Verifying Topological Indices for Higher-Order Rank Deficiencies

R. Baker Kearfott
and
Jianwei Dian
Department of Mathematics
University of Louisiana atLafayette
Lafayette, LA 70504
rbk@louisiana.edu

October 27, 2000

Abstract

It has been known how to use computational fixed point theorems to verify existence and uniqueness of a true solution to a nonlinear system of equations within a small region about an approximate solution. This can be done in $\mathcal{O}\left(n^3\right)$ operations, where n is the number of equations and unknowns. However, these standard techniques are only valid if the Jacobi matrix for the system is nonsingular at the solution. In previous work and a dissertation (of Dian), we have shown, both theoretically and practically, that existence and multiplicity can be verified in a complex setting, and in the real setting for odd multiplicity, when the rank defect of the Jacobi matrix at an isolated solution is 1. Here, after a brief introduction, we discuss the case of higher rank defect.

1 Background

Given a system of nonlinear equations, numerical methods can typically produce an approximation $\check{x}$ to a solution x^*. It is then sometimes desirable to compute bounds

$$\begin{aligned} \boldsymbol{x} &= (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) \\ &= ([\underline{x}_1, \overline{x}_1], [\underline{x}_2, \overline{x}_2], \ldots, [\underline{x}_n, \overline{x}_n], \end{aligned}$$

such that $\check{x} \in \boldsymbol{x}$, and such that a computational fixed point theorem can verify that there is a true solution of the nonlinear system within $\boldsymbol{x}$. Specifically, we examine the problem

$$\boxed{\begin{array}{l}\text{Given } F : \boldsymbol{x} \to \mathbb{R}^n \text{ and } \boldsymbol{x} \in \mathbb{IR}^n,\ \textit{rigorously} \text{ verify:} \\ \quad \bullet \ \text{there exists a unique } x^* \in \boldsymbol{x} \text{ such that } F(x^*) = 0.\end{array}} \tag{1}$$

If the Jacobi matrix $F'(x^*)$ is nonsingular, if $\check{x}$ is a sufficiently accurate approximation to x^*, and if the dimensions of $\boldsymbol{x}$ are chosen appropriately, then interval Newton methods will do (1); there are numerous explanations of how this is done, such as those found in [3], [7], [8], and [5, §1.5]. These interval Newton methods are based on interval versions of linear algebra algorithms, such as the Gauss–Seidel method; such interval Newton methods achieve the verification in $\mathcal{O}\left(n^3\right)$ operations.

Common thinking has been that (1) cannot be done when $F'(x^*)$ is singular or excessively ill-conditioned. However, as explained in [6] and [2], if $\mathbb{R}^n$ in (1) is replaced by $\mathbb{C}^n$, then, in principle, existence and uniqueness can still be verified. The steps of the algorithm for this singular-case verification are outwardly similar to the non-singular case, except that there is an extra low-dimensional search. In [6], we exhibited algorithms for the rank-defect-1 case, i.e. when the null space of $F(x^*)$ has dimension 1; we showed theoretically that these algorithms can verify existence and uniqueness in $\mathcal{O}\left(n^3\right)$ operations, and we illustrated this dependency on dimension with actual computations on a discretization of a model nonlinear eigenvalue problem, with dimension ranging from 2 to 320.

Our algorithms are based on rigorous computation the topological index of the map F over a box $\boldsymbol{x}$ of appropriate size centered at the approximate solution $\check{x}$; see [6] and [2] for a review and references. The algorithm in [6] was specific to the case where this topological index was 2, although we subsequently discovered that the algorithm easily generalizes to arbitrary index. This generalization is explained in [2]; the algorithm is also an $\mathcal{O}\left(n^3\right)$ algorithm.

The dissertation [2] also contains a theoretical study and an algorithm dealing directly with $F : \boldsymbol{x} \to \mathbb{R}^n$, where $\boldsymbol{x} \subset \mathbb{R}^n$, rather than dealing with a complex extension. In particular, a heuristic is effective at guessing the topological index of the complex extension at a solution with a rank-1 singularity; if this topological index happens to be odd, then the topological index in $\mathbb{R}^n$ can be verified to be either 1 or -1, much more efficiently (but still with $\mathcal{O}\left(n^3\right)$ operations) than the corresponding verification in the complex extension. This real-space verification has the additional theoretical advantage that an actual solution to $F(x) = 0$ has been verified to exist within $\boldsymbol{x} \subset \mathbb{R}^n$, while the complex computations only verify that a solution, possibly with imaginary components, exists within a small region in complex space containing $\boldsymbol{x} \subset \mathbb{R}^n$.

At present, we are completing implementations of the techniques for arbitrary topological index and for verifying the topological index in real space; we expect this to be straightforward and successful. However, interesting development remains for the case where the null space of $F(x^*)$ is greater than one. Here, we present this case, pointing out opportunities and difficulties. In §2, we explain our general framework for computing the topological index, while we discuss the higher-dimensional null-space case in §3.

2 Topological Index Computations: The General Setting and the Rank 1 Defect Case

Our computations are based on

1. preconditioning the system $F(x) = 0$ by multiplying by a constant matrix Y so the Jacobi matrix for $YF(x^*)$ is approximately diagonal, except in p rows, where the dimension of the null space of $F'(x^*)$ is p;
2. constructing a box $\boldsymbol{x}$, with astutely chosen coordinate widths, centered at the approximate solution $\check{x}$;
3. computing the Brouwer degree of YF, and hence of F, over $\boldsymbol{x}$ by searching the $(n-1)$-dimensional sides of $\boldsymbol{x}$ to verify solutions of a certain system of equations derived from the components of YF.

Details of these ideas, as well as a review of properties of the Brouwer degree and references to comprehensive introductions, appear in [6] and [2]. Of interest here is the fact that, because of the form of the preconditioned system, the search on the boundary can be greatly streamlined. In particular, previous general algorithms for the topological degree, such as the heuristic algorithms in [10] or [4], as well as the rigorous algorithm in [1], have running times that depend exponentially on n. In contrast, in the preconditioned system YF, one can, in effect, express $(n-p)$ of the variables in terms of p variables. When $p = 1$, if the box dimensions are chosen appropriately, all but four of the $4n$ sides of the box in $\mathbb{C}^n$ (treated as a box in $\mathbb{R}^{2n}$) may be eliminated with simple interval evaluations, and the remaining four $(n-1)$-dimensional sides may be handled with one-dimensional searches. Easily-obtainable approximations to the solutions of the system derived from YF further facilitate these one-dimensional searches. Here, we present those details of that process relevant to studying generalization to $p > 1$.

Following [6] and [2], we observe that, if the rank defect of $F'(x^*)$ is p, then the preconditioner Y can be formed as one would compute an inverse of $F'(x^*)$, except with an incomplete LU-factorization based on full pivoting. The resulting preconditioned Jacobi matrix, to within a column permutation, has the form in Figure 1. Hence, if we assume $F(z) = (f_1(z_1, \ldots, z_n), \ldots, f_n(z_1, \ldots, z_n))$ has already been so preconditioned, then, for the rank-1 defect case $p = 1$, the components of f have the form

$$
\begin{aligned}
f_k(z) &= (z_k - x_k^*) + \frac{\partial f_k}{\partial z_n}(x^*)(z_n - x_n^*) + \mathcal{O}\left(\|z - x^*\|^2\right) && (2)\\
&\quad \text{for } 1 \le k \le n-1, \\
f_n(z) &= \frac{1}{2!}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\frac{\partial^2 f_n}{\partial x_{k_1}\partial x_{k_2}}(x^*)(z_{k_1} - x_{k_1}^*)(z_{k_2} - x_{k_2}^*) + \ldots + && (3)\\
&\quad \frac{1}{d!}\sum_{k_1=1}^{n}\cdots\sum_{k_d=1}^{n}\frac{\partial^d f_n}{\partial x_{k_1}\ldots\partial x_{k_d}}(x^*)(z_{k_1} - x_{k_1}^*)\ldots(z_{k_d} - x_{k_d}^*)
\end{aligned}
$$

$$YF(x^*) = \left(\begin{array}{ccccc} & & & & \overbrace{\hphantom{*\ldots*}}^{p} \\ 1 & 0 & \ldots & 0 & *\ldots* \\ 0 & 1 & 0\ldots & 0 & *\ldots* \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & 1 & *\ldots* \\ 0 & \ldots & 0 & 0 & 0\ldots0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & 0\ldots0 \end{array}\right).$$

Figure 1: A singular system of rank $n-p$ preconditioned with an incomplete LU factorization, where "$*$" represents a non-zero element.

$$+\mathcal{O}\left(\|z - x^*\|^{d+1}\right).$$

(See [2].) If the actual solution x^* is sufficiently close to the approximate solution $\check{x}$, then x^* may be replaced by $\check{x}$ in the above equations. Furthermore, if we use the notation $z_k = x_k + iy_k$, $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, and $f_k(z) = u_k(x,y) + iv_k(x,y)$ for $1 \le i \le n$, then the above forms for f_k and f_n may be expressed as

$$\begin{aligned} u_k(x,y) &= (x_k - x_k^*) + \frac{\partial f_k}{\partial x_n}(x^*)(x_n - x_n^*) + \mathcal{O}\left(\|(x - x^*, y)\|^2\right), && (4) \\ v_k(x,y) &= y_k + \frac{\partial f_k}{\partial x_n}(x^*)y_n + \mathcal{O}\left(\|(x - x^*, y)\|^2\right), && (5) \end{aligned}$$

We further define

$$\begin{aligned} &\boldsymbol{x}_{\underline{k}} \equiv (\boldsymbol{x}_1, \boldsymbol{y}_1, \ldots, \boldsymbol{x}_{k-1}, \boldsymbol{y}_{k-1}, \underline{x}_k, \boldsymbol{y}_k, \boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1}, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_n), \quad \text{and} \\ &\boldsymbol{x}_{\overline{k}} \equiv (\boldsymbol{x}_1, \boldsymbol{y}_1, \ldots, \boldsymbol{x}_{k-1}, \boldsymbol{y}_{k-1}, \overline{x}_k, \boldsymbol{y}_k, \boldsymbol{x}_{k+1}, \boldsymbol{y}_{k+1}, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_n), \end{aligned}$$

and define $\boldsymbol{y}_{\underline{k}}$ and $\boldsymbol{y}_{\overline{k}}$ similarly. With this, define $\tilde{F} : \tilde{\mathbf{D}} \subset \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ by $\tilde{F} = (u_1, v_1, \ldots, u_n, v_n)$, and define

$$\tilde{F}_{\neg u_n}(x,y) \equiv \Big(u_1(x,y), v_1(x,y), \ldots, u_{n-1}(x,y), v_{n-1}(x,y), v_n(x,y)\Big).$$

To compute the degree $\mathrm{d}(\tilde{F}, \boldsymbol{z}, 0)$, we will consider $\tilde{F}_{\neg u_n}$ on the boundary of $\boldsymbol{z}$. The boundary of $\boldsymbol{z}$ consists of the $4n$ faces $\boldsymbol{x}_{\underline{1}}$, $\boldsymbol{x}_{\overline{1}}$, $\boldsymbol{y}_{\underline{1}}$, $\boldsymbol{y}_{\overline{1}}$, …, $\boldsymbol{x}_{\underline{n}}$, $\boldsymbol{x}_{\overline{n}}$, $\boldsymbol{y}_{\underline{n}}$, $\boldsymbol{y}_{\overline{n}}$. The box coordinates $\boldsymbol{x}_k$, $\boldsymbol{y}_k$, $1 \le k \le n$ are chosen so $\boldsymbol{x}_k$ is centered on $\check{x}_k$ and $\boldsymbol{y}_k$ is centered on 0, $1 \le k \le n$, and so the widths $\mathrm{w}(\boldsymbol{x}_n)$ and $\mathrm{w}(\boldsymbol{y}_n)$ obey

$$\begin{aligned} \mathrm{w}(\boldsymbol{x}_n) &\le \frac{1}{2} \min_{1 \le k \le n-1} \left\{ \frac{\mathrm{w}(\boldsymbol{x}_k)}{|\partial f_k / \partial x_n(\check{x})|} \right\}, && (6) \\ \mathrm{w}(\boldsymbol{y}_n) &\le \frac{1}{2} \min_{1 \le k \le n-1} \left\{ \frac{\mathrm{w}(\boldsymbol{y}_k)}{|\partial f_k / \partial x_n(\check{x})|} \right\}. && (7) \end{aligned}$$

Provided $\tilde{F}_{\neg u_n}(x,y) \neq 0$ for $(x,y) \in \boldsymbol{x}_{\underline{k}}, \boldsymbol{x}_{\overline{k}} \boldsymbol{y}_{\underline{k}}, \boldsymbol{y}_{\overline{k}}$, $1 \leq k \leq n-1$, the degree $\mathrm{d}(\tilde{F}, \boldsymbol{z}, 0)$ may be computed with the formula

$$
\begin{aligned}
\mathrm{d}(\tilde{F}, \boldsymbol{z}, 0) \quad = \quad & -\sum_{\substack{x_n = \underline{x}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \operatorname{sgn} \left| \frac{\partial \tilde{F}_{\neg u_n}}{\partial x_1 y_1 \ldots x_{n-1} y_{n-1} y_n}(x,y) \right| \\
& + \sum_{\substack{x_n = \overline{x}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \operatorname{sgn} \left| \frac{\partial \tilde{F}_{\neg u_n}}{\partial x_1 y_1 \ldots x_{n-1} y_{n-1} y_n}(x,y) \right| \\
& + \sum_{\substack{y_n = \underline{y}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \operatorname{sgn} \left| \frac{\partial \tilde{F}_{\neg u_n}}{\partial x_1 y_1 \ldots x_{n-1} y_{n-1} x_n}(x,y) \right| \\
& - \sum_{\substack{y_n = \overline{y}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \operatorname{sgn} \left| \frac{\partial \tilde{F}_{\neg u_n}}{\partial x_1 y_1 \ldots x_{n-1} y_{n-1} x_n}(x,y) \right| .
\end{aligned} \tag{8}
$$

If the coordinate extents are chosen as in (6), then (4) makes it unlikely that $u_k = 0$ on $\boldsymbol{x}_{\underline{k}}$ and $\boldsymbol{x}_{\overline{k}}$, $1 \leq k \leq n-1$; similarly, if the coordinate extents are chosen as in (7), then (5) makes it unlikely that $v_k = 0$ on $\boldsymbol{y}_{\underline{k}}$ and $\boldsymbol{y}_{\overline{k}}$, $1 \leq k \leq n-1$. The verification algorithm verifies these facts by computing the interval values $\boldsymbol{u}(\boldsymbol{x}_{\underline{k}})$, $\boldsymbol{u}(\boldsymbol{x}_{\overline{k}})$, $\boldsymbol{v}(\boldsymbol{y}_{\underline{k}})$, and $\boldsymbol{v}(\boldsymbol{y}_{\overline{k}})$. Formula (8) is then used by systematic search (made rigorous with interval computations) of the four faces $\boldsymbol{x}_{\underline{n}}$, $\boldsymbol{x}_{\overline{n}}$, $\boldsymbol{y}_{\underline{n}}$, and $\boldsymbol{y}_{\overline{n}}$ for solutions to $\tilde{F}_{\neg u_n} = 0$. The search is reduced to a one-dimensional search over the y_n coordinate on $\boldsymbol{x}_{\underline{n}}$ and $\boldsymbol{x}_{\overline{n}}$ and a one-dimensional search over the x_n coordinate on $\boldsymbol{y}_{\underline{n}}$ and $\boldsymbol{y}_{\overline{n}}$ by formally solving the rigorous interval enclosure for v_k corresponding to (7) for $\boldsymbol{y}_k$ in terms of $\boldsymbol{y}_n$ and formally solving the rigorous interval enclosure for u_k corresponding to (6) for $\boldsymbol{x}_k$ in terms of $\boldsymbol{x}_n$. For details, see [6] and [2].

The one-dimensional search is facilitated with accurate a priori approximations to the solutions of $\tilde{F}_{\neg u_n}(x,y) = 0$ on $\boldsymbol{x}_{\underline{n}}$, $\boldsymbol{x}_{\overline{n}}$, $\boldsymbol{y}_{\underline{n}}$, and $\boldsymbol{y}_{\overline{n}}$. Denote

$$
\begin{aligned}
\alpha_k & \equiv \frac{\partial f_k}{\partial x_n}(\check{x}), \qquad 1 \leq k \leq n-1, \\
\alpha_n & \equiv -1, \\
\Delta_1 & \equiv \left| \frac{\partial F}{\partial x_1 \ldots \partial x_n}(\check{x}) \right| . \\
\Delta_d & \equiv \sum_{k_1=1}^{n} \cdots \sum_{k_d=1}^{n} \frac{\partial^d f_n}{\partial x_{k_1} \ldots \partial x_{k_d}}(\check{x}) \alpha_{k_1} \ldots \alpha_{k_d}, \qquad 2 \leq d.
\end{aligned} \tag{9}
$$

Then, consider $f_n(z) = (u_n(x,y), v_n(x,y))$, assume the formulas (2) and (3) to be exact without the Landau symbols, and assume the actual topological index is d. In that case, all terms in (3) except terms of order d and higher vanish. Further assuming $f_k(z) = 0$, $1 \leq k \leq n-1$, solving

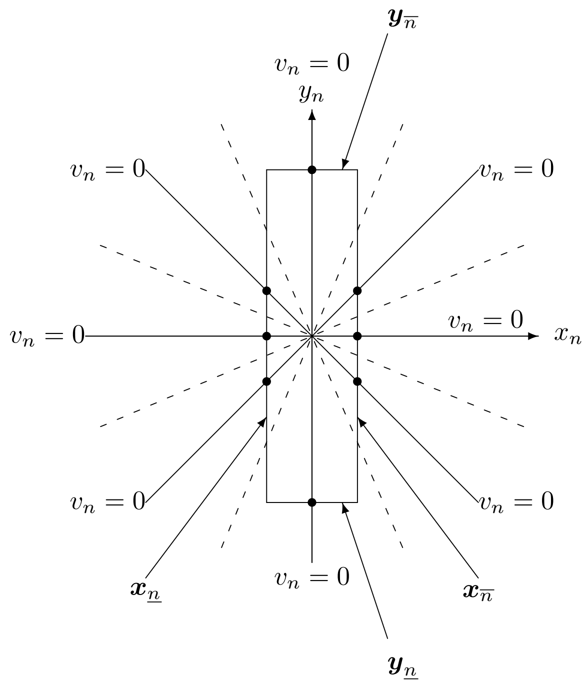

Figure 2: When $d = 4$. $v_n = 0$ on solid lines and $u_n = 0$ on dashed lines. The thick dots represent the solutions of $\tilde{F}_{\neg u_n}(x, y) = 0$ on the boundary.

for z_k in terms of z_n in (2), and plugging into (3) finally gives

$$f_n(z) = \frac{(-1)^d t^d \Delta_d}{d!}(z_n - x_n^*)^d + \mathcal{O}\left(\|z - x^*\|^{d+1}\right). \tag{10}$$

It follows from (10) that solutions of $v_n = \Im f_n = 0$ at those complex values $z_n = (x_n, y_n)$ whose argument θ obeys $\theta^d = \pi/2 + \ell\pi$ for some integer ℓ. This is illustrated for $d = 4$ in Figure 2. For details, see [2].

Thus, in the one-dimensional rank defect case, one-dimensional subintervals of four one-dimensional intervals can be constructed about approximate solutions of $\tilde{F}_{\neg u_n}(x, y) = 0$, these intervals can be rapidly verified to contain unique solutions of $\tilde{F}_{\neg u_n}(x, y) = 0$, and the remainder of the four original one-dimensional intervals can be rapidly but rigorously eliminated with interval evaluations of $\tilde{F}_{\neg u_n} 0$. Actual algorithms appear in [2], numerical results for $d = 2$ appear in [6], and we are presently completing numerical experiments for $d > 2$.

3 The Higher Rank-Defect Case

When the dimension p of the null space of $F'(x^*)$ is greater than 1, the forms corresponding to (2) and (3) are

$$
\begin{aligned}
f_k(z) &= (z_k - x_k^*) + \frac{\partial f_k}{\partial z_n}(x^*)(z_n - x_n^*) + \mathcal{O}\left(\|z - x^*\|^2\right) \qquad (11)\\
&\quad \text{for } 1 \le k \le n-p,\\
f_q(z) &= \frac{1}{2!}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\frac{\partial^2 f_q}{\partial x_{k_1}\partial x_{k_2}}(x^*)(z_{k_1} - x_{k_1}^*)(z_{k_2} - x_{k_2}^*) + \ldots +\\
&\quad \frac{1}{d!}\sum_{k_1=1}^{n}\cdots\sum_{k_d=1}^{n}\frac{\partial^d f_q}{\partial x_{k_1}\ldots\partial x_{k_d}}(x^*)(z_{k_1} - x_{k_1}^*)\ldots(z_{k_d} - x_{k_d}^*)\\
&\quad +\mathcal{O}\left(\|z - x^*\|^{d+1}\right), \qquad \text{for } n-p+1 \le q \le n. \qquad (12)
\end{aligned}
$$

In this more general setting, (11) can be used as before to eliminate variables from (12). However, p variables remain, and there are p equations left. In general, this system is an arbitrary system of p homogeneous degree-d equations in p variables; to see this, let $n = p$ and *specify* the *complete original* system by

$$
\begin{aligned}
f_q(z) &= \frac{1}{d!}\sum_{k_1=1}^{p}\cdots\sum_{k_d=1}^{p}\frac{\partial^d f_q}{\partial x_{k_1}\ldots\partial x_{k_d}}(x^*)(z_{k_1} - x_{k_1}^*)\ldots(z_{k_d} - x_{k_d}^*)\\
&\quad \text{for } 1 \le q \le p, \qquad (13)
\end{aligned}
$$

where the partial derivatives are set arbitrarily, subject only to the condition that corresponding mixed partial derivatives are equal. This implies that, in the analogue of the case when $p = 1$, a p-dimensional space must be searched. Furthermore, for approximate starting solutions, instead of a simple formula as for (10) and Figure 2, all solutions to a general d-homogeneous system of p equations in p unknowns would need to be found. For higher p and d, that could be expensive for a verification step that may be a small part of another overall algorithm.

The general formula from which (8) was derived is [6, Theorem 2.5]). A straightforward change of notation to our complex setting, selection of $s = 1$, and selection of $p = 2n - 1$ (corresponding to the component u_n) in that theorem gives the general formula

$$
\begin{aligned}
\mathrm{d}(F, \boldsymbol{z}, 0) &= \mathrm{d}(\tilde{F}, (\boldsymbol{x}, \boldsymbol{y}), 0) \qquad (14)\\
&= -\sum_{k\in\underline{K_0}}\ \sum_{\substack{z\in\boldsymbol{x}_{\underline{k}}\\ F_{\neg u_n}(z)=0}} \mathrm{sgn}(D_1)\\
&\quad + \sum_{k\in\overline{K_0}}\ \sum_{\substack{z\in\boldsymbol{x}_{\overline{k}}\\ F_{\neg u_n}(z)=0}} \mathrm{sgn}(D_1)\\
&\quad + \sum_{k\in\underline{K_1}}\ \sum_{\substack{z\in\boldsymbol{y}_{\underline{k}}\\ F_{\neg u_n}(z)=0}} \mathrm{sgn}(D_2)\\
&\quad - \sum_{k\in\overline{K_1}}\ \sum_{\substack{z\in\boldsymbol{y}_{\overline{k}}\\ F_{\neg u_n}(z)=0}} \mathrm{sgn}(D_2),
\end{aligned}
$$

where

$$D_1 = \left|\frac{\partial F_{\neg u_n}}{\partial x_1 y_1 x_2 y_2 \dots x_{k-1} y_{k-1} y_k x_{k+1} y_{k+1} \dots x_n y_n}(z)\right| \quad \text{and}$$

$$D_2 = \left|\frac{\partial F_{\neg u_n}}{\partial x_1 y_1 x_2 y_2 \dots x_{k-1} y_{k-1} x_k x_{k+1} y_{k+1} \dots x_n y_n}(z)\right|,$$

and where, following [6, p. 366], $\underline{K_0}$ is that subset of the integers $k \in \{1, \dots, n\}$ such that $F_{\neg u_n} = 0$ has solutions on $\boldsymbol{x}_{\underline{k}}$ and $\mathrm{sgn}(u_n) = 1$ at these solutions, and $\overline{K_0}$ is that subset of the integers $k \in \{1, \dots, n\}$ such that $F_{\neg u_n} = 0$ has solutions on $\boldsymbol{x}_{\overline{k}}$ and $\mathrm{sgn}(u_n) = 1$ at these solutions, $\underline{K_1}$ is that subset of the integers $k \in \{1, \dots, n\}$ such that $F_{\neg u_n} = 0$ has solutions on $\boldsymbol{y}_{\underline{k}}$ and $\mathrm{sgn}(u_n) = 1$ at these solutions, and $\overline{K_1}$ is that subset of the integers $k \in \{1, \dots, n\}$ such that $F_{\neg u_n} = 0$ has solutions on $\boldsymbol{y}_{\overline{k}}$ and $\mathrm{sgn}(u_n) = 1$ at these solutions (Note that now, in contrast to (8), any u_q, $1 \le n-p+1 \le n$ may be chosen for $F_{\neg u_q}$. For simplicity, we continue to choose $q = n$ without loss of generality.)

Now, as in the rank-1 defect case $p = 1$, the box $\boldsymbol{z}$ can be constructed so u_k is probably nonzero on $\boldsymbol{x}_{\underline{k}}$ and $\boldsymbol{x}_{\overline{k}}$, and v_k is probably nonzero on $\boldsymbol{y}_{\underline{k}}$ and $\boldsymbol{y}_{\overline{k}}$ for $1 \le k \le n-p$. In particular, the forms corresponding to (4) and (5) are

$$\begin{aligned} u_k(x,y) &= (x_k - x_k^*) + \sum_{q=n-p+1}^{n} \frac{\partial f_k}{\partial x_q}(x^*)(x_q - x_q^*) \\ &\qquad + \mathcal{O}\left(\|(x-x^*,y)\|^2\right), \end{aligned} \tag{15}$$

$$v_k(x,y) = y_k + \sum_{q=n-p+1}^{n} \frac{\partial f_k}{\partial x_q}(x^*) y_q + \mathcal{O}\left(\|(x-x^*,y)\|^2\right), \tag{16}$$

from which it follows that conditions corresponding to (6) and (7) are

$$\sum_{q=n-p+1}^{n} \left|\frac{\partial f_k}{\partial x_q}(x^*)\right| \mathrm{w}(\boldsymbol{x}_q) \le \frac{1}{2}\mathrm{w}(\boldsymbol{x}_k), \quad 1 \le k \le n-p, \tag{17}$$

$$\sum_{q=n-p+1}^{n} \left|\frac{\partial f_k}{\partial x_q}(x^*)\right| \mathrm{w}(\boldsymbol{y}_q) \le \frac{1}{2}\mathrm{w}(\boldsymbol{y}_k), \quad 1 \le k \le n-p. \tag{18}$$

Therefore, the $4p$ faces $\boldsymbol{x}_{\underline{q}}$ and $\boldsymbol{x}_{\overline{q}}$, $\boldsymbol{y}_{\underline{q}}$, and $\boldsymbol{y}_{\overline{q}}$ for $n-p+1 \le q \le n$ cannot be eliminated from consideration in the sums in (14). In fact, instead of being reduced to (8), (14) can only be reduced to

$$\mathrm{d}(F, \boldsymbol{z}, 0) = \sum_{q=1}^{p} \tag{19}$$

$$\left\{ - \sum_{\substack{x_n = \underline{x}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \mathrm{sgn}(D_1) + \sum_{\substack{x_n = \overline{x}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \mathrm{sgn}(D_1) \right.$$

$$+ \sum_{\substack{y_n = \underline{y}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \operatorname{sgn}(D_2) - \sum_{\substack{y_n = \overline{y}_n \\ \tilde{F}_{\neg u_n}(x,y)=0 \\ u_n(x,y)>0}} \operatorname{sgn}(D_2) \Bigg\} .$$

Thus, instead of four one-dimensional intervals to be searched, $4p$ $2p$-dimensional boxes (corresponding to real and imaginary coordinates of the p variables that could not be eliminated) would need to be searched. All in all, it thus appears that the task increases exponentially with the rank defect p, and the complication of implementation jumps substantially from $p = 1$ to $p = 2$. However, it is conceivable that some kind of preconditioning based on higher-order, in addition to first derivatives, could simplify the process.

4 Uses and Limitations

Verification of topological indices is potentially useful in automatic theorem proving associated with bifurcation theory and practical bifurcation problems. It also could be useful in branch and bound optimization algorithms as described in [5, Ch. 5] or [3], to verify feasibility of a set of constraints that happen to be linearly dependent on isolated parts of the feasible set. Although such linear dependencies appear unlikely (or impossible) from a probabilistic point of view, they do occur in practice. Higher-order rank deficiencies also occur in practice. However, the difficulty of verification appears to increase rapidly with the dimension of the null space, and the implementation becomes significantly more complicated between a one-dimensional and a two-dimensional null space.

In any case, the techniques treated here are in general only applicable to isolated solutions. To see this, note that a condition that the topological degree $\mathrm{d}(F, \boldsymbol{x}, 0)$ be defined is that there be no solutions to $F(x) = 0$ on the boundary of $\boldsymbol{x}$; if a solution x^* to $F(x) = 0$ is not isolated, then $\boldsymbol{x}$ in general cannot be so chosen.

An algorithm for exhaustively analyzing the solution sets of polynomial systems that have higher-dimensional solution sets is described in [9]. Although that algorithm, in its present form, does not claim to rigorously verify, its capabilities are nonetheless impressive.

References

[1] O. Aberth. Computation of topological degree using interval arithmetic, and applications. *Math. Comp.*, 62(205):171–178, January 1994.

[2] J. Dian. *Existence Verification of Higher Degree Singular Zeros of Nonlinear Systems.* PhD thesis, University of Louisiana at Lafayette, 2000.

[3] E. R. Hansen. *Global Optimization Using Interval Analysis.* Marcel Dekker, Inc., New York, 1992.

[4] R. B. Kearfott. An efficient degree-computation method for a generalized method of bisection. *Numer. Math.*, 32:109–127, 1979.

[5] R. B. Kearfott. *Rigorous Global Search: Continuous Problems.* Kluwer, Dordrecht, Netherlands, 1996.

[6] R. B. Kearfott, J. Dian, and A. Neumaier. Existence verification for singular zeros of nonlinear systems. *SIAM J. Numer. Anal.*, 38(2):360–379, 2000.

[7] A. Neumaier. *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge, England, 1990.

[8] H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization.* Wiley, New York, 1988.

[9] A. J. Sommese, J. Verschelde, and C. W. Wampler. Numerical decomposition of the solution sets of polynomial systems into irreducible components, 2000. Accepted for publication in *SIAM J. Numer. Anal.*

[10] F. Stenger. An algorithm for the topological degree of a mapping in $\mathbb{R}^n$. *Numer. Math.*, 25:23–38, 1976.