However, if we choose a unitary matrix R, then it is much easier to compute the inverse of R since $R^{-1} = R^*$.*

It is interesting to notice that, after running one regularized Newton iteration, the last $n-1$ elements of the solution $\hat{\mathbf{z}}$ have already been refined quadratically.

Theorem 3.6. *Suppose $\hat{\mathbf{z}}_e$ and $\hat{\mathbf{z}}$ are defined in (3.10) and (3.11) respectively. Under Assumption 1, we have*

$$|\hat{z}_{1,e} - \hat{z}_1| = \Theta(\varepsilon), \tag{3.16}$$

and

$$|\hat{z}_{i,e} - \hat{z}_i| = \mathrm{O}(\varepsilon^2), \text{ for } i = 2, \ldots, n. \tag{3.17}$$

Proof. From the Taylor expansion of $H(\mathbf{z})$ at $\hat{\mathbf{z}}$, we have

$$H(\hat{\mathbf{z}}_e) = H(\hat{\mathbf{z}}) + H'(\hat{\mathbf{z}})(\hat{\mathbf{z}}_e - \hat{\mathbf{z}}) + \mathrm{O}(\varepsilon^2).$$

Since $H(\hat{\mathbf{z}}_e) = 0$ and $\|H(\hat{\mathbf{z}})\| = \mathrm{O}(\varepsilon^2)$, we have

$$\|H'(\hat{\mathbf{z}})(\hat{\mathbf{z}}_e - \hat{\mathbf{z}})\| = \mathrm{O}(\varepsilon^2).$$

From (3.14) and (3.15), we have

$$\left\| \frac{\partial H(\hat{\mathbf{z}})}{\partial z_1} (\hat{z}_{1,e} - \hat{z}_1) \right\| = \mathrm{O}(\varepsilon^2),$$

and

$$\left\| \left[\frac{\partial H(\hat{\mathbf{z}})}{\partial z_2}, \ldots, \frac{\partial H(\hat{\mathbf{z}})}{\partial z_n} \right] \cdot [\hat{z}_{2,e} - \hat{z}_2, \ldots, \hat{z}_{n,e} - \hat{z}_n]^T \right\| = \mathrm{O}(\varepsilon^2).$$

According to Remark 3.4, the matrix $\left[\frac{\partial H(\hat{\mathbf{z}})}{\partial z_2}, \ldots, \frac{\partial H(\hat{\mathbf{z}})}{\partial z_n} \right]$ is of full column rank, so that (3.17) is correct. The equation (3.16) follows from (3.15) and (3.17). □

If the multiplicity μ is larger than 2, the regularity assumption in [13] will not be satisfied. The violation of the regularity assumption is caused by the existence of the higher order differential condition. It is interesting to notice that the left singular vector of the Jacobian matrix $H'(\hat{\mathbf{z}})$ corresponding to the smallest singular value can be used to prove the following theorem.

THEOREM 3.7. *If the multiplicity of the singular root is larger than* 2*, under Assumption 1, we have*

$$\|L_1(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \left\|\frac{\partial H(\hat{\mathbf{z}})}{\partial z_1}\right\| = \mathrm{O}(\varepsilon^2).$$

Proof. If $\mu > 2$, according to Theorem 2.4 and (2.6), there exists a second order differential condition such that

$$\|L_2(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \left\|\left(\frac{1}{2}\frac{\partial^2}{\partial z_1^2} + a_{2,2}\frac{\partial}{\partial z_2} + \cdots + a_{2,n}\frac{\partial}{\partial z_n}\right)(H)_{\mathbf{z}=\hat{\mathbf{z}}}\right\| = \mathrm{O}(\varepsilon). \tag{3.18}$$

Suppose $\mathbf{u}_n$ is the left singular vector of $H'(\hat{\mathbf{z}})$ corresponding to the smallest singular value σ_n' and $\|\mathbf{u}_n\| = 1$, then

$$\left|\mathbf{u}_n^* \frac{\partial H(\hat{\mathbf{z}})}{\partial z_i}\right| = \mathrm{O}(\varepsilon),\ 1 \le i \le n. \tag{3.19}$$

From (3.18) and (3.19), we have

$$\left|\mathbf{u}_n^* \frac{\partial^2 H(\hat{\mathbf{z}})}{\partial z_1^2}\right| = \mathrm{O}(\varepsilon). \tag{3.20}$$

Therefore, we get

$$\left|\mathbf{u}_n^* \frac{\partial^2 H(\hat{\mathbf{z}})}{\partial z_1^2}(\hat{z}_{1,e} - \hat{z}_1)^2\right| = \mathrm{O}(\varepsilon^3). \tag{3.21}$$

From the Taylor expansion of $H(\mathbf{z})$ at $\hat{\mathbf{z}}$, we have

$$H(\hat{\mathbf{z}}_e) = H(\hat{\mathbf{z}}) + H'(\hat{\mathbf{z}})(\hat{\mathbf{z}}_e - \hat{\mathbf{z}}) + H''(\hat{\mathbf{z}})(\hat{\mathbf{z}}_e - \hat{\mathbf{z}})^2 + \mathrm{O}(\varepsilon^3), \tag{3.22}$$

where $(\hat{\mathbf{z}}_e - \hat{\mathbf{z}})^2$ denotes the vector of all monomials with degree 2 and $H''(\hat{\mathbf{z}})$ consists of all second order derivatives of H evaluated at $\hat{\mathbf{z}}$.

According to Theorem 3.6, all elements in $(\hat{\mathbf{z}}_e - \hat{\mathbf{z}})^2$ are $\mathrm{O}(\varepsilon^3)$ except the first one. Combining with (3.21), we have

$$\left|\mathbf{u}_n^* H''(\hat{\mathbf{z}})(\hat{\mathbf{z}}_e - \hat{\mathbf{z}})^2\right| = \mathrm{O}(\varepsilon^3). \tag{3.23}$$

On the other hand, the Taylor expansion of $H(\mathbf{z})$ at $\hat{\mathbf{z}}_e$ shows that

$$H(\hat{\mathbf{z}}) = H(\hat{\mathbf{z}}_e) + H'(\hat{\mathbf{z}}_e)(\hat{\mathbf{z}} - \hat{\mathbf{z}}_e) + H''(\hat{\mathbf{z}}_e)(\hat{\mathbf{z}} - \hat{\mathbf{z}}_e)^2 + \mathrm{O}(\varepsilon^3).$$

Since the corank of $H'(\hat{\mathbf{z}}_e)$ is one, suppose $\mathbf{u}_e$ is the left null vector of $H'(\hat{\mathbf{z}}_e)$ and $\|\mathbf{u}_e\| = 1$, then

$$\mathbf{u}_e^* H'(\hat{\mathbf{z}}_e) = 0.$$

Notice that

$$\|\mathbf{u}_e^* H'(\hat{\mathbf{z}})\| = \|\mathbf{u}_e^*[H'(\hat{\mathbf{z}}) - H'(\hat{\mathbf{z}}_e)]\| \le \|H'(\hat{\mathbf{z}}) - H'(\hat{\mathbf{z}}_e)\| = \mathrm{O}(\varepsilon),$$

and $H'(\hat{\mathbf{z}})$ has corank one approximately, so that $\|\mathbf{u}_n - \mathbf{u}_e\| = \mathrm{O}(\varepsilon)$. Moreover, using the same analysis above, we obtain that

$$|\mathbf{u}_e^* H''(\hat{\mathbf{z}}_e)(\hat{\mathbf{z}} - \hat{\mathbf{z}}_e)^2| = \mathrm{O}(\varepsilon^3).$$

Hence, we have $|\mathbf{u}_e^* H(\hat{\mathbf{z}})| = \mathrm{O}(\varepsilon^3)$. Noticing $\|H(\hat{\mathbf{z}})\| = \mathrm{O}(\varepsilon^2)$, we get

$$|\mathbf{u}_n^* H(\hat{\mathbf{z}})| \le |(\mathbf{u}_n - \mathbf{u}_e)^* H(\hat{\mathbf{z}})| + |\mathbf{u}_e^* H(\hat{\mathbf{z}})| = \mathrm{O}(\varepsilon^3). \tag{3.24}$$

Combining (3.22), (3.23) and (3.24), we have

$$|\mathbf{u}_n^* H'(\hat{\mathbf{z}})(\hat{\mathbf{z}}_e - \hat{\mathbf{z}})| = \mathrm{O}(\varepsilon^3), \tag{3.25}$$

which is equivalent to

$$|\sigma_n' \, \mathbf{v}_n^* (\hat{\mathbf{z}}_e - \hat{\mathbf{z}})| = \mathrm{O}(\varepsilon^3), \tag{3.26}$$

where $\mathbf{v}_n = [1, 0, \ldots, 0]^T$ is the right singular vector of $H'(\hat{\mathbf{z}})$ corresponding to σ_n'. Hence, $|\sigma_n'(\hat{z}_{1,e} - \hat{z}_1)| = \mathrm{O}(\varepsilon^3)$. Based on (3.16), we have

$$\sigma_n' = \mathrm{O}(\varepsilon^2). \tag{3.27}$$

Moreover, from (3.14), we have

$$\|L_1(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \left\| \frac{\partial H(\hat{\mathbf{z}})}{\partial z_1} \right\| = \mathrm{O}(\varepsilon^2). \tag{3.28}$$

□

It is amazing to notice that not only the first order differential condition computed according to Theorem 2.4 satisfies (3.28), but also all other differential conditions up to the order $\mu - 2 \ge 0$ satisfy similar conditions:

$$\|L_i(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2), \text{ for } i = 0, \ldots, \mu - 2. \tag{3.29}$$

3.2. An Augmented Polynomial System. To prove (3.29) inductively, we need to introduce an augmented polynomial system and prove the following theorem.

Theorem 3.8. *Let us assume that $H(\mathbf{z})$ is a polynomial system which has $\hat{\mathbf{z}}_e$ as an isolated exact singular solution with the multiplicity μ, the corank of $H'(\hat{\mathbf{z}}_e)$ is one. Let I be the ideal generated by polynomials in H and $\{L_0, L_1, \ldots, L_{\mu-1}\}$ be a closed basis of $\triangle_{\hat{\mathbf{z}}_e}(I)$, where $L_0 = D(0, \ldots, 0), L_1 = D(1, 0, \ldots, 0)$ and $L_k = P_k + a_{k,2} D(0, 1, \ldots, 0) + \cdots + a_{k,n} D(0, \ldots, 1)$ constructed according to Theorem 2.4. The augmented polynomial system*

$$G(\mathbf{z}, \lambda) := \begin{cases} H(\mathbf{z}), \\ H'(\mathbf{z}) \cdot \lambda, \\ \lambda_1 - 1, \end{cases} \tag{3.30}$$

where $\lambda = [\lambda_1, \ldots, \lambda_n]^T$ has an isolated singular solution $(\hat{\mathbf{z}}_e, \hat{\lambda}_e)$ with the multiplicity $\mu - 1$, where $\hat{\lambda}_e = [1, 0, \ldots, 0]^T$. If $\mu \ge 3$ then the Jacobian matrix $G'(\hat{\mathbf{z}}_e, \hat{\lambda}_e)$ has corank one and

$$\tilde{L}_1 = \frac{\partial}{\partial z_1} + 2a_{2,2} \frac{\partial}{\partial \lambda_2} + \cdots + 2a_{2,n} \frac{\partial}{\partial \lambda_n} \tag{3.31}$$

satisfies $\tilde{L}_1(G)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} = 0$. *Moreover, starting from* $\tilde{L}_0 = D(0,\ldots,0)$ *and* $\tilde{L}_1$, *for* $2 \le k \le \mu - 2$, *the* k*-th order differential condition of* G *at* $(\hat{\mathbf{z}}_e, \hat{\lambda}_e)$ *retaining the closedness has the following form:*

$$\tilde{L}_k = \tilde{P}_k + a_{k,2}\frac{\partial}{\partial z_2} + \cdots + a_{k,n}\frac{\partial}{\partial z_n} + (k+1)a_{k+1,2}\frac{\partial}{\partial \lambda_2} + \cdots + (k+1)a_{k+1,n}\frac{\partial}{\partial \lambda_n} \quad (3.32)$$

where

$$\tilde{P}_k = P_k + \Psi_{n+2}(Q_{k,n+2}) + \Psi_{n+3}(Q_{k,n+3})_{\alpha_{n+2}=0} + \cdots + \Psi_{2n}(Q_{k,2n})_{\alpha_{n+2}=\cdots=\alpha_{2n-1}=0} \quad (3.33)$$

and

$$Q_{k,n+j} = \Phi_{n+j}(\tilde{P}_k) = 2a_{2,j}\tilde{L}_{k-1} + \cdots + ka_{k,j}\tilde{L}_1, \; 2 \le j \le n. \quad (3.34)$$

Proof. The Jacobian matrix of $G(\mathbf{z},\lambda)$ at $(\hat{\mathbf{z}}_e, \hat{\lambda}_e)$ is

$$G'(\hat{\mathbf{z}}_e, \hat{\lambda}_e) = \begin{bmatrix} H'(\hat{\mathbf{z}}_e) & 0 \\ H''(\hat{\mathbf{z}}_e)\cdot\hat{\lambda}_e & H'(\hat{\mathbf{z}}_e) \\ 0 & \hat{\lambda}_e^T \end{bmatrix},$$

where $H''(\hat{\mathbf{z}}_e)\cdot\hat{\lambda}_e = \left[\frac{\partial^2 H(\hat{\mathbf{z}}_e)}{\partial z_1^2}, \ldots, \frac{\partial^2 H(\hat{\mathbf{z}}_e)}{\partial z_1 \partial z_n}\right]$. Since the corank of $H'(\hat{\mathbf{z}}_e)$ is one and $L_1 = D(1,0,\ldots,0) \in \triangle^{(1)}_{\hat{\mathbf{z}}_e}(I)$, the first column of $H'(\hat{\mathbf{z}}_e)$ is a zero vector and the remaining columns of $H'(\hat{\mathbf{z}}_e)$ are linearly independent. Moreover, since $\hat{\lambda}_e^T = [1,0,\ldots,0]$, the last $2n-1$ columns of $G'(\hat{\mathbf{z}}_e,\hat{\lambda}_e)$ are linearly independent and its corank is less than one.

If $\mu \ge 3$, the second order differential condition of H at $\hat{\mathbf{z}}_e$ has the form $L_2 = D(2,0,\ldots,0) + a_{2,2}D(0,1,0,\ldots,0) + \cdots + a_{2,n}D(0,\ldots,0,1)$. From $L_2(H)_{\mathbf{z}=\hat{\mathbf{z}}_e} = 0$, we have

$$\frac{1}{2}\frac{\partial^2 H(\hat{\mathbf{z}}_e)}{\partial z_1^2} + a_{2,2}\frac{\partial H(\hat{\mathbf{z}}_e)}{\partial z_2} \cdots + a_{2,n}\frac{\partial H(\hat{\mathbf{z}}_e)}{\partial z_n} = 0.$$

The vector $\mathbf{v} = [1,0,\ldots,0,2a_{2,2},\ldots,2a_{2,n}]^T$ is a null vector of $G'(\hat{\mathbf{z}}_e,\hat{\lambda}_e)$. Therefore, the Jacobian matrix $G'(\hat{\mathbf{z}}_e,\hat{\lambda}_e)$ has corank one and the first order differential operator $\tilde{L}_1$ in (3.31) satisfies

$$\tilde{L}_1(H'(\mathbf{z})\cdot\lambda)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} = 2L_2(H)_{\mathbf{z}=\hat{\mathbf{z}}_e} = 0. \quad (3.35)$$

Hence, we have

$$\tilde{L}_1(G)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} = 0. \quad (3.36)$$

Using similar arguments in [20] for proving Theorem 2.4, we can show that the differential operators $\tilde{L}_k$ defined by formulas (3.32), (3.33) and (3.34) retain the closedness. It should also be noticed that $\tilde{L}_k$ always contains the differential monomial $D(k,0,\ldots,0)$ and there are no differential monomials $D(i,0,\ldots,0)$ for $i < k$ contained in $\tilde{L}_k$. Otherwise, we can reduce them by $\tilde{L}_i$. Moreover, $\frac{\partial}{\partial\lambda_1}$ is not contained in any $\tilde{L}_k$, otherwise, $\tilde{L}_k(\lambda_1)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} \neq 0$. Hence, due to the closedness, there are no differential operators $D(\alpha_1,\ldots,\alpha_n,\alpha_{n+1},\ldots,\alpha_{2n})$ with $\alpha_{n+1} > 0$ contained in any $\tilde{L}_k$.

Now let us show that the constructed differential operators $\tilde{L}_k$ satisfy

$$\tilde{L}_k(G)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} = 0, \text{ for } 1 \le k \le \mu - 2. \tag{3.37}$$

From (3.36), we can see that (3.37) is true for $k = 1$. Moreover, it is easy to check that

$$\tilde{L}_k(H)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} = \left(P_k + a_{k,2}\frac{\partial}{\partial z_2} + \cdots + a_{k,n}\frac{\partial}{\partial z_n}\right)(H)_{\mathbf{z}=\hat{\mathbf{z}}_e} = 0, \tag{3.38}$$

and

$$\tilde{L}_k(\lambda_1)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} = 0. \tag{3.39}$$

Based on formulas (3.32), (3.33) and (3.34), we have

$$\tilde{L}_k(H'(\mathbf{z})\cdot\lambda)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} =$$
$$\left(L_k\frac{\partial}{\partial z_1} + \sum_{j=2}^{n}(2\,a_{2,j}L_{k-1} + \cdots + k\,a_{k,j}L_1 + (k+1)\,a_{k+1,j})\frac{\partial}{\partial z_j}\right)(H)_{\mathbf{z}=\hat{\mathbf{z}}_e}.$$

Let us set

$$Q_{k+1} = L_k\frac{\partial}{\partial z_1} + \sum_{j=2}^{n}(2\,a_{2,j}L_{k-1} + \cdots + k\,a_{k,j}L_1)\frac{\partial}{\partial z_j}. \tag{3.40}$$

We prove inductively that

$$Q_{k+1} = (k+1)P_{k+1}. \tag{3.41}$$

Hence, we have

$$\begin{aligned} &\tilde{L}_k(H'(\mathbf{z})\cdot\lambda)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}}_e,\hat{\lambda}_e)} \\ &= \left((k+1)P_{k+1} + (k+1)a_{k+1,2}\tfrac{\partial}{\partial z_2} + \cdots + (k+1)a_{k+1,n}\tfrac{\partial}{\partial z_n}\right)(H)_{\mathbf{z}=\hat{\mathbf{z}}_e} \\ &= (k+1)L_{k+1}(H)_{\mathbf{z}=\hat{\mathbf{z}}_e} = 0. \end{aligned} \tag{3.42}$$

From (3.38), (3.39) and (3.42), we derive that (3.37) is true for $1 \le k \le \mu - 2$. ∎

Corollary 3.9. *Suppose $F(\mathbf{x})$ is a polynomial system which has $\hat{\mathbf{x}}_e$ as an isolated exact singular solution with the multiplicity μ and the corank of $F'(\hat{\mathbf{x}}_e)$ is one. Let $\mathbf{r}_1$ be the null vector of $F'(\hat{\mathbf{x}}_e)$ and $\|\mathbf{r}_1\| = 1$. For any random vector $\mathbf{h} \in \mathbb{C}^n$ satisfying $\mathbf{h}^*\mathbf{r}_1 \neq 0$, the augmented polynomial system*

$$J(\mathbf{x},\nu) := \begin{cases} F(\mathbf{x}), \\ F'(\mathbf{x})\cdot\nu, \\ \mathbf{h}^*\nu - 1, \end{cases} \tag{3.43}$$

has $(\hat{\mathbf{x}}_e, \frac{\mathbf{r}_1}{\mathbf{h}^*\mathbf{r}_1})$ as an isolated singular solution with the multiplicity $\mu - 1$.

Proof. Let $\{\mathbf{r}_1,\ldots,\mathbf{r}_n\}$ be a normal orthogonal basis of $\mathbb{C}^n$, then $\mathbf{h} = (\mathbf{h}^*\mathbf{r}_1)\mathbf{r}_1 + \cdots + (\mathbf{h}^*\mathbf{r}_n)\mathbf{r}_n$. If $\mathbf{h}^*\mathbf{r}_1 \neq 0$, performing the linear transformation

$$\mathbf{x} = R\,\mathbf{z}, \quad \nu = R\,\lambda,$$

where $R = \left[\frac{\mathbf{r}_1}{\mathbf{h}^*\mathbf{r}_1}, \mathbf{r}_2 - \frac{\mathbf{h}^*\mathbf{r}_2}{\mathbf{h}^*\mathbf{r}_1}\mathbf{r}_1, \ldots, \mathbf{r}_n - \frac{\mathbf{h}^*\mathbf{r}_n}{\mathbf{h}^*\mathbf{r}_1}\mathbf{r}_1\right]$ is a regular matrix, we obtain the augmented polynomial system

$$G(\mathbf{z}, \lambda) := \begin{cases} H(\mathbf{z}), \\ H'(\mathbf{z}) \cdot \lambda, \\ \lambda_1 - 1, \end{cases}$$

where

$$H(\mathbf{z}) = F(R\,\mathbf{z}), \quad H'(\mathbf{z}) \cdot \lambda = F'(\mathbf{x}) \cdot R \cdot R^{-1}\nu, \quad \lambda_1 - 1 = \mathbf{h}^*\nu - 1.$$

According to Theorem 3.8, we know that $(\hat{\mathbf{z}}_e, \hat{\lambda}_e)$ is an isolated singular solution of G with the multiplicity $\mu - 1$, where $\hat{\mathbf{z}}_e = R^{-1}\hat{\mathbf{x}}_e$ and $\hat{\lambda}_e = [1, 0, \ldots, 0]^T$. Hence, by Theorem 3.3 and Remark 3.5, $\left(\hat{\mathbf{x}}_e, \frac{\mathbf{r}_1}{\mathbf{h}^*\mathbf{r}_1}\right)$ is an isolated singular solution of $J(\mathbf{x}, \nu)$ with the multiplicity $\mu - 1$. ▯

Remark 3.10. *It is well known that the augmented polynomial system $J(\mathbf{x}, \nu)$ defined in (3.43) has an isolated singular solution $\left(\hat{\mathbf{x}}_e, \frac{\mathbf{r}_1}{\mathbf{h}^*\mathbf{r}_1}\right)$ with the multiplicity less than μ, see [18, 5]. Here, we proved the conjecture in [5] that the multiplicity of the singular solution of the augmented polynomial system (3.43) drops by one exactly in the breadth one case.*

Remark 3.11. *For the system $H(\mathbf{z})$ and its approximate singular solution $\hat{\mathbf{z}}$ defined in (3.9) and (3.11), the augmented polynomial system defined in (3.30) has $(\hat{\mathbf{z}}, \hat{\lambda})$ ($\hat{\lambda} = [1, 0, \ldots, 0]^T$) as an approximate solution. Suppose $\{L_0, \ldots, L_{\mu-1}\}$ is a closed basis of the approximate local dual space of the system H at $\hat{\mathbf{z}}$ constructed according to Theorem 2.4, from $L_0 = D(0, \ldots, 0)$ and $L_1 = D(1, 0, \ldots, 0)$, then $\{\tilde{L}_0, \tilde{L}_1, \ldots, \tilde{L}_{\mu-2}\}$ constructed according to Theorem 3.8 is a closed basis of the approximate local dual space of the system G at $(\hat{\mathbf{z}}, \hat{\lambda})$, satisfying*

$$\begin{cases} \|\tilde{L}_k(H)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}},\hat{\lambda})}\| = \|L_k(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon), \\ \|\tilde{L}_k(H'\lambda)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}},\hat{\lambda})}\| = \|(k+1)L_{k+1}(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon), \\ \|\tilde{L}_k(\lambda_1)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}},\hat{\lambda})}\| = 0, \end{cases}$$

for $1 \le k \le \mu - 2$.

Theorem 3.12. *Let $F(\mathbf{x})$ be a polynomial system which has $\hat{\mathbf{x}}_e$ as an isolated exact singular solution with the multiplicity μ and the breadth one. Suppose $\hat{\mathbf{x}}$ is an approximate solution of F which satisfies*

$$\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \Theta(\varepsilon) \text{ and } \|F(\hat{\mathbf{x}})\| = \mathrm{O}(\varepsilon^2), \tag{3.44}$$

for a small positive number ε. Let $\sigma_1, \ldots, \sigma_n$ be the singular values of $F'(\hat{\mathbf{x}})$ satisfying $\sigma_i = \Theta(1)$, $1 \le i \le n-1$ and $\sigma_n = \mathrm{O}(\varepsilon)$. Suppose $\mathbf{r}_1$ is the right singular vector corresponding to σ_n. We form a unitary matrix $R = [\mathbf{r}_1, \ldots, \mathbf{r}_n]$ and set $H(\mathbf{z}) = F(R\,\mathbf{z})$. Suppose $\{L_0, \ldots, L_{\mu-1}\}$ is a closed basis of the approximate local dual space of the system H at $\hat{\mathbf{z}} = R^{-1}\,\hat{\mathbf{x}}$ constructed according to Theorem 2.4 from $L_0 = D(0, \ldots, 0)$ and $L_1 = D(1, 0, \ldots, 0)$, then

$$\|L_i(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2), \text{ for } i = 0, \ldots, \mu - 2.$$

Remark 3.13. *Under Assumption 1, according to Theorem 3.2, we can always perform the regularized Newton iteration to obtain an approximate singular solution* $\hat{\mathbf{x}}$ *satisfying (3.44). Moreover, it should also be noticed that all discussions in Section 3.1 after Theorem 3.2 are valid if we start with an approximate singular solution satisfying (3.44).*

Proof. According to (3.13) and Theorem 3.7, we know that Theorem 3.12 is true for $\mu = 2$ and $\mu = 3$.

Now let us assume that Theorem 3.12 is true for $\mu = k$ and $k \geq 3$. For $\mu = k+1$, we form the augmented polynomial system $G(\mathbf{z}, \lambda) = \{H(\mathbf{z}), H'(\mathbf{z}) \cdot \lambda, \lambda_1 - 1\}$.

According to Theorem 3.3, the root $\hat{\mathbf{z}}_e$ defined in (3.10) is an exact singular solution of $H(\mathbf{z})$ with the multiplicity μ and the corank of $H'(\hat{\mathbf{z}}_e)$ is one. Let $\mathbf{v}$ be the null vector of $H'(\hat{\mathbf{z}}_e)$ and $\|\mathbf{v}\| = 1$. Since

$$\|H'(\hat{\mathbf{z}})\mathbf{v}\| = \|[H'(\hat{\mathbf{z}}) - H'(\hat{\mathbf{z}}_e)]\mathbf{v}\| = \mathrm{O}(\varepsilon), \tag{3.45}$$

$\left[\frac{\partial H(\hat{\mathbf{z}})}{\partial z_2}, \ldots, \frac{\partial H(\hat{\mathbf{z}})}{\partial z_n}\right]$ is of full column rank, combining with (3.14), we derive that

$$v_1 = \Theta(1), \text{ and } v_i = \mathrm{O}(\varepsilon), \text{ for } 2 \leq i \leq n. \tag{3.46}$$

Set $\mathbf{h} = [1, 0, \ldots, 0]^T$, we have $\mathbf{h}^*\mathbf{v} = v_1 = \Theta(1) \neq 0$.

According to Corollary 3.9, the augmented polynomial system $G(\mathbf{z}, \lambda)$ has $\left(\hat{\mathbf{z}}_e, \frac{\mathbf{v}}{\mathbf{h}^*\mathbf{v}}\right) = (\hat{\mathbf{z}}_e, \hat{\lambda}_e)$, where

$$\hat{\lambda}_e = \left[1, \frac{v_2}{v_1}, \ldots, \frac{v_n}{v_1}\right]^T,$$

as an isolated singular solution with the multiplicity $\mu - 1$, which is equal to k. According to Remark 3.11, $(\hat{\mathbf{z}}, \hat{\lambda})$ ($\hat{\lambda} = [1, 0, \ldots, 0]^T$) is an approximate solution of $G(\mathbf{z}, \lambda)$. Moreover, by (3.15) and (3.46), we have

$$\|(\hat{\mathbf{z}}, \hat{\lambda}) - (\hat{\mathbf{z}}_e, \hat{\lambda}_e)\| = \sqrt{\|\hat{\mathbf{z}} - \hat{\mathbf{z}}_e\|^2 + \|\hat{\lambda} - \hat{\lambda}_e\|^2} = \Theta(\varepsilon).$$

Furthermore, from (3.13) and (3.28), we have

$$\|G(\hat{\mathbf{z}}, \hat{\lambda})\| = \sqrt{\|H(\hat{\mathbf{z}})\|^2 + \left\|\frac{\partial H(\hat{\mathbf{z}})}{\partial z_1}\right\|^2} = \mathrm{O}(\varepsilon^2).$$

We have assumed that Theorem 3.12 is true when the multiplicity is equal to k. Therefore, for the augmented polynomial system $G(\mathbf{z}, \lambda)$, we can form a unitary matrix $\bar{R}$ with $\mathbf{r}_1 = \frac{1}{a}[1, 0, \ldots, 0, 2a_{2,2}, \ldots, 2a_{2,n}]^T$ as its first column, where $a = \sqrt{1 + 4(a_{2,2}^2 + \cdots + a_{2,n}^2)}$, then generating a new system $J(\mathbf{w}) = G(\bar{R}\,\mathbf{w})$ which has an approximate singular solution $\hat{\mathbf{w}}$ with the multiplicity k. By the inductive assumption, we have

$$\|\bar{L}_i(J)_{\mathbf{w}=\hat{\mathbf{w}}}\| = \mathrm{O}(\varepsilon^2), \text{ for } 0 \leq i \leq k-2,$$

where $\bar{L}_i$ is the i-th differential condition of J at $\hat{\mathbf{w}}$ constructed by Theorem 2.4 from $\bar{L}_0 = D(0, \ldots, 0)$ and $\bar{L}_1 = D(1, 0, \ldots, 0)$. According to Theorem 3.3,

$$\bar{L}_i(J)_{\mathbf{w}=\hat{\mathbf{w}}} = \Gamma_{\bar{R}}(\bar{L}_i)(G)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}},\hat{\lambda})}.$$

Since $\{\tilde{L}_0, \tilde{L}_1, \ldots, \tilde{L}_{k-1}\}$ and $\{\Gamma_{\bar{R}}(\bar{L}_0), \Gamma_{\bar{R}}(\bar{L}_1), \ldots, \Gamma_{\bar{R}}(\bar{L}_{k-1})\}$ are both closed basis of the approximate local dual space of the system G at $(\hat{\mathbf{z}}, \hat{\lambda})$, and

$$\Gamma_{\bar{R}}(\bar{L}_0) = \tilde{L}_0, \text{ and } \Gamma_{\bar{R}}(\bar{L}_1) = \frac{1}{a}\tilde{L}_1,$$

we derive that $\Gamma_{\bar{R}}(\bar{L}_i)$ is a linear combination of $\{\tilde{L}_0, \tilde{L}_1, \ldots, \tilde{L}_i\}$. Hence, we have $\|\tilde{L}_i(G)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}},\hat{\lambda})}\| = \mathrm{O}(\varepsilon^2)$, and

$$\|L_{i+1}(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \left\| \frac{1}{i+1}\tilde{L}_i(H'\lambda)_{(\mathbf{z},\lambda)=(\hat{\mathbf{z}},\hat{\lambda})} \right\| = \mathrm{O}(\varepsilon^2).$$

Therefore, Theorem 3.12 is true for $\mu = k+1$. ☐

3.3. An Algorithm for Refining Approximate Singular Solutions.

ALGORITHM 1. MultipleRootRefinerBreadthOne

Input: *An approximate solution $\hat{\mathbf{x}}$ of a polynomial system F which is close to an isolated exact singular solution of F with the multiplicity μ in the breadth one case, and a tolerance τ.*

Output: *Refined solution $\hat{\mathbf{x}}$.*

1. Regularized Newton Iteration: *Solve the regularized least squares problem*

$$(A^*A + \sigma_n I_n)\hat{\mathbf{y}} = A^*\mathbf{b},$$

where $\mathbf{b} = -F(\hat{\mathbf{x}})$, A^* is the Hermitian (conjugate) transpose of $A = F'(\hat{\mathbf{x}})$, I_n is the $n \times n$ identity matrix and σ_n is the smallest singular value of A.

2. *Compute the null vector $\mathbf{r}_1$ of $F'(\hat{\mathbf{x}} + \hat{\mathbf{y}})$ with respect to τ, form a unitary matrix R with $\mathbf{r}_1$ as its first column and perform the linear transformation*

$$H(\mathbf{z}) := F(R\,\mathbf{z}),$$

and set $\hat{\mathbf{z}} := R^{-1}(\hat{\mathbf{x}} + \hat{\mathbf{y}})$.

3. *Construct a closed basis of the approximate local dual space of $I = (h_1, \ldots, h_n)$ at $\hat{\mathbf{z}}$ with respect to τ:*

$$\triangle_{\hat{\mathbf{z}}}^{(\mu-1)}(I) := \mathrm{Span}(L_0, L_1, \ldots, L_{\mu-1})$$

by Algorithm MultiplicityStructureBreadthOneNumeric *in [20].*

4. *Solve the linear system*

$$\left[P_\mu(H)_{\mathbf{z}=\hat{\mathbf{z}}}, \frac{\partial H(\hat{\mathbf{z}})}{\partial z_2}, \ldots, \frac{\partial H(\hat{\mathbf{z}})}{\partial z_n} \right] \mathbf{v} = -L_{\mu-1}(H)_{\mathbf{z}=\hat{\mathbf{z}}}, \qquad (3.47)$$

where $\mathbf{v} = [v_1, \cdots, v_n]^T$ and P_μ is the differential operator of order μ computed by formulas in Theorem 2.4. Set $\delta := \frac{v_1}{\mu}$.

5. *Return*

$$\hat{\mathbf{x}} := \hat{\mathbf{x}} + \hat{\mathbf{y}} + \delta\,\mathbf{r}_1.$$

REMARK 3.14. *The size of matrices involved in the algorithm* MultipleRootRefinerBreadthOne *is bounded by $n \times n$, whereas the size of matrices used in the deflation method is bounded by $(\mu\, n) \times (\mu\, n)$ [5, 18].*

REMARK 3.15. *In fact, in order to keep the sparse structure of the original polynomial system, we should avoid performing the linear transformation. Moreover, it is expensive to compute and store all differential conditions. Since we only need their evaluations to solve (3.47), it's possible to compute and store only the necessary evaluations of these differential conditions. We will discuss these issues in forthcoming papers.*

3.4. Quadratic Convergence of the Algorithm.

THEOREM 3.16. *Under Assumptions 1, the refined singular solution* $\hat{\mathbf{x}}$ *returned by Algorithm* MultipleRootRefinerBreadthOne *satisfies*

$$\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \mathrm{O}(\varepsilon^2). \tag{3.48}$$

Proof. According to Theorem 3.12, we have $L_i(H)_{\mathbf{z}=\hat{\mathbf{z}}} = \mathrm{O}(\varepsilon^2)$, for $0 \le i \le \mu-2$. Since

$$\Phi_k(L_i) \in \mathrm{Span}(L_0, \ldots, L_{\mu-2}), \text{ for } 1 \le k \le n,$$

we have

$$\|L_i((z_k - \hat{z}_k)H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \|\Phi_k(L_i)(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2), \text{ for } 1 \le k \le n, 0 \le i \le \mu-1. \tag{3.49}$$

The matrix in (3.47) is of full rank. We solve the linear system (3.47) to obtain the vector $\mathbf{v} = [v_1, \ldots, v_n]^T$ such that $L_\mu(H)_{\mathbf{z}=\hat{\mathbf{z}}} = 0$ for

$$L_\mu := L_{\mu-1} + v_1 \cdot P_\mu + v_2 \cdot \frac{\partial}{\partial z_2} + \cdots + v_n \cdot \frac{\partial}{\partial z_n}. \tag{3.50}$$

It should be noticed that the vector $\mathbf{v}$ satisfies $\|\mathbf{v}\| = \mathrm{O}(\varepsilon)$ since $\|L_{\mu-1}(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon)$. Moreover,

$$\Phi_k(L_\mu) \in \mathrm{Span}(L_0, \ldots, L_{\mu-2}), \text{ for } 2 \le k \le n,$$

we have

$$\|L_\mu((z_k - \hat{z}_k)H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \|\Phi_k(L_\mu)(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2), \text{ for } 2 \le k \le n. \tag{3.51}$$

For $k = 1$, since $\|\Phi_1(v_1 P_\mu)(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \|v_1 L_{\mu-1}(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2)$, we have

$$\|L_\mu((z_1 - \hat{z}_1)H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \|\Phi_1(L_\mu)(H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2). \tag{3.52}$$

From (3.49) and (3.52), for $i = 0, 1, \ldots, \mu-2, \mu$, we have

$$\|L_i(p \cdot H)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2), \ \forall p \in \{(z_1 - \hat{z}_1)^{\alpha_1} \cdots (z_n - \hat{z}_n)^{\alpha_n}, \alpha_1 \ge 0, \ldots, \alpha_n \ge 0\}.$$

Especially, we have

$$\|M_{\mu+1} \cdot L_i(\mathbf{v}(\mathbf{z})_\mu)_{\mathbf{z}=\hat{\mathbf{z}}}\| = \mathrm{O}(\varepsilon^2),$$

where $M_{\mu+1}$ is the coefficient matrix of the Taylor expansion of the system H and all its prolongations up to the degree μ at $\hat{\mathbf{z}}$, and

$$\mathbf{v}(\mathbf{z})_\mu = \left[(z_1 - \hat{z}_1)^\mu, (z_1 - \hat{z}_1)^{\mu-1}(z_2 - \hat{z}_2), \ldots, z_1 - \hat{z}_1, \ldots, z_n - \hat{z}_n, 1\right]^T.$$

It is important to notice that, based on the closedness conditions, we obtain the null space of $M_{\mu+1}$ with matrices of size $n \times n$ instead of generating the big matrix $M_{\mu+1}$. Similarly to the analysis in [39, Remark 18], the trace of the multiplication matrix $\widetilde{M}_{z_1}$ formed from approximate null vectors $L_i(\mathbf{v}(\mathbf{z})_\mu)_{\mathbf{z}=\hat{\mathbf{z}}}$ has the following property

$$\frac{1}{\mu}\mathrm{Tr}(\widetilde{M}_{z_1}) = \frac{1}{\mu}\mathrm{Tr}(M_{z_1}) + \mathrm{O}(\varepsilon^2) = -\hat{z}_{1,\epsilon} + \mathrm{O}(\varepsilon^2). \tag{3.53}$$

It is interesting to notice that, by using the approximate basis $\{L_0, \ldots, L_{\mu-2}, L_\mu\}$ and the normal set $\left\{1, \frac{\partial}{\partial z_1}, \ldots, \frac{\partial^{\mu-1}}{\partial z_1^{\mu-1}}\right\}$, we can form the multiplication matrix

$$\widetilde{M}_{z_1} \cdot \begin{bmatrix} l_0 & 0 & \cdots & 0 \\ 0 & l_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & l_{\mu-1} \end{bmatrix} = \begin{bmatrix} 0 & l_1 & 0 & \cdots & 0 \\ 0 & 0 & l_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & l_{\mu-1} \\ 0 & \cdots & \cdots & 0 & v_1 \cdot l_{\mu-1} \end{bmatrix},$$

where l_i is the coefficient of $\frac{\partial^i}{\partial z_1^i}$ in L_i. Hence, the trace of $\widetilde{M}_{z_1}$ is v_1. Therefore, there is no need to form the multiplication matrix! According to (3.53), we have

$$\frac{v_1}{\mu} + \hat{z}_{1,\epsilon} = \mathrm{O}(\varepsilon^2). \tag{3.54}$$

Since the last $n-1$ elements of $\hat{\mathbf{z}}$ have already been refined quadratically, by updating $\hat{z}_1 := \hat{z}_1 + \delta$ for $\delta := \frac{v_1}{\mu}$, we have

$$\|\hat{\mathbf{x}} - \hat{\mathbf{x}}_e\| = \|R\,(\hat{\mathbf{z}} - \hat{\mathbf{z}}_e)\| = \mathrm{O}(\varepsilon^2).$$

□

REMARK 3.17. *The algorithm* MultipleRootRefinerBreadthOne *also works well for some overdetermined polynomial systems, i.e., the number of polynomials is bigger than the number of variables, see the last example Menzel1 in Table 4.*

4. Experiments.

The following experiments are done in Maple 15 under Windows 7 for Digits := 14. Let t and s be the number of polynomials and variables respectively, μ be the multiplicity. The last column lists the increase in the number of correct digits from the initial guess to the final approximation, and gives the number of iterations necessary to get the desired precision. Systems DZ3 and Dayton2 are quoted from [3, 5]. Menzel1 and SY5 are cited from [23] and [33] respectively. Example 1 is kindly provided by the reviewer.

EXAMPLE 1. *The system*

$$\{32y - 24z + 8u - v - 16, -x^2 + y, -y^2 + z, -yz + u, -z^2 + v\}$$

has a 4*-fold solution:*

$$\{u = 8, v = 16, x = \sqrt{2}, y = 2, z = 4.\}$$

The last three examples GLSY1, GLSY2, GLSY3 are corresponding to the Example 2 in [10] for $N = 5, 10, 20$ respectively. Other examples are cited from the PHCpack demos by Jan Verschelde.

REMARK 4.1. *It should be noticed that the tolerance has to be chosen carefully in order to obtain the true multiplicity of the singular root. You may increase or decrease the tolerance to achieve the quadratical convergence. For the example DZ3, we choose the tolerance to be* 10^{-1}*. For Example 1, the tolerance is set to be* 10^{-8} *to achieve the quadratical convergence. For other examples listed in Table 4, we choose the tolerance to be* 10^{-2}*. Since the tolerance can range from* 0.1 *to* 10^{-10}*, it is not easy for a user to choose a correct tolerance to obtain the correct multiplicity and achieve the quadratic*

System	Zero	t	s	μ	# Digits
Ojika1	$(1,2)$	2	2	3	$2 \to 3 \to 8 \to 16$
Ojika2	$(1,0,0)$	3	3	2	$2 \to 3 \to 7 \to 14$
Ojika3	$(-2.5,2.5,1)$	3	3	2	$2 \to 4 \to 7 \to 16$
Ojika4	$(0,0,10)$	3	3	3	$3 \to 4 \to 8 \to 17$
Decker2	$(0,0)$	2	2	4	$3 \to 8 \to 25$
DZ3	$(\frac{2\sqrt{7}}{5}+\frac{\sqrt{5}}{5}, -\frac{\sqrt{7}}{5}+\frac{2\sqrt{5}}{5})$	2	2	5	$3 \to 8 \to 12$
Dayton2	$(0,0,0)$	3	3	5	$3 \to 5 \to 10 \to 20$
SY5	$(1,1)$	2	2	2	$2 \to 4 \to 9 \to 19$
Menzel1	$(1,1)$	3	2	2	$2 \to 4 \to 9 \to 20$
Example 1	$(8,16,\sqrt{2},2,4)$	4	5	5	$3 \to 8 \to 13$
GLSY1	$(0,0,0)$	3	3	4	$3 \to 5 \to 11 \to 16$
GLSY2	$(0,0,0)$	3	3	4	$3 \to 6 \to 12 \to 15$
GLSY3	$(0,0,0)$	3	3	4	$3 \to 5 \to 11 \to 16$

Algorithm Performance

convergence. We are going to improve our algorithm in future so that it will not be very sensitive to a user chosen tolerance.

The last three examples from [10] show that our algorithm works also well for slightly perturbed systems. The output of the algorithm converges quadratically to the origin. Although it might not be meaningful to compute a solution near a cluster to 15 *digits, it is still interesting to see that our algorithm actually converges quadratically to the true singular solution* $(0,0,0)$ *of the nearby singular system for* $N=+\infty$.

The Maple code of the algorithm and test results are available `http://www.mmrc.iss.ac.cn/~lzhi/Research/hybrid/rootrefinerbreadthone`.

5. Conclusion. It is a challenge problem to solve the polynomial systems with singular solutions. Various symbolic-numeric methods have been proposed for refining an approximate singular solution to high accuracy [2, 4, 5, 9, 10, 17, 18, 27, 38, 39]. The breadth one case root refinement has been studied in [4, 5, 10, 13]. In this paper, we show how to apply strategies in [20] to reduce the size of matrices that appear in [5, 38] to obtain a more efficient algorithm for refining an approximately known multiple root for this special case. We have proved the quadratic convergence of the new algorithm when the approximate solution is close to the isolated exact singular solution. We also notice that when the singular solution $\hat{\mathbf{x}}_e$ is not well separated from other solutions of F, it is difficult to ensure that the approximate solution $\hat{\mathbf{x}}$ will converge to $\hat{\mathbf{x}}_e$. In [32], they described an algorithm for computing verified error bounds for double roots of polynomial systems. We will explore ways of computing the certified bound for ε to guarantee the convergence of our algorithm. It is also interesting to see whether the approach in the paper can be generalized to refine singular solutions when the Jacobian matrix is not of corank one.

Acknowledgments. The authors wish to thank the reviewers for valuable suggestions and comments that have improved the presentation of the paper.

REFERENCES

[1] Xiaojun Chen, Zuhair Nashed, and Liqun Qi, *Convergence of Newton's method for singular smooth and nonsmooth equations using adaptive outer inverses*, SIAM J. on Optimization, 7 (1997), pp. 445–462.

[2] R. Corless, P. Gianni, and B. Trager, *A reordered Schur factorization method for zero-dimensional polynomial systems with multiple roots*, in Proc. 1997 Internat. Symp. Symbolic Algebraic Comput. ISSAC'97, Küchlin, ed., New York, 1997, ACM Press, pp. 133–140.

[3] B. Dayton, *Numerical local rings and local solutions of nonlinear systems*, in SNC'07 Proc. 2007 Internat. Workshop on Symbolic-Numeric Comput., Jan Verschelde and Stephen M. Watt, eds., New York, N. Y., 2007, ACM Press, pp. 79–86.

[4] B. Dayton, T. Li, and Z. Zeng, *Multiple zeros of nonlinear systems*, Mathematics of Computation, 80 (2011), pp. 2143–2168.

[5] B. Dayton and Z. Zeng, *Computing the multiplicity structure in solving polynomial systems*, in ISSAC'05 Proc. 2005 Internat. Symp. Symbolic Algebraic Comput., Manuel Kauers, ed., New York, N. Y., 2005, ACM Press, pp. 116–123.

[6] D. W. Decker and C. T. Kelley, *Newton's method at singular points. i*, SIAM Journal on Numerical Analysis, 17 (1980), pp. 66–70.

[7] ———, *Newton's method at singular points. ii*, SIAM Journal on Numerical Analysis, 17 (1980), pp. 465–471.

[8] ———, *Convergence acceleration for Newton's method at singular points*, SIAM Journal on Numerical Analysis, 19 (1982), pp. 219–229.

[9] 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), pp. 257–311.

[10] ———, *On location and approximation of clusters of zeros: Case of embedding dimension one*, Found. Comput. Math., 7 (2007), pp. 1–58.

[11] G. Golub and C. Van Loan, *Matrix computations*, Johns Hopkins University Press, 3nd edition, 1996.

[12] Andreas Griewank, *Analysis and modification of Newton's method at singularities*, thesis, Australian National University, 1980.

[13] A. Griewank, *On solving nonlinear equations with simple singularities or nearly singular solutions*, SIAM Review, 27 (1985), pp. 537–563.

[14] Andreas Griewank and M. R. Osborne, *Newton's method for singular problems when the dimension of the null space is* > 1, SIAM Journal on Numerical Analysis, 18 (1981), pp. 145–149.

[15] Andreas Griewank and M. R. Osborne, Analysis of *Newton's method for irregular singularities*, SIAM Journal on Numerical Analysis, 20 (1983), pp. 747–743.

[16] L.V. Kantorovich and G.P. Akilov, *Functional analysis in normed spaces*, Pergamon, New York, 1964.

[17] G. Lecerf, *Quadratic Newton iteration for systems with multiplicity*, Foundations of Computational Mathematics, 2 (2002), pp. 247–293.

[18] A. Leykin, J. Verschelde, and A. Zhao, *Newton's method with deflation for isolated singularities of polynomial systems*, Theoretical Computer Science, 359 (2006), pp. 111–122.

[19] N. Li, *An improved method for evaluating Max Noether conditions: case of breadth one*, in Marc Moreno Maza, editor, SNC'11 Proc. 2011 Internat. Workshop on Symbolic-Numeric Comput., pages 102–103, New York, N. Y., 2011. ACM Press.

[20] N. Li and L. Zhi, *Compute the multiplicity structure of an isolated singular solution: case of breadth one*, MM Research Preprints, 28 (2009), pp. 93–103. Accepted for publication in J. Symbolic Comput.
`http://www.mmrc.iss.ac.cn/~lzhi/Publications/jsc_LiZhi2010.pdf`.

[21] M. Marinari, T. Mora, and H. Möller, *Gröbner bases of ideals defined by functionals with an application to ideals of projective points*, Applicable Algebra in Engineering, Communication and Computing, 4 (1993), pp. 103–145.

[22] ———, *On multiplicities in polynomial system solving*, Trans. Amer. Math. Soc., 348 (1996), pp. 3283–3321.

[23] R. Menzel, *On solving nonlinear least-squares problems in case of rankdeficient Jacobians*, Computing, 34 (1985), pp. 63–72.

[24] H. Möller and H. Stetter, *Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems*, Numer. Math., 70 (1995), pp. 311–329.

[25] H. Möller and R. Tenberg, *Multivariate polynomial system solving using intersections of eigenspaces*, J. Symbolic Comput., 30 (2001), pp. 1–19.

[26] B. Mourrain, *Isolated points, duality and residues*, J. of Pure and Applied Algebra, 117 & 118 (1996), pp. 469–493.

[27] T. Ojika, *Modified deflation algorithm for the solution of singular problems*, J. Math. Anal.

Appl., 123 (1987), pp. 199–221.
[28] 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), pp. 463–479.
[29] L.B. Rall, *Convergence of the Newton process to multiple solutions*, Numer. Math., 9 (1966), pp. 23–37.
[30] G. W. Reddien, *On Newton's method for singular problems*, SIAM Journal on Numerical Analysis, 15 (1978), pp. 993–996.
[31] ———, *Newton's method and high order singularities*, Comput. Math. Appl, 5 (1980), pp. 79–86.
[32] S.M. Rump and S. Graillat, *Verified error bounds for multiple roots of systems of nonlinear equations*, Numerical Algorithms, 54 (2009), pp. 359–377.
[33] Yun-Qiu Shen and Tjalling J. Ypma, *Newton's method for singular nonlinear equations using approximate left and right nullspaces of the Jacobian*, Applied Numerical Mathematics, 54 (2005), pp. 256 – 265. 6th IMACS.
[34] R.P. Stanley, *Hilbert function of graded algebras*, Advances in Math., 28 (1973), pp. 57–83.
[35] H. Stetter, *Numerical Polynomial Algebra*, SIAM, Philadelphia, 2004.
[36] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill Posed Problems*, V. H. Winston and Sons, 1977.
[37] van der Waerden B. L., *Algebra*, Frederick Ungar Pub. Co., 1970.
[38] X. Wu and L. Zhi, *Computing the multiplicity structure from geometric involutive form*, in Proc. 2008 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'08), David Jeffrey, ed., New York, N. Y., 2008, ACM Press, pp. 325–332.
[39] ———, *Determining singular solutions of polynomial systems via symbolic-numeric reduction to geometric involutive form*, MM Research Preprints, 27 (2008), pp. 104–122. Accepted for publication in J. Symbolic Comput.
http://www.mmrc.iss.ac.cn/~lzhi/Publications/jsc_zhiwu.pdf.