Computing Isolated Singular Solutions of Polynomial Systems Accurately: Case of Breadth One

Nan Li

Key Laboratory of Mathematics Mechanization
Chinese Academy of Sciences

Joint work with Lihong Zhi

SCAN2010, Lyon, France, September 27-30
A Motivating Example [Ojika 1987]

\[ F = \{ f_1 = x_1^2 + x_2 - 3, \ f_2 = x_1 + \frac{1}{8} x_2^2 - \frac{3}{2} \} \]

• \( \hat{x} = (1, 2) \) is a 3-fold solution since \( F(\hat{x}) = 0 \) and

\[
\left( \frac{\partial}{\partial x_1} - 2 \frac{\partial}{\partial x_2} \right) F(\hat{x}) = 0,
\]

\[
\left( \frac{\partial}{\partial x_2} - \frac{\partial^2}{2 \partial x_1^2} + \frac{2 \partial^2}{\partial x_1 \partial x_2} - \frac{2 \partial^2}{\partial x_2^2} \right) F(\hat{x}) = 0.
\]
A Motivating Example [Ojika 1987]

\[ F = \{ f_1 = x_1^2 + x_2 - 3, \quad f_2 = x_1 + \frac{1}{8} x_2^2 - \frac{3}{2} \} \]

- \( \hat{x} = (1, 2) \) is a 3-fold solution since \( F(\hat{x}) = 0 \) and

\[
\left( \frac{\partial}{\partial x_1} - \frac{2\partial}{\partial x_2} \right) F(\hat{x}) = 0, \\
\left( \frac{\partial}{\partial x_2} - \frac{\partial^2}{2\partial x_1^2} + \frac{2\partial^2}{\partial x_1 \partial x_2} - \frac{2\partial^2}{\partial x_2^2} \right) F(\hat{x}) = 0.
\]

- Newton’s method with \textbf{Digits} = 14 in Maple and an initial solution \([0., 3.]\), after
  - 10 iterations, we get \([0.9920798, 2.0158017]\);
  - 100 iterations, we get \([1.0000363, 1.9999274]\);
  - 1000 iterations, we still get \([1.0000363, 1.9999274]\).
Related Previous Work

- Compute the multiplicity structure of a singular solution
  Zeng’09, Damiano-Sabadini-Struppa’07, Dayton-Zeng’05, Möller-Tenber’01, Mourrain’96, Möller-Stetter’95, Marinari-Mora-Möller’95.

- Refine an approximate singular solution
  Leykin-Verschelde-Zhao’05,06, Lercerf’03, Ojika-Watanabe-Mitsui’83, Decker-Keller-Kelley’80,83, Griewank-Osborne’81, Reddien’78, Rall’66.

- A special case: Jacobian matrix has corank one
  Giusti-Lercerf-Salvy-Yakoubsohn’07, Dayton-Zeng’05, Griewank’85.
### Notations

- Consider a polynomial system $F \in \mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, \ldots, x_n]$, 
  
  $$
  F : \begin{cases}
    f_1(x_1, \ldots, x_n) = 0, \\
    f_2(x_1, \ldots, x_n) = 0, \\
    \vdots \\
    f_n(x_1, \ldots, x_n) = 0.
  \end{cases}
  $$

- Let $I = (f_1, \ldots, f_n)$ be the ideal generated by $f_1, \ldots, f_n$.

- Let $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_n)$ be an isolated singular solution of $F$ with the multiplicity $\mu$. 

Max Noether Conditions

- Let \( D(\alpha) = D(\alpha_1, \ldots, \alpha_n) : \mathbb{C}[x] \to \mathbb{C}[x] \) 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 \( \alpha = [\alpha_1, \ldots, \alpha_n] \).
Max Noether Conditions

• Let $D(\alpha) = D(\alpha_1, \ldots, \alpha_n) : \mathbb{C}[x] \to \mathbb{C}[x]$ 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 $\alpha = [\alpha_1, \ldots, \alpha_n]$.

• Let $\mathcal{D} = \{D(\alpha) | |\alpha| \geq 0\}$, we define the Max Noether space associated to $I$ at $\hat{x}$ as

$$\triangle_{\hat{x}}(I) := \{L \in \text{Span}_\mathbb{C}(\mathcal{D}) | L(f)_{x=\hat{x}} = 0, \forall f \in I\}.$$

$\forall L \in \triangle_{\hat{x}}(I)$ is called a Max Noether condition and $\dim_{\mathbb{C}}(\triangle_{\hat{x}}(I)) = \mu$. 

Closedness Condition

Two morphisms on \( \mathcal{D} = \{ D(\alpha)| |\alpha| \geq 0 \} \)

- Anti-differentiation operator

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

- Differential operator

\[
\Psi_j(D(\alpha)) := D(\alpha_1, \ldots, \alpha_j + 1, \ldots, \alpha_n).
\]
Closedness Condition

Two morphisms on $\mathcal{D} = \{D(\alpha) \mid |\alpha| \geq 0\}$

- Anti-differentiation operator

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

- Differential operator

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

**Theorem.** [Stetter’04] A subspace $\triangle_{\hat{x}}$ of $\text{Span}_\mathbb{C}(\mathcal{D})$ is said to be closed if and only if its dimension is finite and

$$L \in \triangle_{\hat{x}} \iff \Phi_j(L) \in \triangle_{\hat{x}}, \ j = 1, \ldots, n.$$
Case of Breadth One

**Lemma.** Suppose corank of the Jacobian matrix $F'(\hat{x})$ is one, i.e., $\dim \text{Nullspace}(F'(\hat{x})) = 1$ and for a nonnegative integer $k$, let $\triangle^{(k)}(I) = \{ L \in \triangle(\hat{x})(I), \text{order}(L) \leq k \}$, then for $1 \leq k \leq \mu - 1$,

$$\dim(\triangle^{(k)}(I)) = \dim(\triangle^{(k-1)}(I)) + 1,$$

and for $k \geq \mu$,

$$\dim(\triangle^{(k)}(I)) = \dim(\triangle^{(\mu-1)}(I)).$$
Max Noether Conditions in Breadth One Case

**Theorem.** [Li and Zhi’09] Suppose corank\((F'(\hat{x})) = 1\), and \(L_0 = 1_{id}, L_1 = \frac{\partial}{\partial x_1} \in \triangle^{(1)}_{\hat{x}}(I)\). We construct \(L_k\) incrementally for \(k\) from 2 to \(\mu - 1\) by

\[
L_k = P_k + a_{k,2} \frac{\partial}{\partial x_2} + \cdots + a_{k,n} \frac{\partial}{\partial x_n},
\]

where \(P_k\) is obtained from \(L_1, \ldots, L_{k-1}\) by

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

where for \(2 \leq j \leq n\),

\[
Q_{k,j} = \Phi_j(P_k) = a_{2,j}L_{k-2} + \cdots + a_{k-1,j}L_1.
\]

and \(a_{k,j}\) are determined by solving \(L_k(F)_{x=\hat{x}} = 0\), i.e.,

\[
\begin{bmatrix}
\frac{\partial F(\hat{x})}{\partial x_2}, & \ldots, & \frac{\partial F(\hat{x})}{\partial x_n}
\end{bmatrix}
\cdot
[a_{k,2}, \ldots, a_{k,n}]^T = -P_k(F)_{x=\hat{x}}.
\]
Assumptions on Approximate Solutions

Suppose we are given an approximate solution

\[ \hat{x} = \hat{x}_e + \hat{x}_\varepsilon, \]

where

- \( \hat{x}_e \) denotes the exact singular solution of \( F \) with the multiplicity \( \mu \) and \( \text{corank}(F'(\hat{x}_e)) = 1 \).

- \( \hat{x}_\varepsilon \) denotes the error in the solution, \( \|\hat{x}_\varepsilon\| = \varepsilon \ll 1 \).

- \( \sigma_i \) denotes the singular values of \( F'(\hat{x}) \) satisfying

  \[ \sigma_i = \Theta(1), \ 1 \leq i \leq n - 1, \quad \sigma_n = O(\varepsilon). \]
Regularized Newton Iterations

**Theorem.** Under the assumptions, let $A = F'(\hat{x})$, the solution $\hat{y}$ of the following regularized least squares problem

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

will satisfy

$$\|\hat{y}\| = O(\varepsilon), \|F(\hat{x} + \hat{y})\| = O(\varepsilon^2),$$

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

• Compute the right singular vector $\mathbf{r}_1$ of $F' (\hat{\mathbf{x}} + \hat{\mathbf{y}})$ satisfying

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

• Form a unitary matrix $R = [\mathbf{r}_1, \ldots, \mathbf{r}_n]$ and let

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

• $\hat{\mathbf{z}}_e = R^{-1} \hat{\mathbf{x}}_e$ is an exact root of $H$.

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

• $\left\| \frac{\partial H(\hat{\mathbf{z}})}{\partial z_1} \right\| = \left\| F'(\hat{\mathbf{x}} + \hat{\mathbf{y}}) \mathbf{r}_1 \right\| = \sigma'_n = O(\varepsilon).$
Consequence of the Regularized Newton Iteration

**Theorem.** Under the same assumptions, after running one regularized Newton iteration, we have

\[ |\hat{z}_{1,e} - \hat{z}_1| = \Theta(\varepsilon), \]

and

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

**Theorem.** Under the same assumptions, if the multiplicity \( \mu \geq 2 \), then we have

\[ \|L_i(H)_{z=\hat{z}}\| = O(\varepsilon^2), \text{ for } i = 0, \ldots, \mu - 2. \]
An Augmented Polynomial System

**Theorem.** Under the same assumptions, $\forall h \in \mathbb{C}^n$ satisfying $h^*r_1 \neq 0$, $F'(\hat{x}_e)r_1 = 0$, the augmented polynomial system

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

has $(\hat{x}_e, \frac{r_1}{h^*r_1})$ as an isolated singular solution with the multiplicity $\mu - 1$. 
An Augmented Polynomial System

**Theorem.** Under the same assumptions, $\forall h \in \mathbb{C}^n$ satisfying $h^*r_1 \neq 0$, $F'(\hat{x}_e)r_1 = 0$, the augmented polynomial system

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

has $(\hat{x}_e, \frac{r_1}{h^*r_1})$ as an isolated singular solution with the multiplicity $\mu - 1$.

Ojika et al.’82, Leykin et al’05: The multiplicity drops at least one.

Dayton, Zeng’05 Conjecture: The multiplicity drops exact one.
Refining Approximate Singular Solutions

- Run one regularized Newton iteration to $F$ and $\hat{x}$ to get $\hat{y}$. 
Refining Approximate Singular Solutions

- Run one regularized Newton iteration to $F$ and $\hat{x}$ to get $\hat{y}$.
- Compute the null vector $r_1$ of $F'(\hat{x} + \hat{y})$, form a unitary matrix $R = [r_1, \ldots, r_n]$, and set

\[ H(z) := F(Rz), \hat{z} := R^{-1}(\hat{x} + \hat{y}). \]
Refining Approximate Singular Solutions

- Run one regularized Newton iteration to $F$ and $\hat{x}$ to get $\hat{y}$.
- Compute the null vector $r_1$ of $F'(\hat{x} + \hat{y})$, form a unitary matrix $R = [r_1, \ldots, r_n]$, and set
  \[
  H(z) := F(Rz), \quad \hat{z} := R^{-1}(\hat{x} + \hat{y}).
  \]
- Compute the closed approximate Max Noether basis $\{L_0, L_1, \ldots, L_{\mu-1}\}$ of $H$ at $\hat{z}$, and obtain $P_\mu$. 
Refining Approximate Singular Solutions

• Run one regularized Newton iteration to \( F \) and \( \hat{x} \) to get \( \hat{y} \).

• Compute the null vector \( r_1 \) of \( F'(\hat{x} + \hat{y}) \), form a unitary matrix \( R = [r_1, \ldots, r_n] \), and set

\[
H(z) := F(Rz), \quad \hat{z} := R^{-1}(\hat{x} + \hat{y}).
\]

• Compute the closed approximate Max Noether basis \( \{L_0, L_1, \ldots, L_{\mu - 1}\} \) of \( H \) at \( \hat{z} \), and obtain \( P_\mu \).

• Solve

\[
\begin{bmatrix}
P_\mu(H)_{z=\hat{z}}, \frac{\partial H(\hat{z})}{\partial z_2}, \ldots, \frac{\partial H(\hat{z})}{\partial z_s}
\end{bmatrix} v = -L_{\mu - 1}(H)_{z=\hat{z}}
\]

to obtain \( v \) and set \( \delta := \frac{v_1}{\mu} \).
Refining Approximate Singular Solutions

• Run one regularized Newton iteration to \( F \) and \( \hat{x} \) to get \( \hat{y} \).

• Compute the null vector \( \mathbf{r}_1 \) of \( F'(\hat{x} + \hat{y}) \), form a unitary matrix \( R = [\mathbf{r}_1, \ldots, \mathbf{r}_n] \), and set

\[
H(z) := F(Rz), \; \hat{z} := R^{-1}(\hat{x} + \hat{y}).
\]

• Compute the closed approximate Max Noether basis \( \{L_0, L_1, \ldots, L_{\mu-1}\} \) of \( H \) at \( \hat{z} \), and obtain \( P_\mu \).

• Solve

\[
\begin{bmatrix}
P_\mu(H)_{z=\hat{z}} \frac{\partial H(\hat{z})}{\partial z_2}, & \cdots, & \frac{\partial H(\hat{z})}{\partial z_s}
\end{bmatrix}
\mathbf{v} = -L_{\mu-1}(H)_{z=\hat{z}}
\]

to obtain \( \mathbf{v} \) and set \( \delta := \frac{v_1}{\mu} \).

• Return \( \hat{x} := \hat{x} + \hat{y} + \delta \mathbf{r}_1 \).
Rate of Convergence

**Theorem.** Suppose \( \| \hat{x}_\varepsilon \| = O(\varepsilon) \ll 1 \), \( \| F(\hat{x}) \| = O(\varepsilon) \) and \( F'(\hat{x}) \) has corank one approximately. Then the solution \( \hat{x} \) computed by our algorithm satisfies

\[
\| \hat{x} - \hat{x}_e \| = O(\varepsilon^2).
\]

Our refining algorithm converges **quadratically**.
Rate of Convergence

**Theorem.** Suppose $\|\hat{x}_\epsilon\| = O(\varepsilon) \ll 1$, $\|F(\hat{x})\| = O(\varepsilon)$ and $F'(\hat{x})$ has corank one approximately. Then the solution $\hat{x}$ computed by our algorithm satisfies

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

*Our refining algorithm converges quadratically.*

**Remark:** The biggest size of matrices appeared in our approach is $n \times n$, much smaller than $\mu n \times \mu n$ in [Dayton and Zeng’05] and $2^{\mu-1} n \times 2^{\mu-1} n$ [Ojika et al., Leykin et al.].
Example ([Ojika 1987] continued)

Given $\hat{x} = (1.001, 2.002)$ and $F(\hat{x}) \approx (0.4 \times 10^{-2}, 0.2 \times 10^{-2})$.

- Apply one regularized Newton iteration to $F$ and $\hat{x}$ to get

  $\hat{x} + \hat{y} \approx (0.99940, 2.0011998)$

  $F(\hat{x} + \hat{y}) \approx (0.2588 \times 10^{-5}, 0.1295 \times 10^{-5})$;
Example ([Ojika 1987] continued)

Given \( \hat{x} = (1.001, 2.002) \) and \( F(\hat{x}) \approx (0.4 \times 10^{-2}, 0.2 \times 10^{-2}) \).

- Apply one regularized Newton iteration to \( F \) and \( \hat{x} \) to get
  \[
  \hat{x} + \hat{y} \approx (0.99940, 2.0011998) \\
  F(\hat{x} + \hat{y}) \approx (0.2588 \times 10^{-5}, 0.1295 \times 10^{-5});
  \]

- Perform the linear transformation to get
  \[
  \hat{z}_e \approx (0.1341 \times 10^{-2}, 0.1319 \times 10^{-5}); \\
  \left\| \frac{\partial H(\hat{z})}{\partial z_1} \right\| \approx 0.3 \times 10^{-6}.
  \]
Example ([Ojika 1987] continued)

Given $\hat{x} = (1.001, 2.002)$ and $F(\hat{x}) \approx (0.4 \times 10^{-2}, 0.2 \times 10^{-2})$.

- Apply one regularized Newton iteration to $F$ and $\hat{x}$ to get
  
  \[
  \hat{x} + \hat{y} \approx (0.99940, 2.0011998) \\
  F(\hat{x} + \hat{y}) \approx (0.2588 \times 10^{-5}, 0.1295 \times 10^{-5});
  \]

- Perform the linear transformation to get
  
  \[
  \hat{z}_\epsilon \approx (0.1341 \times 10^{-2}, 0.1319 \times 10^{-5}); \quad \left\| \frac{\partial H(\hat{z})}{\partial z_1} \right\| \approx 0.3 \times 10^{-6}.
  \]

- Let $L_0 = D(0, 0)$, $L_1 = D(1, 0)$, and compute
  \[
  L_2 = -0.9960D(2, 0) + 0.0892D(0, 1), \\
  L_2(H)_{z=\hat{z}} = [-0.7168 \times 10^{-4}, 0.1433 \times 10^{-3}]^T.
  \]
Example ([Ojika 1987] continued)

- Form the matrix: \( K = \begin{bmatrix} -0.0714 & 2.2350 \\ 0.0089 & 1.1182 \end{bmatrix} \). Solve the linear system \( K \mathbf{v} = -L_2(H)_{z=\hat{z}} \) to obtain

\[
\mathbf{v} = [-0.4014 \times 10^{-2}, -0.9611 \times 10^{-4}]^T.
\]

The refined solution is

\[
\hat{\mathbf{x}} = (1 - 0.14 \times 10^{-6}, 2 + 0.3 \times 10^{-5}).
\]
Example ([Ojika 1987] continued)

• Form the matrix: \( K = \begin{bmatrix}
-0.0714 & 2.2350 \\
0.0089 & 1.1182
\end{bmatrix} \). Solve the linear system \( K \mathbf{v} = -L_2(H)_{z=\hat{z}} \) to obtain

\[
\mathbf{v} = \begin{bmatrix}
-0.4014 \times 10^{-2} \\
-0.9611 \times 10^{-4}
\end{bmatrix}^T.
\]

The refined solution is

\[
\hat{\mathbf{x}} = (1 - 0.14 \times 10^{-6}, 2 + 0.3 \times 10^{-5}).
\]

• After running the algorithm again, we get

\[
\hat{\mathbf{x}} = (1 - 0.18 \times 10^{-11}, 2 + 0.7 \times 10^{-11}).
\]
**Algorithm Performance in Maple**  \( \text{Digits} = 14 \)

<table>
<thead>
<tr>
<th>System</th>
<th>Zero</th>
<th>( t )</th>
<th>( s )</th>
<th>( \mu )</th>
<th># Digits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ojika1</td>
<td>((1, 2))</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>(2 \rightarrow 5 \rightarrow 11 \rightarrow 15)</td>
</tr>
<tr>
<td>Ojika2</td>
<td>((1, 0, 0))</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>(2 \rightarrow 5 \rightarrow 10 \rightarrow 14)</td>
</tr>
<tr>
<td>Ojika3</td>
<td>((-2.5, 2.5, 1))</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>(2 \rightarrow 4 \rightarrow 9 \rightarrow 14)</td>
</tr>
<tr>
<td>Ojika4</td>
<td>((0, 0, 10))</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>(2 \rightarrow 3 \rightarrow 7 \rightarrow 13)</td>
</tr>
<tr>
<td>Decker2</td>
<td>((0, 0))</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>(2 \rightarrow 5 \rightarrow 15)</td>
</tr>
<tr>
<td>DLZ</td>
<td>((0, 0))</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>(2 \rightarrow 5 \rightarrow 16)</td>
</tr>
<tr>
<td>DZ3</td>
<td>((\frac{2\sqrt{7}}{5} + \frac{\sqrt{5}}{5}, -\frac{\sqrt{7}}{5} + \frac{2\sqrt{5}}{5}))</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>(2 \rightarrow 5 \rightarrow 13)</td>
</tr>
<tr>
<td>Dayton2</td>
<td>((0, 0, 0))</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>(2 \rightarrow 3 \rightarrow 7 \rightarrow 13)</td>
</tr>
<tr>
<td>SY5</td>
<td>((1, 1))</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>(2 \rightarrow 5 \rightarrow 11 \rightarrow 14)</td>
</tr>
<tr>
<td>Menzel1</td>
<td>((1, 1))</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>(2 \rightarrow 5 \rightarrow 10 \rightarrow 14)</td>
</tr>
</tbody>
</table>
THANK YOU!