MATHEMATICS OF COMPUTATION
Volume 80, Number 273, January 2011, Pages 345–377
S 0025-5718(2010)02399-3
Article electronically published on June 9, 2010

REGENERATION HOMOTOPIES FOR SOLVING SYSTEMS OF POLYNOMIALS

JONATHAN D. HAUENSTEIN, ANDREW J. SOMMESE, AND CHARLES W. WAMPLER

Abstract. We present a new technique, based on polynomial continuation, for solving systems of n polynomials in N complex variables. The method allows equations to be introduced one-by-one or in groups, obtaining at each stage a representation of the solution set that can be extended to the next stage until finally obtaining the solution set for the entire system. At any stage where positive dimensional solution components must be found, they are sliced down to isolated points by the introduction of hyperplanes. By moving these hyperplanes, one may build up the solution set to an intermediate system in which a union of hyperplanes "regenerates" the intersection of the component with the variety of the polynomial (or system of polynomials) brought in at the next stage. The theory underlying the approach guarantees that homotopy paths lead to all isolated solutions, and this capability can be used to generate witness supersets for solution components at any dimension, the first step in computing an irreducible decomposition of the solution set of a system of polynomial equations. The method is illustrated on several challenging problems, where it proves advantageous over both the polyhedral homotopy method and the diagonal equation-by-equation method, formerly the two leading approaches to solving sparse polynomial systems by numerical continuation.

1. Introduction

A classical approach to computing the intersection of two algebraic sets A and B in some larger algebraic set X is to replace A by a set A', such that $A' \cap B$ is easier to work with and may be deformed to $A \cap B$. Such results, often called Moving Lemmas [8, §11.4], underlie most of the traditional homotopies used to compute

Received by the editor September 23, 2008 and, in revised form, July 24, 2009 and October 27, 2009.

2010 *Mathematics Subject Classification.* Primary 65H10; Secondary 13P05, 14Q99, 68W30.

Key words and phrases. Algebraic set, component of solutions, diagonal homotopy, embedding, equation-by-equation solver, generic point, homotopy continuation, irreducible component, numerical irreducible decomposition, numerical algebraic geometry, path following, polynomial system, witness point, witness set.

The first author was supported by the Duncan Chair of the University of Notre Dame, the University of Notre Dame Center for Applied Mathematics, NSF Grant DMS-0410047, NSF Grant DMS-0712910, and the Fields Institute.

The second author was supported by the Duncan Chair of the University of Notre Dame, NSF Grant DMS-0410047, and NSF Grant DMS-0712910.

The third author was supported by NSF Grant DMS-0410047 and NSF Grant DMS-0712910.

numerical solutions of polynomial systems. For example, consider a system

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix} = 0 \tag{1.1}$$

of n polynomials of degrees $d_1, \ldots, d_n$, respectively, on complex Euclidean space, $\mathbb{C}^N$, and let $\mathrm{Var}(f)$ denote its solution set. When the system is square, i.e., $n = N$, a common objective is to find all isolated solutions, if any. A basic "total degree" approach degenerates the solution set $\mathrm{Var}(f_i)$ of the i-th equation to a union of d_i linear hyperplanes, for $i = 1, \ldots, N$. Assuming some genericity, the intersection of the hyperplanes gives the total degree, $D = d_1 \cdots d_N$, where N is the number of start points. Homotopy continuation deforms the hyperplanes back to the original polynomials, implicitly defining D solution paths emanating from the start points. The endpoints of these paths all lie in the closure of $\mathrm{Var}(f)$ and include all of its isolated solutions. Numerical polynomial continuation computes approximations to these endpoints by numerically tracking the paths from the start points.

A related problem, more representative of the current contribution, begins with a representation of the solution set $B = \mathrm{Var}(f_1, \ldots, f_k)$, in terms of a witness set (explained in more detail below) and seeks the isolated points in $\mathrm{Var}(f)$ by deforming from $A' \cap B$ to $A \cap B$, where $A = \mathrm{Var}(f_{k+1}, \ldots, f_N)$. This approach hinges on finding an appropriate form of A' such that $A' \cap B$ is easier to solve than $A \cap B$ and yet the final deformation yields all isolated points in $A \cap B$. An algorithm which solves this problem can be adapted to also treat cases where $n \neq N$ and to find witness sets for all solution components of any dimension.

Besides a total degree homotopy of the sort outlined in the opening paragraph, there exist other homotopies that take advantage of various special properties that may be observed in the target system f. Such homotopies reduce the number of paths to be tracked, hence reducing the computational cost of obtaining the solution set. Notable examples are multihomogeneous homotopies [18], linear product homotopies [35], product decomposition homotopies [20], and polyhedral homotopies [9, 36, 15]. These homotopies deform all of the polynomials in f simultaneously to find its isolated solutions in one stage. The method in [40] uses a two-stage approach to solving mixed polynomial-trigonometric systems: the final stage is a product decomposition homotopy whose start system is solved by multiple polyhedral homotopies. This has some resemblance to the more general method introduced here, which we call "regeneration".

Regeneration finds a set of points that includes all isolated solutions in $\mathrm{Var}(f)$, but does so in several successive stages. Instead of deforming the polynomials all at once, we replace some of the $\mathrm{Var}(f_i)$ by hyperplanes (that is, replace f_i by a linear function) and find the isolated solutions of this simpler system. The hyperplanes can subsequently be moved to a succession of positions to *regenerate* f_i. The technique offers great flexibility in the order and number of the original polynomials that are regenerated at any stage. In particular, one may choose to regenerate the polynomials one-by-one to find the isolated points of $\mathrm{Var}(f)$ after n stages of regeneration. Thus regeneration is an "equation-by-equation" method [31].

It is important to note that the solution paths in a homotopy can all be tracked independently. Consequently, homotopy algorithms tend to parallelize efficiently

by distributing path tracking assignments to the available processors. This is an advantage over symbolic methods, which tend to be difficult to parallelize.

Numerical algebraic geometry also includes methods to compute the numerical irreducible decomposition of the solution set of a system of n polynomial equations on $\mathbb{C}^N$ [25, 33]. In this approach, a k-dimensional solution component, say $Z_k \subset \mathrm{Var}(f)$, is represented by a *witness set* whose main constituent is the set of isolated points $Z_k \cap L$, where L is a generic k-codimensional linear space. A *witness superset* for Z_k is a finite set of points, which is contained in $L \cap \mathrm{Var}(f)$ and which contains $Z_k \cap L$. Regeneration can be applied to find witness supersets, after which other methods are used to remove junk points to obtain the true witness set for each dimension and then break these into irreducible components using monodromy [26] verified by linear traces [27]. Specifically, an equation-by-equation form of regeneration builds a witness set for $\mathrm{Var}(f_1, \ldots, f_{k+1})$ out of one previously found for $\mathrm{Var}(f_1, \ldots, f_k)$, proceeding in this fashion to ultimately find a witness set for $\mathrm{Var}(f_1, \ldots, f_n)$.

An existing method, diagonal homotopy [29, 30], can also find witness sets working equation by equation [31]. We compare regeneration to a slightly improved version of diagonal homotopy. We also compare it to polyhedral homotopy, currently considered the most efficient method for attacking sparse polynomial systems.

This paper is organized as follows. After reviewing some background material in §2, we formally state the problems addressed in this paper in §3. Briefly, given a system of polynomials, we may seek all nonsingular isolated solutions, or all isolated solutions, or witness sets for all solution components at every dimension. In §4, we define the notions of a *trackable path* and a *complete homotopy*. §5 reviews the basic constructions that are combined to form the new regeneration method: parameter homotopy and product decomposition. These pieces are brought together in §6 to form the regeneration method for finding isolated roots and in §7 for finding sets at every dimension. In §8, we introduce some improvements to an existing algorithm, equation-by-equation diagonal homotopy, in preparation for comparing it to regeneration. Then, in §9, we compare the performance of regeneration with the diagonal and polyhedral homotopies on some test examples.

2. Background

The book [33] overviews the entire field of numerical algebraic geometry, while the survey [15] is a good reference on solving for isolated solutions, especially using polyhedral homotopy.

2.1. **Genericity.** In this article, we often say that for a generic choice of a point in an irreducible algebraic set Q, such as $\mathbb{C}^M$, some property holds true. This is shorthand for saying that there is a nonempty Zariski open subset of Q for which the property is true. An exception to this is when we say that some property holds true for a general γ from $\mathbb{S}^1 := \{z \in \mathbb{C} \mid |z| = 1\}$; in that case, we mean that the property holds true except for a finite set of $\gamma \in \mathbb{S}^1$.

Genericity, e.g., generic points for a property of an irreducible algebraic set, and Bertini Theorems are discussed thoroughly in [33], starting with [33, Chapter 4]. Often a homotopy depends on some parameters which must be chosen generically for the desirable properties of the homotopy to hold. In practice, we select these parameters using a random number generator, which results in a homotopy that has the desired property with probability one.

2.2. **Varieties and multiplicity.** For a system of polynomials f on $\mathbb{C}^N$, we use the notation $\mathrm{Var}(f)$, read as "variety of f", to mean

$$\mathrm{Var}(f) = \{x \in \mathbb{C}^N | f(x) = 0\}. \tag{2.1}$$

Hence, $\mathrm{Var}(f)$ is just a set of points that carries no multiplicity information. We use $f^{-1}(0)$ to denote $\mathrm{Var}(f)$ with its natural structure of a possibly nonreduced scheme. Though multiplicities may be defined, e.g., [8, §4.3], for any irreducible algebraic subset $Z \subset f^{-1}(0)$, we only need the multiplicity for $Z \subset f^{-1}(0)$, where Z is an irreducible component of $\mathrm{Var}(f)$. When Z is an isolated point x^*, the multiplicity is the complex dimension of the algebraic local ring at x^* coming from the polynomials on $\mathbb{C}^N$ quotiented by the ideal generated by the functions f. This equals the quotient of the ring of convergent power series on $\mathbb{C}^N$ centered at x^* quotiented by the ideal generated by the functions f. Effective numerical methods to compute multiplicities are given in [2, 7].

For an irreducible k-dimensional component Z of $\mathrm{Var}(f)$, the multiplicity of Z as a component of $f^{-1}(0)$ equals the multiplicity of any of the isolated solutions $L \cap Z$ for a general $(N-k)$-dimensional linear space Z, i.e.,

(1) choose generic affine linear equations $L_1, \ldots, L_{N-k}$ on $\mathbb{C}^N$;
(2) choose a point $x^* \in Z \cap \mathrm{Var}(L_1, \ldots, L_{N-k})$;
(3) compute the multiplicity of x^* as a component of $(f \cup \{L_1, \ldots, L_{N-k}\})^{-1}(0)$.

2.3. **Numerical irreducible decomposition.** For a system of polynomials f on $\mathbb{C}^N$, the set $Z := \mathrm{Var}(f)$ is an affine algebraic set, and it decomposes into a union of equidimensional algebraic sets Z_i, where $\dim Z_i = i$, each of which decomposes into a finite number of distinct irreducible components Z_{ij}:

$$Z = \bigcup_{i=1}^{\dim Z} Z_i = \bigcup_{i=1}^{\dim Z} \left(\bigcup_{j \in \mathcal{I}_i} Z_{i,j} \right), \tag{2.2}$$

where

(1) for each i, $Z_i := \bigcup_{j \in \mathcal{I}_i} Z_{i,j}$;
(2) the sets $\mathcal{I}_i$ are finite and each $Z_{i,j}$ is irreducible of dimension i;
(3) $Z_{i,j}$ is not contained in a union of a collection of the $Z_{a,b}$ unless $Z_{i,j}$ occurs in the collection.

The breakup of Z into the $Z_{i,j}$ is called the irreducible decomposition of Z.

Corresponding to the irreducible decomposition is the concept of a numerical irreducible decomposition, which consists of a numerical witness set for each irreducible component. Recall from the introduction that the main constituent of a witness set for a k-dimensional algebraic set $Z_k \subset \mathbb{C}^N$ is the set of $\deg Z_k$ isolated points $Z_k \cap L$, where $L \subset \mathbb{C}^N$ is a generic linear space of codimension k. In addition to the points $Z_k \cap L$, the remaining elements of a witness set are the linear space L and a system f of which Z_k is a component. Any collection C of irreducible components $Z_{i,j}$ of $\mathrm{Var}(f)$ having the same dimension, i, can be numerically represented by a witness set $\{f, L, C \cap L\}$. An algebraic set having components of different dimensions can be represented by a list of witness sets, at least one for each dimension. A "numerical irreducible decomposition" is such a list having one witness set $W_{i,j}$ for the reduction of each irreducible component $Z_{i,j}$. As there is no confusion, we refer to such lists of witness sets also as witness sets, with the

witness set for a single component being a list with just one element. In this way, we may represent any algebraic set by a witness set.

For an equidimensional component Z having witness set $W = \{f, L, Z \cap L\}$, a witness superset is any set of the form $\{f, L, S\}$, where $S \subset \mathrm{Var}(f) \cap L$ is a finite list of points such that $Z \cap L \subset S$. Generally, to compute a witness set, one first computes a witness superset and then discards the junk points, which are any points in S that are not isolated points in $Z \cap L$. Clearly, nonsingular points are isolated, but a singular point in S might be isolated or it might be a member of a higher dimensional component of $\mathrm{Var}(f)$. At present, the only published way to determine which points are junk is to find all components of $\mathrm{Var}(f)$ at every dimension so that one may perform membership tests, such as a homotopy membership test (see below). In principle, the dimension at a point x of $\mathrm{Var}(f)$ could be determined from purely local information, that is, from the germ of f at x. An algorithm to accomplish this is provided in [1].

2.4. **Membership tests.** A membership test for an algebraic set Y determines if a given point, say x^*, is in Y. There are at least two types of membership tests that may be used in the procedures we discuss. The simplest is if Y is given as $\mathrm{Var}(\xi)$ for some (finite) set of polynomials ξ on $\mathbb{C}^N$. Then, the test of whether point x^* is in Y is merely to check if $\xi(x^*) = 0$. Another possibility is that Y is a component of $\mathrm{Var}(\xi)$ given by a witness set, which might in turn be a collection of witness sets for several dimensions. In this case, membership testing is done by a homotopy membership test; see [26] and [33, §15.4]. That is, the witness points are tracked as the linear subspace that slices out the witness set is moved by continuation to a generic linear space containing x^*. If and only if at least one of these paths end at x^*, then x^* is in Y.

2.5. **Randomizations.** Suppose we have a system of n polynomials f on $\mathbb{C}^N$. A randomization of f to size $m \leq n$ is a new polynomial system, say g, of the form $g = Pf$, where $P \in \mathbb{C}^{m \times n}$ is a generic matrix and f and g are treated as column matrices whose entries are polynomials. For any nonsingular $m \times m$ matrix Q, $\mathrm{Var}(Qg) = \mathrm{Var}(g)$. Let $P = \begin{bmatrix} P_1 & P_2 \end{bmatrix}$, where P_1 is $m \times m$. Since P is generic, P_1 is invertible, and so

$$\mathrm{Var}(g) = \mathrm{Var}(Pf) = \mathrm{Var}(P_1^{-1}Pf) = \mathrm{Var}(\begin{bmatrix} I & R \end{bmatrix} f), \tag{2.3}$$

where $R \in \mathbb{C}^{m \times (n-m)}$. Thus, any full randomization $g = Pf$ has a special randomization of the form of $\begin{bmatrix} I & R \end{bmatrix} f$ that gives the same variety. The two forms of randomization give the same generic properties, so we need not distinguish between them. In practice, the special form $g = \begin{bmatrix} I & R \end{bmatrix} f$ is favored for requiring fewer operations to evaluate and potentially having lower degree equations.

Randomization's usefulness derives from the following properties; see [33, Theorem 13.5.1]. Assume that $A \subset \mathbb{C}^N$ is an irreducible algebraic set. Then, for a generic randomization g of f to size $m < n$:

(1) if $\dim A > N - m$, then A is an irreducible component of $\mathrm{Var}(f)$ if and only if it is an irreducible component of $\mathrm{Var}(g)$;
(2) if $\dim A = N - m$, then A is an irreducible component of $\mathrm{Var}(f)$ implies that it is also an irreducible component of $\mathrm{Var}(g)$;

(3) if A is an irreducible component of $\mathrm{Var}(f)$ with $\dim A \geq N - m$, its multiplicity as a component of $\mathrm{Var}(g)$ is greater than or equal to its multiplicity as a component of $\mathrm{Var}(f)$ with equality if either multiplicity is 1.

Of particular importance to us is the fact that for $m \geq N$, subject to genericity, the isolated (respectively, nonsingular) points in $\mathrm{Var}(g)$ include the isolated (respectively, nonsingular) points in $\mathrm{Var}(f)$ [32], [33, §13.5]. When g is "square", that is, when $m = N$, some of the isolated points in $\mathrm{Var}(g)$ may be extraneous, that is, not in $\mathrm{Var}(f)$; these can be eliminated by casting out any point, say x^*, that gives a nonzero evaluation $f(x^*)$. Consequently, to find all isolated (resp. nonsingular) points in $\mathrm{Var}(f)$, we may "square up" f to form g, a generic randomization f to size N and find all isolated (resp. nonsingular) points in $\mathrm{Var}(g)$. After eliminating extraneous roots, one has the desired points.

3. Problem statement

Regeneration is applicable to several basic problems in numerical algebraic geometry.

Problem 1 (Isolated Roots)**.** Let f be a square polynomial system, that is, f consists of N polynomials on $\mathbb{C}^N$, let Y be a proper algebraic subset of $\mathbb{C}^N$, and let Z_0 be the set of isolated points in $\mathrm{Var}(f) \setminus Y$. Given f and a membership test for Y, we identify two problems:

(a) find all nonsingular points in Z_0;
(b) find a finite set $S \subset \mathrm{Var}(f)$ such that $Z_0 \subset S$.

When seeking only the nonsingular isolated roots in Problem 1(a), one may shortcut the complexities of working with sets of multiplicity greater than one, e.g., avoid using deflation described in §4.1. Moreover, in practical problems, the nonsingular roots are often the ones of highest interest. In contrast, Problem 1(b) asks us to find all isolated roots, including those of multiplicity greater than one. As stated, it is enough to find a superset S of these roots; eliminating the junk points (the points in S that are not isolated in $\mathrm{Var}(f)$) is a post-processing step not addressed here.

As noted in the final paragraph of §2.5, the equivalent of Problem 1 for an overdetermined system, i.e., one for which f is a system of n polynomials on $\mathbb{C}^N$ with $n > N$, can be treated by squaring up to size N. For an underdetermined system, i.e., $n < N$, there can be no isolated roots. However, one could append $N - n$ generic linear equations to square up the system, thereby obtaining a witness superset for the $(N - n)$-dimensional component of $\mathrm{Var}(f)$, which, if it exists, is the component of lowest dimension.

Problem 2 (Witness Superset)**.** Let Y, Z be algebraic subsets of $\mathbb{C}^N$. Given a witness set for $(Z \setminus Y) \subset \mathbb{C}^N$, a polynomial system f on $\mathbb{C}^N$, and a membership test for Y, find a witness superset for $(Z \cap \mathrm{Var}(f)) \setminus Y$.

Whereas Problem 1 seeks only isolated roots, Problem 2 seeks witness supersets for components at every dimension, including the zero-dimensional points. Note that if one desires a witness superset for $\mathrm{Var}(f)$, this comes as a special case of Problem 2 with $Z = \mathbb{C}^N$ and $Y = \emptyset$. (A witness set for $\mathbb{C}^N$ is always available; any generic point in $\mathbb{C}^N$ suffices.) One way to find a witness superset for $\mathrm{Var}(f)$ is to attack each dimension independently by appending the requisite number of

general linear equations, squaring up, and applying any algorithm that solves Problem 1 to the squared-up system. We will give a more efficient algorithm based on regeneration.

4. Homotopy

The purpose of a homotopy is to provide a finite set of one-real-dimensional paths, i.e., parameterized by one real parameter, that lead to a desired zero-dimensional algebraic set. The target set is not explicitly known at the outset, but rather is defined implicitly by algebraic conditions on the variables. To be useful, the homotopy specification must include a start point on each path so that a path tracking algorithm can trace the path from the start points to the target set. It is commonly the case that there are extra paths in the homotopy; the homotopy is considered complete if the set of path endpoints is a superset of the target algebraic set. Before formalizing the definition of a *complete homotopy*, we need a definition of a *trackable path*, and some trackable paths require deflation, as described next.

4.1. **Deflation.** Suppose that we have a parameterized system

$$F(x,q) : \mathbb{C}^N \times \mathbb{C}^M \to \mathbb{C}^N \tag{4.1}$$

that is polynomial in x and complex analytic in q. Suppose further that for a fixed $q^* \in \mathbb{C}^M$ we have a point x^* that is an isolated solution of $F(x,q^*) = 0$. It is common that we would like to numerically continue the solution of $F(x,q) = 0$ from x^* as q varies in the neighborhood of q^*. If x^* is nonsingular, this is straightforwardly accomplished by predictor-corrector methods. When x^* is singular and the multiplicity of nearby solutions of $F(x,q) = 0$ as q varies in a neighborhood of q^* equals the multiplicity of x^* as a solution of $F(x,q^*) = 0$, we know the continuation exists, but since the Jacobian of $F(x,q^*)$ is rank deficient at x^*, predictor-corrector methods do not directly work.

One approach to handle such singular points is by using deflation. Deflation is a regularization operation for polynomial systems in several variables. Introduced by Ojika, Watanabe, and Mitsui [23], it was refined by Ojika [22], and brought to its current form for isolated roots, which includes a proof of termination, by Leykin, Verschelde, and Zhao [12, 13] (see also [10]).

If x^* is an isolated solution of a polynomial system $f : \mathbb{C}^N \to \mathbb{C}^N$, then the deflated system is an associated polynomial system $\widehat{f}(x,\xi) = 0$ consisting of $N+N'$ polynomials on $\mathbb{C}^{N+N'}$ with a single nonsingular solution (x^*,ξ^*) lying over x^*.

A parameterized approach that works for components is presented in [33, §13.3.2, §15.2.2]. This approach works if x^* has the same multiplicity structure as the other points in its neighborhood in the sense that the ranks of a finite number of associated matrices for x^* are the same as the corresponding ranks of the matrices for the nearby points. This condition is true for general points, but we do not know if it is true for certain points of interest arising in this article.

There is a second approach [13] to deflating systems with isolated solutions based on using the Dayton-Zeng multiplicity matrix [7, 41]. This method takes "larger steps" than with the usual deflation procedure in the hopes of reducing the computational work involved in deflation. Extending this to components along the exact same lines of [33, §13.3.2, §15.2.2], deflation works for the whole component when done at a generic point of the component, but still can fail when carried out at a specific point on the component.

Happily, a straightforward modification of the procedure with the multiplicity matrix works generically for a component when carried out at a specific point on the component satisfying the mild conditions that hold in the situations of this article.

The Macaulay matrix. Let x^* be an isolated solution of a polynomial system $F : \mathbb{C}^N \to \mathbb{C}^n$ with variables $x = (x_1, \ldots, x_N)$. For a multi-index $\alpha = (k_1, \ldots, k_N)$ with $k_1, \ldots, k_N$ nonnegative integers, let

$$\begin{array}{rcl} |\alpha| & = & k_1 + \cdots + k_N, \\ \alpha! & = & k_1! \cdots k_N!, \\ D^\alpha & = & \frac{1}{\alpha!} \left(\frac{\partial}{\partial x_1}\right)^{k_1} \cdots \left(\frac{\partial}{\partial x_N}\right)^{k_N}, \\ x^\alpha & = & x_1^{k_1} \cdots x_N^{k_N}, \\ (x - x^*)^\alpha & = & (x_1 - x_1^*)^{k_1} \cdots (x_N - x_N^*)^{k_N}. \end{array}$$

The multiplicity matrix of Dayton and Zeng [7] of order d is

$$\mathcal{A}_d(x) := \left[\left(D^\alpha \left((x - x^*)^\beta \cdot F_j \right) \right)(x) \right], \tag{4.2}$$

where the rows are indexed by the indices (β, j) and the columns by the indices α, where $|\beta| \le d-1$, $1 \le j \le n$, and $|\alpha| \le d$. We denote the evaluation of $\mathcal{A}_d(x)$ at x^* by $\mathcal{A}_d$. Letting $P_{d,N}$ denote the dimension of the vector space of polynomials of degree at most d in N variables, this is an $(n \cdot P_{d-1,N}) \times P_{d,N}$ matrix. This matrix has a lot of structure and its apparent size is greatly reduced in calculations [41].

In [13], a slight variant of $\mathcal{A}_d$ is used with the entries changed to $D^\alpha \left(x^\beta \cdot F_j \right)(x^*)$ and with α restricted to not have all entries zero. For us it is convenient to allow the α to have all entries zero. By the *Macaulay matrix*, we mean the matrix

$$\mathcal{M}_d(x) := \left[\left(D^\alpha \left(x^\beta \cdot F_j \right) \right)(x) \right], \tag{4.3}$$

where the rows are indexed by the indices (β, j) and the columns by the indices α, where $|\beta| \le d-1$, $1 \le j \le n$, and $|\alpha| \le d$. We let $\mathcal{M}_d$ denote $\mathcal{M}_d(x^*)$.

Let $\mu(x^*)$ denote the multiplicity of x^* as a solution of $F(x) = 0$. It is known that the coranks c_d of the $\mathcal{A}_d$ are strictly increasing with d until d^*, which equals the smallest $d \ge 1$ with $c_d = c_{d+1}$, and for $d \ge d^*$, the c_d equals the multiplicity $\mu(x^*)$.

Consider the solutions ξ of

$$\mathcal{A}_d \cdot \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{P_{d,N}} \end{bmatrix} = 0. \tag{4.4}$$

Note that $\mathcal{A}_{d+1} = \begin{bmatrix} \mathcal{A}_d & B \\ 0 & C \end{bmatrix}$. From this we see that each solution ξ^* of $\mathcal{A}_d \cdot \xi^* = 0$ gives rise to a solution $\xi = \begin{bmatrix} \xi^* \\ 0 \end{bmatrix}$ of $\mathcal{A}_{d+1} \cdot \xi = 0$. Thus, if $c_{d^*+1} = c_{d^*}$, then the solutions obtained by tacking zeros onto the solutions ξ of $\mathcal{A}_{d^*} \cdot \xi = 0$ are the solutions of $\mathcal{A}_d \cdot \xi = 0$ for any $d > d^*$.

The matrices $\mathcal{M}_d(x)$ may be used instead of the $\mathcal{A}_d(x)$ because the coranks are easily seen to be the same. Indeed $\mathcal{M}_d(x) = G_d \cdot \mathcal{A}_d(x)$, where G_d is a lower triangular matrix with 1's along the diagonal.

Motivated by the partial deflation system used in [13], we have the following result.

Theorem 4.1. *Let x^* denote an isolated solution of a system of polynomials $F : \mathbb{C}^N \to \mathbb{C}^n$. Let c_d be the corank of the Macaulay matrix $\mathcal{M}_d$ for $d \geq 0$. Let $L_1(\xi), \ldots, L_{c_d}(\xi)$ be generic inhomogeneous linear functions. If $c_d = c_{d+1}$, then the polynomial system*

$$\mathcal{F}(x,\xi) := \begin{bmatrix} F(x) \\ \left(D^\alpha \left(x^\beta \cdot F_j\right)\right)(x) \cdot \xi \\ L_1(\xi) \\ \vdots \\ L_{c_d}(\xi) \end{bmatrix} = 0, \tag{4.5}$$

with $1 \leq j \leq n$, $1 \leq |\alpha| \leq d$, and $|\beta| \leq d$, has a unique nonsingular solution of the form (x^*, ξ^*) over x^*.

Remark 4.2. Note that unlike the deflation system in [13], $|\beta|$ may equal d in this result.

Proof. Let ξ^* denote the unique solution of the system $\mathcal{F}(x^*, \xi)$ on $\{x^*\} \times \mathbb{C}^{c_d - 1}$. It suffices to show that the Jacobian matrix of $\mathcal{F}(x, \xi^*)$ has an empty nullspace at (x^*, ξ^*).

Suppose the nullspace is not empty, that is, suppose that there exists a vector $v \neq 0$ with

$$\left(v_1 \frac{\partial \mathcal{F}(x,\xi^*)}{\partial x_1} + \cdots + v_N \frac{\partial \mathcal{F}(x,\xi^*)}{\partial x_N}\right)_{x=x^*} = 0.$$

This relation corresponds to a solution ξ' of

$$\mathcal{M}_{d+1} \cdot \xi' = 0$$

with ξ' having at least one nonzero coordinate corresponding to a D^α with $|\alpha| = d+1$. This contradicts the observations made about Eq. (4.4) discussed right after that equation. □

Remark 4.3. Note that the converse of the above result is true.

Now let's consider the parameterized version of Theorem 4.1.

In applications we have an M-dimensional irreducible component Z of the solution set of a polynomial system $F : \mathbb{C}^{N+M} \to \mathbb{C}^N$ and a general N-dimensional linear space L meeting Z in a finite number of points. By genericity we know that

(1) the points $Z \cap L$ are smooth points of Z considered as a reduced algebraic set;
(2) no point of $Z \cap L$ belongs to any other component of $\mathrm{Var}(F)$;
(3) the multiplicities of the points $Z \cap L$ as isolated solutions of F restricted to L equal the multiplicity of Z as a component of the solution set of $F(x) = 0$.

These properties remain true if we replace $L = (L_1(x), \ldots, L_M(x))$ with any of the parallel N-dimensional linear spaces $\mathrm{Var}(L_1(x) - q_1, \ldots, L_M(x) - q_M)$ for $q := (q_1, \ldots, q_M)$ in a sufficiently small ball around the origin in $\mathbb{C}^M$. The following result applies precisely to this situation.

Theorem 4.4. *Let Z denote an M-dimensional irreducible component of $\mathrm{Var}(F)$, where $F : \mathbb{C}^N \times \mathbb{C}^M \to \mathbb{C}^N$ is a system of polynomials. Let $\pi_1 : \mathbb{C}^N \times \mathbb{C}^M \to \mathbb{C}^N$ and $\pi_2 : \mathbb{C}^N \times \mathbb{C}^M \to \mathbb{C}^M$ be the projection maps $(x, q) \mapsto x$ and $(x, q) \mapsto q$, respectively. Let $F_q : \mathbb{C}^N \to \mathbb{C}^N$ denote $F(x, q)$ for a fixed point $q \in \mathbb{C}^M$. For $z = (x, q) \in Z$ and*

$k \geq 0$, let $c_k(z)$ be the corank of the Macaulay matrix $\mathcal{M}_k$ for $F_q(x)$. Let $L_1(\xi)$, $\ldots$, $L_{c_d}(\xi)$ be generic inhomogeneous linear functions. Assume that $(x^*, q^*) \in Z$, that there are no other irreducible components of $\mathrm{Var}(F)$ containing (x^*, q^*), and that the map $\pi_{2,Z} : Z \to \mathbb{C}^M$ induced by restricting π_2 to Z gives an isomorphism from a complex neighborhood U of (x^*, q^*) to its image in $\mathbb{C}^M$. Then, we have the following.

(1) *For every $z \in Z \cap U$, the multiplicity of $\pi_1(z)$ as an isolated solution of $F_{\pi_2(z)}(x) = 0$ is equal to the multiplicity of Z as a component of $F^{-1}(0)$.*
(2) *There is a finite integer d such that the polynomial system*

$$\mathcal{F}(x, q, \xi) := \begin{bmatrix} F(x, q) \\ \left(D^\alpha \left(x^\beta \cdot F_j\right)\right)(x, q) \cdot \xi \\ L_1(\xi) \\ \vdots \\ L_{c_d}(\xi) \end{bmatrix} = 0, \tag{4.6}$$

with $1 \leq j \leq n$, $1 \leq |\alpha| \leq d$, and $|\beta| \leq d$, has an M-dimensional component Z^* mapping generically one-to-one to Z under the projection $(x, q, \xi) \to (x, q)$.
(3) *There is an open set $\mathcal{U} \subset U$ that contains (x^*, q^*) such that for $z \in \mathcal{U}$, the inverse image of z on Z^* is a nonsingular isolated solution of $\{\mathcal{F}(x, q, \xi), q - \pi_2(z)\} = 0$.*
(4) *Z^* has multiplicity one as a component of the solution set of $\mathcal{F}(x, q, \xi) = 0$.*

Proof. Item (1) is true for generic $z \in Z \cap U$ by the definition of the multiplicity of a component. By assumption, there are no other components of $\mathrm{Var}(F)$ passing through (x^*, q^*) and since F has the same number of polynomials, N, as the codimension of Z in $\mathbb{C}^N \times \mathbb{C}^M$, there can be no embedded components either. Hence, the multiplicity must be constant for every $z \in Z \cap U$.

For item (2), Theorem 4.1 implies that for each point $z \in Z \cap U$, there is a d such that $\mathcal{F}(x, \pi(z), \xi) = 0$ has a nonsingular solution that maps to z under the projection $(x, q, \xi) \mapsto (x, q)$. Since d is an integer determined by the algebraic condition $c_d(z) = c_{d+1}(z)$, it must be constant on a Zariski open subset W of $Z \cap U$. Call this value d^*. By upper semi-continuity, on the complement of W, the codimensions $c_k(z)$ of $\mathcal{M}_k(z)$ for $k < d^*$ can increase, but cannot decrease, compared to the corresponding $c_k(z)$ on W. But by item (1), for large enough k, $c_k(z)$ is equal for all $z \in Z \cap U$. This means that $d \leq d^*$ on $(z \cap U) \setminus W$. But as discussed after Eq. (4.4), for d larger than the minimum k where $c_k(z) = c_{k+1}(z)$, the system has the same nonsingular root as at the minimum such k, only with zeros tacked onto ξ. Thus, the system $\mathcal{F}(x, q, \xi)$ for $d = d^*$ gives a nonsingular root for all $z \in Z \cap U$ that projects to z. Since ξ is unique over each z, these form an M-dimensional nonsingular component Z^* that maps one-to-one to $Z \cap U$ and hence generically one-to-one on all of Z. The remaining items are consequences of the nonsingularity of Z^*. □

To apply the above deflation, we can proceed as follows.

(1) Compute the coranks c_d at point (x^*, q^*) until $c_d = c_{d+1}$.
(2) Construct the system $\mathcal{F}(x, q, \xi)$ and randomize it to a system of $N + P_{d,N}$ equations in the $N + M + P_{d,N}$ variables (x, q, ξ).

(3) Use homotopy continuation on the randomized system to track points on Z for a short distance starting at (x^*, q^*) over a random line through q^*.
(4) If the tracked points are on Z, the deflation system has been found. If not, increment d by one and go back to step 2.

Theorem 4.4 implies that the procedure will terminate with $d = d^*$, where d^* is as defined in the proof of the theorem.

4.2. **Trackable paths.**

Definition 4.5 (Trackable Path)**.** Let $H(x,t) : \mathbb{C}^N \times \mathbb{C} \to \mathbb{C}^N$ be polynomial in x and complex analytic in t and let x^* be an isolated solution of $H(x,1) = 0$. We say that x^* is *trackable* (or equivalently we say that we can *track* x^*) for $t \in (0,1]$ from $t = 1$ to $t = 0$ using $H(x,t)$ if

(1) when x^* is nonsingular, there is a smooth map $\psi_{x^*} : (0,1] \to \mathbb{C}^N$ such that $\psi_{x^*}(1) = x^*$ and $\psi_{x^*}(t)$ is a nonsingular isolated solution of $H(x,t) = 0$ for $t \in (0,1]$;
(2) when x^* is isolated singular, letting $\widehat{H}(x,t,\xi) = 0$ denote the system that arises through deflation, and letting (x^*, ξ^*) denote the nonsingular isolated solution of $\widehat{H}(x,1,\xi) = 0$ over x^*, we can track the nonsingular path starting at (x^*, ξ^*) for $\widehat{H}(x,1,\xi)$ for $t \in (0,1]$ from $t = 1$ to $t = 0$; i.e., there is a smooth map $\psi_{x^*} : (0,1] \to \mathbb{C}^N \times \mathbb{C}^{N'}$ such that $\psi_{x^*}(1) = (x^*, \xi^*)$ and $\psi_{x^*}(t)$ is a nonsingular isolated solution of $H(x,t,\xi) = 0$ for $t \in (0,1]$.

By the *limit of the tracking using* $H(x,t) = 0$ of the point x^* as t goes to 0, we mean $\lim_{t\to 0} \psi_{x^*}(t)$ in case (1) and the x coordinates of $\lim_{t\to 0} \psi_{x^*}(t)$ in case (2).

The actual tracking of a path can be carried out by numerical path tracking methods, usually predictor-corrector algorithms. Notice that in Definition 4.5, the path need only be nonsingular for $t \in (0,1]$, and thus the usual predictor-corrector methods, based on Newton's method, are not guaranteed to converge all the way to $t = 0$. Endgame algorithms that can compute the limits $\lim\limits_{t\to 0} \psi_{x^*}(t)$ of the paths overcome this difficulty [33, Chapter 10].

4.3. **Endpoints at infinity.** The endpoints of some paths in a homotopy may diverge to infinity. Such paths are numerically difficult to track and they are infinitely long. Fortunately, there is a simple maneuver that tames these paths [17]: homogenize polynomials on $\mathbb{C}^N$ to get polynomials on $\mathbb{P}^N$ and perform computations on a generic patch of $\mathbb{P}^N$, which means restricting to a generic hyperplane in $\mathbb{C}^{N+1}$. In this way, the formerly infinitely long paths to infinity become finite length and their endpoints can be computed accurately. This allows us to clearly distinguish between finite endpoints of large magnitude and true endpoints at infinity. The endpoints at infinity may then be discarded.

4.4. **Complete homotopy.** With the above definition of a trackable path, we are ready to formalize the definition of a complete homotopy.

Definition 4.6 (Complete Homotopy)**.** Let $H(x,t) : \mathbb{C}^N \times \mathbb{C} \to \mathbb{C}^N$ be polynomial in x and complex analytic in t. Let S be a finite set of points in $\mathrm{Var}(H(x,1))$. Then, $H(x,t)$ with S is a *complete homotopy* for an algebraic set $Y \subset \mathbb{C}^N$ if

(1) every point in S is trackable, and
(2) every isolated point in Y is the limit of at least one such path.

A complete homotopy is a theoretical construct that begets a corresponding numerical homotopy method. In the numerical method, we begin with numerical approximations to the points in S and obtain numerical approximations to a set of points that is a superset of the isolated points in Y. It is understood that any singular points in S are handled by deflation as described in Definition 4.5, case (2).

5. Building blocks

In this section, we review two theoretical building blocks that are the foundation for regeneration: parameter continuation and product decomposition.

5.1. **Parameter continuation.** We have the following basic result.

Theorem 5.1 (Parameter Continuation)**.** *Suppose that $f(x,q) : \mathbb{C}^N \times \mathbb{C}^M \to \mathbb{C}^N$ is a system of polynomials. Let U be a Zariski open dense subset of $\mathbb{C}^N$. There is a Zariski open set $\mathcal{U} \subset \mathbb{C}^M$ of points q where the maximal number not counting multiplicities of isolated (respectively, nonsingular) points in $\mathrm{Var}(f(x,q)) \cap U$ for $q \in \mathbb{C}^M$ is taken on. Let S be the set of isolated (respectively, nonsingular) solutions to $f(x,q_1) = 0$ on U for $q_1 \in \mathcal{U}$. Then, given any $q_0 \in \mathbb{C}^M$ and all but a finite number of $\gamma \in S^1$, the homotopy function $H(x,t) = f(x, tq_1 + (1-t)q_0)$, with $t = \gamma\tau/[1 + (\gamma - 1)\tau]$, $\tau \in (0,1]$ with start points S is a complete homotopy for finding the isolated (respectively, nonsingular) solutions of $f(x,q_0) = 0$ contained in U.*

The nonsingular case was proven in [19] and the isolated case is proven in [33, Theorem 7.1.6] for general q_1; see also [33, Corollary A.14.2]. The case here with $q_1 \in \mathcal{U}$ follows from [33, Theorem 7.1.6] and the deflation algorithms after Theorem 4.4.

These results also show the stronger fact, that we do not use in this article, that there is a Zariski open set $\mathcal{U}$ of q, where the maximal number of isolated solutions of $f_q(x) = 0$ of given multiplicity contained in U for $q \in \mathbb{C}^M$ is taken on.

There are considerably more general versions [33, §A.14], replacing $\mathbb{C}^N$ and $\mathbb{C}^M$ with more general sets and allowing the $f(x,q)$ to be merely complex analytic in the q variables.

In the homotopy of Theorem 5.1, the reparameterization of t in terms of τ avoids the possibility that the path in parameter space might intersect $\mathbb{C}^M \setminus \mathcal{U}$ enroute from q_1 to q_0 [33, Lemma 7.1.3 ("Gamma Trick")]. If q_1 is chosen randomly in $\mathbb{C}^M$, then the straight-line parameter path $tq_1 + (1-t)q_0$, $t \in (0,1]$, suffices with probability one [33, Lemma 7.1.2].

Let $\#_{\mathrm{isol}}(U)$ (respectively, $\#_{\mathrm{reg}}(U)$) be the maximal number, not counting multiplicities, of isolated (respectively, nonsingular) points in $\mathrm{Var}(f(x,q)) \cap U$. Let $\#_{\mathrm{isol}}(q^*, U)$ (respectively $\#_{\mathrm{reg}}(q^*, U)$) denote the number, not counting multiplicities, of isolated (respectively, nonsingular) points in $\mathrm{Var}(f(x,q^*)) \cap U$.

Remark 5.2. There are two ways in which Theorem 5.1 is typically used.

The first use, and the focus of [19], is to justify a procedure to solve systems with parameters which need to be solved many times. The procedure is to solve $F(x,q_1) = 0$ once for a generic q_1 and thereafter only track the solutions found for $q = q_1$ using homotopies over curves in the q-parameter space starting at q_1 and ending at the q_0. In engineering problems it is common for the cost of solving the

system for special values after solving for a general q^* to be orders of magnitude cheaper than the solution of the general q^*, e.g., [39].

The second use, which is the way we will use this result, is to construct a homotopy to find the isolated or nonsingular solutions on a nonempty Zariski open set $U \subset \mathbb{C}^N$ of a specific system $f(x) = 0$, where $f(x) : \mathbb{C}^N \to \mathbb{C}^N$ is a polynomial system. This is done by embedding $f(x)$ into a family of polynomial systems, i.e., constructing a polynomial system $F(x, q) : \mathbb{C}^N \times \mathbb{C}^M \to \mathbb{C}^N$ with $F(x, q_0) = f(x)$, and then showing that for some $q_1 \in \mathbb{C}^M$ with $F(x, q_1)$ of a particular form, $\#_{\text{isol}}(q_1, U) = \#_{\text{isol}}(U)$ (respectively $\#_{\text{reg}}(q_1, U) = \#_{\text{reg}}(U)$).

5.2. **Product decomposition.** The idea of a homotopy based on a product decomposition was introduced in [20] with related ideas in [35]. Suppose that V_1 and V_2 are vector spaces of polynomials on $\mathbb{C}^N$; in other words, each polynomial in V_1 is a linear combination, with coefficients in $\mathbb{C}^N$, of a given set of basis polynomials, and similarly for V_2. It is convenient to write $V_1 = \langle \alpha_1, \ldots, \alpha_k \rangle$, where the α_i are the basis polynomials for V_1. Similarly, let's write $V_2 = \langle \beta_1, \ldots, \beta_\ell \rangle$. Then, the image of $V_1 \otimes V_2$ in the space of polynomials is another vector space of polynomials whose basis is all products $\alpha_i \beta_j$, $\alpha_i \in V_1$, $\beta_j \in V_2$.

By a *product decomposition* of polynomials on $\mathbb{C}^N$, or a product decomposition on $\mathbb{C}^N$ for short, we mean a list $\mathbb{V} = \{V, V_1, \ldots, V_m\}$ of vector spaces $V, V_1, \ldots, V_m$ of polynomials on $\mathbb{C}^N$ such that V is the image of $V_1 \otimes \cdots \otimes V_m$ in the space of polynomials. We say that a product decomposition $\mathbb{V} := \{V, V_1, \ldots, V_m\}$ is a *product decomposition* of a polynomial f on $\mathbb{C}^N$ if $f \in V$.

Suppose we have a product decomposition of polynomials $\{V, V_1, \ldots, V_m\}$ and form a polynomial g as a product of one generic polynomial from each V_i. Clearly, g is also in the image V. Call any such g a *generic product member* of V.

For example, consider $f(x_1, x_2) = 3 - 2x_1 + 5x_2 - x_1x_2 + 3x_1x_2^2$. Let $V = \langle 1, x_1, x_2, x_1x_2, x_1x_2^2 \rangle$, $V_1 = \langle 1, x_1, x_1x_2 \rangle$, and $V_2 = \langle 1, x_2 \rangle$. Then, $\mathbb{V} = \{V, V_1, V_2\}$ is a product decomposition of f, and $g = (a_1 + a_2x_1 + a_3x_1x_2)(b_1 + b_2x_2)$ is a generic product member of V provided that $a_i, b_j \in \mathbb{C}$ are random for $i = 1, 2, 3$ and $j = 1, 2$.

Another product decomposition for this polynomial f is $\widehat{\mathbb{V}} = \{\widehat{V}, \widehat{V}_1, \widehat{V}_2, \widehat{V}_3\}$, where $\widehat{V} = \langle 1, x_1, x_2, x_1x_2, x_2^2, x_1x_2^2 \rangle$, $\widehat{V}_1 = \langle 1, x_1 \rangle$, and $\widehat{V}_2 = \widehat{V}_3 = \langle 1, x_2 \rangle$. This product decomposition is called a linear product decomposition since each vector space forming the product has a basis that is a subset of $\{1, x_1, x_2\}$. In the regeneration method as described below, we will use a linear product decomposition in the product decomposition phase. It should be noted that such linear products are nearly identical to the set structures described in [35], although the theory presented there only covers nonsingular solutions on $X = \mathbb{C}^N$ and each set must contain 1.

Given a set $\mathcal{D} = \{\mathbb{V}_1, \ldots, \mathbb{V}_k\}$ of product decompositions $\mathbb{V}_i = \{V_i, V_{i,1}, \ldots, V_{i,d_i}\}$ of polynomials on $\mathbb{C}^N$ for $i = 1, \ldots, k$, there is a universal family $F(x, v)$ of polynomial systems on $\mathbb{C}^N$. The parameter space for the system is the complex Euclidean space $\mathcal{V}(\mathcal{D}) = V_1 \times \cdots \times V_k$. Noting that each $v_i \in V_i$ is a polynomial on $\mathbb{C}^N$ and setting $F_i(x, v) = v_i(x)$ for $1 \le i \le k$, we have the universal system for this parameter space:

$$F(x, v) = \{F_1(x, v), \ldots, F_k(x, v)\}.$$

We also have the irreducible subset $\mathcal{P}(\mathcal{D})$ of the parameter space equal to the image of $\prod_{i=1}^{k}\left(\prod_{j=1}^{d_i} V_{i,j}\right)$ in $\mathcal{V}$. We call $\mathcal{P}(\mathcal{D})$ the space of decomposable systems of the family $F(x,v)$.

If some of the V_i are one dimensional, that is, if some of the V_i are of the form $\langle f_i(x)\rangle$, it is convenient to simplify the parameter spaces, assuming that $k < N$ and that we have a set $\mathcal{D} = \{\mathbb{V}_1, \ldots, \mathbb{V}_k, f_{k+1}, \ldots, f_N\}$ where the $\mathbb{V}_i$ are product decompositions on $\mathbb{C}^N$ and where the f_j are polynomials on $\mathbb{C}^N$. Then $\mathcal{V}(\mathcal{D}) = V_1 \times \cdots \times V_k$ is the parameter space for the polynomial systems

$$F(x,v) = \{F_1(x,v), \ldots, F_k(x,v), f_{k+1}(x), \ldots, f_N(x)\}.$$

The irreducible set of decomposable systems of the family is the image $\mathcal{P}(\mathcal{D})$ of $\prod_{i=1}^{k}\left(\prod_{j=1}^{d_i} V_{i,j}\right)$ in $\mathcal{V}(\mathcal{D})$.

Let g be a generic product member of V. Since g is already factored into a product, unlike a general member of V which is a sum of products, replacing f with g in a system of polynomials makes a new polynomial system that is easier to solve. Product decomposition methods use this new system as the start system in a homotopy to solve the original system that contains f. We state this precisely in the following theorem.

Theorem 5.3 (Product Decomposition)**.** *Let $F = \{f_1(x), \ldots, f_N(x)\}$ be polynomials on $\mathbb{C}^N$. Fix an integer k in $[1, N]$. Let U be a nonempty Zariski open subset of $\mathbb{C}^N$. For $i = 1, \ldots, k$, let $\mathbb{V}_i = \{V_i, V_{i,1}, \ldots, V_{i,d_i}\}$ be a product decomposition for polynomials on $\mathbb{C}^N$ such that $f_i(x) \in V_i$, $i = 1, \ldots, k$, and let g_i be a generic product member of V_i. Assume that for any point $x \in U$ and any $1 \le i \le k$, there is at least one element of V_i not zero at x. Let S be the set of isolated (respectively, nonsingular isolated) points in $\mathrm{Var}(\{g_1, \ldots, g_k, f_{k+1}, \ldots, f_N\}) \cap U$. Then, for all but a finite set of $\gamma \in \mathbb{S}^1$, the homotopy*

$$H(x,t) = \{(1-t)f_1(x) + t\gamma g_1(x), \ldots, (1-t)f_k(x) + t\gamma g_k(x), f_{k+1}(x), \ldots, f_N(x)\} = 0 \tag{5.1}$$

with start set S is a complete homotopy for the isolated (respectively, nonsingular isolated) points in $\mathrm{Var}(F) \cap U$.

Proof. The theory presented in [20] covers the case where $k = N$ and we seek only nonsingular isolated solutions.

The theorem for $k = 1$ is shown in Appendix A below. Here we reduce to that result.

Fixing k, we have a universal family of polynomial systems as above consisting of

(1) the parameter space $\mathcal{V}_k = V_1 \times \cdots \times V_k$;
(2) the universal system $F(x,v) = \{F_1(x,v), \ldots, F_k(x,v), f_{k+1}(x), \ldots, f_N(x)\}$;
(3) the irreducible set $\mathcal{P}_k$ of decomposable systems, which is equal to the image of $\prod_{i=1}^{k}\left(\prod_{j=1}^{d_i} V_{i,j}\right)$ in $\mathcal{V}_k$.

By Theorem 5.1 it suffices to prove that for at least one $p \in \mathcal{P}_k$, $\#_{isol}(p, U) = \#_{isol}(U)$. Note if we show it for one, then the Zariski open set of $\mathcal{V}$ where $\#_{isol}(U)$ is taken on will have nonempty intersection with the irreducible algebraic set $\mathcal{P}_k$, and therefore will meet it in a nonempty Zariski open set.

To show that $\#_{isol}(p, U) = \#_{isol}(U)$, it suffices to show that the theorem is true where $v = \{f_1(x), \ldots, f_k(x)\}$ is a generic point of $\mathcal{P}_k$. Indeed, if the theorem is true for this system, then there is at least one $p \in \mathcal{P}_k$ with the number of isolated solutions of the system $\{p_1(x), \ldots, p_k(x), f_{k+1}, \ldots, f_N(x)\}$ on U at least equal to $\#_{isol}(v, U)$, which is equal to $\#_{isol}(U)$.

To prove this we can use induction. The essential case when $k = 1$ is proven in Appendix A. Starting with general $v = \{f_1(x), \ldots, f_k(x)\}$, we successively construct k homotopies of the sort in the theorem: from the first homotopy, which starts at $v_1 = \{p_1(x), f_2(x) \ldots, f_k(x)\}$ and ends at v, up to the last homotopy, which starts at $v_k = \{p_1(x), \ldots, p_k(x)\}$ and ends at v_{k-1}. We have

$$\#_{isol}(U) = \#_{isol}(v, U) \leq \#_{isol}(v_1, U) \leq \cdots \leq \#_{isol}(v_k, U).$$

□

6. Regeneration for isolated roots

This section addresses Problem 1, in which we seek the isolated roots of a square system, that is, a system where the number of polynomials is equal to the number of variables. We have already explained in §3 that the ability to solve a square system extends to the ability to also solve nonsquare systems.

6.1. **Support linears.** An intermediate step to an efficient solution of Problem 1 requires a new definition. In practice, in a system of polynomials on $\mathbb{C}^N$, say $f(x) = \{f_1(x), \ldots, f_n(x)\}$, it often happens that not all variables $\{x_1, \ldots, x_N\}$ appear in all the polynomials. This leads to the following definition.

Definition 6.1. Let $g(x_1, \ldots, x_N)$ be a polynomial. The *support base* of g is:

- if g is homogeneous, the subset of variables that appear in g;
- otherwise, the union of the subset of variables that appear in g with 1.

Definition 6.2. The *support base* of a collection of polynomials is the union of the support bases of the individual polynomials.

Recall that the support of g is the set of all monomials that appear in g. For sets of monomials C and D, let $C \otimes D$ denote the set consisting of the products of monomials in C and D. If G is the support of g, B is the support base of g, and d is the degree of g, then

$$G \subset \underbrace{B \otimes \cdots \otimes B}_{d \text{ times}}.$$

Associated to the concept of a support base is a support hyperplane.

Definition 6.3. For a collection of polynomials $g_1, \ldots, g_k \in \mathbb{C}[x_1, \ldots, x_N]$, a *linear support set* for $g_1, \ldots, g_k$ is any subset of $\{1, x_1, \ldots, x_N\}$ that includes the support base $g_1, \ldots, g_k$. Associated to a linear support set, say Σ, is a *linear support vector space*, say $V = \langle \Sigma \rangle$; i.e., the basis elements of V are the monomials in Σ. A *support linear* for $g_1, \ldots, g_k$ is a linear function in a linear support vector space for $g_1, \ldots, g_k$, and it is said to be a *generic support linear* if its coefficients are generic. The zero set of a support linear is called a *support hyperplane*. A *minimal support linear* for $g_1, \ldots, g_k$ is one whose monomials are exactly the support base of $g_1, \ldots, g_k$, and its zero set is called a *minimal support hyperplane*.

As an example, suppose $g_1 = 2x_1^3 + 3x_1x_3 + 1.2$ and $g_2 = x_1^3 + 2x_1x_3^2 + x_3^2x_4$ in $\mathbb{C}[x_1, \ldots, x_4]$. The support of g_1 is $G_1 = \{1, x_1^3, x_1x_3\}$ with support base $B_1 = \{1, x_1, x_3\}$. Additionally, the support of g_2 is $G_2 = \{x_1^3, x_1x_3^2, x_3^2x_4\}$ with support base $B_2 = \{x_1, x_3, x_4\}$. The support of g_1, g_2 is $G = \{1, x_1^3, x_1x_3, x_1x_3^2, x_3^2x_4\}$ with support base $B = \{1, x_1, x_3, x_4\}$. One may easily confirm that $G \subset B \otimes B \otimes B$. A minimal support hyperplane for g_1, g_2 is $\mathrm{Var}(\alpha + \beta x_1 + \gamma x_3 + \delta x_4)$ for any constants $(\alpha, \beta, \gamma, \delta) \in \mathbb{C}^4$, and $\mathrm{Var}(\alpha + \beta x_1 + \gamma x_2 + \delta x_3 + \zeta x_4)$ is a support hyperplane of g_1, g_2 but not a minimal one.

6.2. **Incremental regeneration.** Our strategy for solving a square polynomial system will consist of several stages of regeneration, starting with a subset of the polynomials and bringing in new ones at each subsequent stage until finally we have the solutions to the full system. Each regeneration stage has two main steps. First, parameter continuation is used to obtain start points for a product decomposition homotopy. Tracking the corresponding paths for this product decomposition homotopy completes the regeneration stage.

The parameter continuation step regenerates a linear product form related to the new polynomials to be introduced at that stage. This regeneration step is summarized in the following lemma.

Lemma 6.4 (Regeneration of a Linear Product)**.** *Let $U \subset \mathbb{C}^N$ be a nonempty Zariski open set, let $f_1, \ldots, f_m$ be polynomials on $\mathbb{C}^N$, and suppose that for $i = m+1, \ldots, \widehat{m}$, $m < \widehat{m} \leq N$, we have $g_i = \ell_{i,1} \cdots \ell_{i,d_i}$, where $d_i \geq 1$ and each $\ell_{i,j}$ is a linear function on $\mathbb{C}^N$. Further, let S_m be the isolated (resp., nonsingular isolated) points of*

$$\mathrm{Var}(f_1, \ldots, f_m, h_{m+1}, \ldots, h_N) \cap U,$$

where for $i = m+1, \ldots, \widehat{m}$, h_i is a generic support linear for $\ell_{i,1}, \ldots, \ell_{i,d_i}$ and $h_{\widehat{m}+1}, \ldots, h_N$ are polynomials on $\mathbb{C}^N$. Let $T_{m,\widehat{m}}$ be the isolated (resp., nonsingular isolated) points of

$$\mathrm{Var}(f_1, \ldots, f_m, g_{m+1}, \ldots, g_{\widehat{m}}, h_{\widehat{m}+1}, \ldots, h_N) \cap U.$$

Finally, let $\mathcal{I}_{m,\widehat{m}} \in \mathbb{N}^{\widehat{m}-m+1}$ be the index set $[1, d_{m+1}] \times \cdots \times [1, d_{\widehat{m}}]$. Then, for any particular $a = (a_{m+1}, \ldots, a_{\widehat{m}}) \in \mathcal{I}_{m,\widehat{m}}$, the start points S_m, and the homotopy function

$$\begin{aligned} H^{\mathrm{parm}}_{m,\widehat{m},a}(x,t) = \{&f_1, \ldots, f_m, \\ &(1-t)\ell_{m+1,a_{m+1}} + th_{m+1}, \ldots, (1-t)\ell_{\widehat{m},a_{\widehat{m}}} + th_{\widehat{m}}, \\ &h_{\widehat{m}+1}, \ldots, h_N\} = 0 \end{aligned} \tag{6.1}$$

form a complete homotopy for $T_{m,\widehat{m},a}$, the isolated (resp., nonsingular isolated) points of

$$\mathrm{Var}(f_1, \ldots, f_m, \ell_{m+1,a_{m+1}}, \ldots, \ell_{\widehat{m},a_{\widehat{m}}}, h_{\widehat{m}+1}, \ldots, h_N) \cap U.$$

Furthermore, $T_{m,\widehat{m}}$ is contained in $\bigcup_{a \in \mathcal{I}_{m,\widehat{m}}} T_{m,\widehat{m},a}$.

Remark 6.5. The homotopy $H^{\mathrm{parm}}_{m,\widehat{m},a}$ moves the linear slice defined by $h_{m+1} = \cdots = h_{\widehat{m}} = 0$ to the linear slice defined by $\ell_{m+1,a_{m+1}} = \cdots = \ell_{\widehat{m},a_{\widehat{m}}} = 0$.

The proof follows immediately from Theorem 5.1, since each homotopy at (6.1) is a parameter homotopy in the coefficients of the linear functions $h_{m+1}, \ldots, h_{\widehat{m}}$.

The procedure implied by Lemma 6.4 allows us to extend a solution for $f_1, \ldots, f_m$ into one for $f_1, \ldots, f_{\widehat{m}}$, $\widehat{m} > m$. The following lemma establishes the secondary step of regeneration that accomplishes this.

Lemma 6.6 (Incremental Product Decomposition)**.** *Adopt all the notation of Lemma* 6.4. *Further, let* $\mathbb{V}_i = \{V_i, V_{i,1}, \ldots, V_{i,d_i}\}$ *be a linear product decomposition for* f_i, $i = m+1, \ldots, \widehat{m}$, *and assume that each* g_i, $i = m+1, \ldots, \widehat{m}$, *is a generic product member of* V_i. *Then, the start set* $T_{m,\widehat{m}}$ *with the homotopy function*

$$H^{\text{prod}}_{m,\widehat{m}}(x,t) = \{f_1, \ldots, f_m, \\ (1-t)f_{m+1} + tg_{m+1}, \ldots, (1-t)f_{\widehat{m}} + tg_{\widehat{m}}, \\ h_{\widehat{m}+1}, \ldots, h_N\} = 0 \tag{6.2}$$

is a complete homotopy for $S_{\widehat{m}}$.

This lemma follows immediately from Theorem 5.3.

To apply Lemma 6.6, we need a linear product decomposition $V_{i,1} \otimes \cdots \otimes V_{i,d_i}$ for each f_i with $d_i \geq \deg f_i$, $i = m+1, \ldots, \widehat{m}$. (Usually we choose $d_i = \deg f_i$, but, for example, the trilinear quadratic $1 + xy + yz + zx$ admits the decomposition $\langle 1, x\rangle \otimes \langle 1, y\rangle \otimes \langle 1, z\rangle$, which might be useful in some instances.) We know that it is sufficient to choose each $V_{i,j}$ as the vector space whose elements are the support base f_i, but often some of the $V_{i,j}$ may omit some variables that appear in f_i and still suffice. For example, the polynomial $xy + 1$ admits the linear product decomposition $\langle x, 1\rangle \otimes \langle y, 1\rangle$, whereas its support base is $\{x, y, 1\}$.

6.3. **Extrinsic vs. intrinsic homotopy.** In our applications of Eq. (6.1) and Eq. (6.2), $h_{\widehat{m}+1}, \ldots, h_N$ are all linear functions that do not change during the path tracking from $t = 1$ to $t = 0$. We refer to this as the extrinsic regeneration homotopy. The intrinsic formulation, which is more efficient to use when $\widehat{m} \ll N$, proceeds by forming a linear basis for the $\widehat{m}$-dimensional linear space $\text{Var}(h_{\widehat{m}+1}, \ldots, h_N)$. That is, we use linear algebra once at the beginning of each incremental stage to find $A \in \mathbb{C}^{N \times \widehat{m}}$, $b \in \mathbb{C}^N$ such that $\operatorname{rank} A = \widehat{m}$ and $h_i(Au + b) = 0$ for all $u \in \mathbb{C}^{\widehat{m}}$, $i = \widehat{m}+1, \ldots, N$. Then, the homotopies $H(x,t) = 0$ (where H is either H^{parm} or H^{prod}) can be replaced by $\tilde{H}(u,t) = H(Au + b, t) = 0$, whereupon the linear functions are always zero and may be dropped. This reduces the number of functions to be tracked to $\widehat{m}$ instead of N. For efficiency, the polynomials should not be expanded, but evaluated in a straightline manner, e.g., evaluate $\phi = Au + b$ and then evaluate $H(\phi, t)$. When $\widehat{m}$ is not sufficiently small, the extra arithmetic in evaluating ϕ cancels out the savings of tracking on $\mathbb{C}^{\widehat{m}}$ instead of $\mathbb{C}^N$, so it is better to work extrinsically. When $\widehat{m}$ is small enough for the straightline intrinsic formulation to be advantageous, our software package Bertini [3] automatically invokes it.

6.4. **Full regeneration.** With Lemmas 6.4 and 6.6 in hand, it is straightforward to solve Problem 1. One merely specifies any set of strictly increasing integers ending at N, say $0 = m_0 < m_1 < \cdots < m_r = N$. Then, one solves r incremental problems for $(m, \widehat{m}) = (0, m_1), (m_1, m_2), \ldots, (m_{r-1}, N)$, using the isolated (or nonsingular) solutions of one incremental problem as the start points for the next incremental problem. To be most clear, we summarize the steps in pseudocode below.

Theorem 6.7 (Regeneration of Isolated Roots)**.** *Subject to genericity, the procedure* **Regenerate** *below solves Problem* 1*.*

The validity of each homotopy step in **Regenerate** is established by Lemmas 6.4 and 6.6. The only elaboration necessary is to observe that some of the endpoints of the homotopies (Eq. (6.1) and Eq. (6.2)) might lie on higher-dimensional sets, so these must be cast out to obtain just the set of isolated solutions needed for the subsequent homotopy. When it is needed, [1] gives a local dimension test that can differentiate between the isolated and nonisolated solutions. Without a local dimension test, we can only solve the more limited, but highly relevant, case of finding just the nonsingular solutions at each stage. These are easily identified by checking the rank of the Jacobian matrix of partial derivatives for each point. "Subject to genericity" acknowledges that the algorithm must make generic choices of coefficients in the linear functions $h_{m_1}, \ldots, h_N$, the linear functions that form the linear products $g_1, \ldots, g_N$, and generic choices required in any homotopy membership test.

Procedure $S = \textbf{Regenerate}(F, Y, \sigma)$**:**
Inputs: A set $F = \{f_1, \ldots, f_N\}$ of N polynomials on $\mathbb{C}^N$, a proper subset Y of $\mathbb{C}^N$ in a form suitable for a membership test, and $\sigma \in \{\text{True}, \text{False}\}$.
Output: When $\sigma = \text{True}$ (resp., when $\sigma = \text{False}$), the set S of all isolated (resp., nonsingular isolated) points in $\text{Var}(F) \cap X$, where $X = \mathbb{C}^N \setminus Y$.
Begin:

(1) Reorder the polynomials $f_1, \ldots, f_N$ in any advantageous order (see §6.5).
(2) Pick a set of $r+1$ strictly increasing integers starting at 0 and ending at N, say $0 = m_0 < m_1 < \cdots < m_r = N$.
(3) Specify a linear product decomposition $V_{i,1} \otimes \cdots \otimes V_{i,d_i}$ for each f_i, $i = 1, \ldots, N$. One alternative that always suffices is $d_i = \deg f_i$ with $V_{i,j}$, $j = 1, \ldots, d_i$, generated by the support base of f_i.
(4) Choose a generic linear form $\ell_{i,j}$ in each $V_{i,j}$, $i = 1, \ldots, N$, $j = 1, \ldots, d_i$. Let $g_i = \prod_{j=1}^{d_i} \ell_{i,j}$.
(5) For $i = 1, \ldots, N$, choose a generic linear h_i that supports all $\ell_{i,j}$, $j = 1, \ldots, d_i$.
(6) For $i = 1, \ldots, r$, let $(m, \widehat{m}) = (m_{i-1}, m_i)$, let

$$G_{m,\widehat{m}} = \{f_1, \ldots, f_m, g_{m+1}, \ldots, g_{\widehat{m}}, h_{\widehat{m}+1}, \ldots, h_N\}, \tag{6.3}$$

$$F_{\widehat{m}} = \{f_1, \ldots, f_{\widehat{m}}, h_{\widehat{m}+1}, \ldots, h_N\}, \tag{6.4}$$

and do the following:

(a) Solve for $T_{m,\widehat{m}}$, a superset of the set of isolated (resp., nonsingular isolated) points of $\text{Var}(G_{m,\widehat{m}}) \cap X$. There are two cases, as follows.

Case $m = 0$**:** Use numerical linear algebra to solve $G_{0,\widehat{m}}$. Since each g_i is a product of d_i linear factors, there are at most $D_{1,m_1} = \prod_{i=1}^{m_1} d_i$ solutions, all of which can be found by linear algebra. Since the linear factors may be sparse, there may be fewer than D_{1,m_1} solutions. The solution set is called T_{0,m_1}.

Otherwise: Use the homotopies $H^{\text{parm}}_{m,\widehat{m},a}$ from Eq. (6.1) with start set S_m. There is a homotopy for each $a \in \mathcal{I}_{m,\widehat{m}}$, and $T_{m,\widehat{m}}$ is the union of the solutions from all of these.

(b) Use a membership test to expunge any points of $T_{m,\widehat{m}}$ that are in Y.

(c) If $\sigma =$ True, use a local dimension test to expunge any singular points that are not isolated.

(d) If $\sigma =$ False, eliminate any singular points from $T_{m,\widehat{m}}$.

(e) Solve for $S_{\widehat{m}}$, a superset of the set of all isolated (resp., nonsingular isolated) points of $\text{Var}(F_{\widehat{m}}) \cap X$ using the product homotopy $H^{\text{prod}}_{m,\widehat{m}}$ from Eq. (6.2) with start solutions $T_{m,\widehat{m}}$.

(f) Use a membership test to expunge any points of $S_{\widehat{m}}$ that are in Y.

(g) If $\sigma =$ True, use a local dimension test to expunge any singular points of $S_{\widehat{m}}$ that are not isolated.

(h) If $\sigma =$ False, eliminate any singular points from $S_{\widehat{m}}$.

(i) Expunge any points from $S_{\widehat{m}}$ that solve any of $f_{\widehat{m}+1}, \ldots, f_N$.

Return $S = S_N$:

Step 6(i) is not necessary, but rather is a measure added to improve efficiency; it is justified because such points can never lead to isolated solutions of the final system.

If we are seeking all isolated roots (i.e., if $\sigma =$ True), then Steps 6(c) and 6(g) may require a local dimension test (see [1]). Without such a test, one can only reliably find all nonsingular isolated roots.

Steps (1)–(3) of the regeneration procedure allow many freedoms that can be used to adapt the procedure to more efficiently solve a given problem. We discuss how to use these freedoms in the next few paragraphs.

6.5. **Ordering of the functions.** At Step 1, one may choose to reorder the polynomials. In general, this changes the number of paths that need to be tracked. One way to attempt to minimize the number of paths is to minimize the maximum number of possible paths to track. Suppose we are working equation by equation (that is, $r = N$) and that the linear product decompositions have $d_j = \deg f_j$ factors. Then, the maximum number of paths to track is $p = d_1 + d_1 d_2 + \cdots + d_1 d_2 \cdots d_N$. By reordering the functions so that $d_1 \leq d_2 \leq \cdots \leq d_N$, the maximum number of paths p is minimized.

It is common that some endpoints at intermediate stages are cast out for lying on positive dimensional components or on the excluded set Y. In fact, it is to our advantage to arrange for this to happen as early and as often as possible. This goal may sometimes conflict with an ordering having monotonically increasing degrees. It is generally impossible to know ahead of time how the number of paths depends on the ordering, but one simple observation seems to help. When the functions are sparse, often only a subset of the variables appears in some equations. A good strategy is to order the functions so that the functions involving roughly the same collection of variables are introduced consecutively, thereby enforcing the maximum number of conditions on the variables involved in the functions.

When these two strategies are compatible, a good ordering of the polynomials is easily decided. (There may be more than one equally good ordering.) Unfortunately, we do not yet have good rules for picking an ordering when the strategies

conflict. We suggest first ordering by degree, and if some polynomials have the same degree, order them to minimize the rate of accumulation of the new variables. When neither of these criteria decides the ordering of some subgroup of the polynomials, our early experience indicates that the ordering within such a group has a minimal effect.

6.6. **Equation grouping.** At Step 2, one may choose how many polynomials to introduce at each stage. One far extreme is to choose $r = 1$, in which case we introduce all of the polynomials at once, resulting in only one stage of homotopy that is effectively a traditional linear product homotopy on the whole system. At the other extreme, one may choose $r = N$, which means $m_0, \ldots, m_r = 0, 1, 2, \ldots, N-1, N$. We call this "solving equation by equation", because only one new polynomial from F is introduced at each pass through the main loop. We often prefer to take this extreme, but sometimes equations appear in related subgroups that we elect to introduce group by group. The example in Section 9.4 has this character: the polynomials arise naturally as subsystems, each consisting of 2 polynomials. For that problem, introducing the equations two at a time results in fewer paths to track than an equation-by-equation approach.

Another consideration comes into play in an implementation on multiple parallel processors. The number of paths to track usually increases at each stage (often dramatically so), and if there are many processors available, it could happen that some of them sit idle in the early stages. To put this resource to best use, it may also be advantageous to introduce groups of equations in the early stages to make enough paths to keep all processors busy, then drop back to working equation by equation as the solution set increases in size.

6.7. **Choosing linear products.** The freedom to choose a linear product decomposition at Step 3 can have a noticeable effect. To illustrate this, there is a comparison between two different linear products in Section 9.2. One may also take advantage of multilinearity and other forms of sparseness and structure here. For example, regeneration was set up to exploit the 4-homogeneous structure and two-fold symmetry for the polynomial system presented in Section 9.3

7. Regeneration for witness sets

Let us now consider Problem 2, in which one seeks the witness sets for all solution components, both positive dimensional ones and the isolated roots. Recall that at the outset, a witness set is provided for an algebraic set $Z \setminus Y$, and we wish to update this to a witness set for $(Z \cap \mathrm{Var}(f)) \setminus Y$. It is sufficient to be able to do this when f is a single polynomial, because to address a system $f = \{f_1, \ldots, f_n\}$, one may repeat the procedure, introducing one new equation each time. Let $Z_k = (Z \cap \mathrm{Var}(f_1, \ldots, f_k)) \setminus Y$. The procedure **RegenWitness** (see below) for intersecting with a single polynomial generates a witness set for Z_k from one for Z_{k-1} (where $Z_0 = Z$). We use this to successively generate witness sets for Z_1 through Z_n.

Except where stated otherwise, in this section, we assume that f is a single polynomial on $\mathbb{C}^N$. Then, $\mathrm{Var}(f)$ is either $\mathbb{C}^N$ (if f is trivial) or it is a hypersurface. If a generic point w of an irreducible algebraic set X satisfies $f(w) = 0$, then $X \subset \mathrm{Var}(f)$ and so $X \cap \mathrm{Var}(f) = X$.

Recall that a witness set for an algebraic set Z is a collection of witness sets for each k-dimensional component Z_k, as in Eq. (2.2). A witness set W_k for $Z_k \subset \mathbb{C}^N$

is in the form $W_k = \{P, L_k, S\}$, where P is a polynomial system on $\mathbb{C}^N$ such that Z_k is one of its components, $L_k = \{h_1, \ldots, h_k\}$ is a set of k generic linear functions on $\mathbb{C}^N$, and S is the finite set of points $Z_k \cap \mathrm{Var}(L_k)$. In the following, we assume that the same linear functions h_i, $i = 1, \ldots, N$, are used in every witness set, that is, $L_1 = \{h_1\}$ for every one-dimensional component, $L_2 = \{h_1, h_2\}$ for every two-dimensional component, and so on. If one is given a composite witness set where this is not true, it can be made so by choosing a generic $L_N = \{h_1, \ldots, h_N\}$ and moving each of the given linear sets to the appropriate subset of L_N using parameter continuation.

The notation $w \in W = \{P, L, S\}$ means $w = \{P, L, s\}$, where $s \in S$. We use the shorthand $f(w)$ to mean the evaluation $f(s)$.

Procedure $\widehat{W} =$ RegenWitness(W, f, Y):
Inputs: A witness set W for a quasiprojective algebraic set $(Z \setminus Y) \subset \mathbb{C}^N$, where Y, Z are algebraic sets. A polynomial f on $\mathbb{C}^N$. A membership test for Y. Let W_k be the k-dimensional component of W.
Output: A witness set $\widehat{W}$ for $(Z \cap \mathrm{Var}(f)) \setminus Y$.
Begin:

(1) Initialize $\widehat{W}_k$, $k = 0, \ldots, N$, as empty.
(2) Specify a linear product decomposition $V_1 \otimes \cdots \otimes V_d$ for f. One alternative that always suffices is $d = \deg f$ with each V_i generated by the support base of f.
(3) Choose a generic linear form ℓ_i in each V_i, $i = 1, \ldots, d$. This means that ℓ_i is a linear function with generic coefficients. Let $g = \prod_{i=1}^{d} \ell_i$.
(4) For $k = N, N-1, \ldots, 0$, do the following:
  (a) For each $w = \{P, L_k, w^*\} \in W_k$ (P is a polynomial system, L_k is a set of k generic linear functions, and $w^* \in \mathrm{Var}(P, L_k)$) do the following:
    (i) Evaluate $e = f(w^*)$.
    (ii) If $e = 0$, then append $\{\{P, f\}, L_k, w^*\}$ to $\widehat{W}_k$.
    (iii) If $e \neq 0$ and $k = 0$, discard w.
    (iv) If $e \neq 0$ and $k > 0$, do the following:
      (A) If necessary, square down P to $N - k$ polynomials (see Eq. (2.3)). For simplicity, call the result P again.
      (B) For $i = 1, \ldots, d$, start at w and track the continuation path of

$$H_1(x,t) = \{P(x), h_1(x), \ldots, h_{k-1}(x), \; t h_k(x) + (1-t)\ell_i(x)\} = 0. \tag{7.1}$$

      Let T be the set of d endpoints of these paths.
      (C) Use a membership test to expunge any points of T that are in Y.
      (D) Use a local dimension test [1] to expunge any points of T that are not isolated points of $\mathrm{Var}(H_1(x, 0))$.
      (E) Track all paths starting at T for the homotopy function

$$H_2(x,t) = \{P(x), h_1(x), \ldots, h_{k-1}(x), \; t g(x) + (1-t) f(x)\} = 0. \tag{7.2}$$

      Call the set of endpoints S.

(F) Use a membership test to expunge any points of S that are in Y.
(G) Append $\{\{P, f\}, \{h_1, \ldots, h_{k-1}\}, S\}$ to $\widehat{W}_{k-1}$.

(b) Remove from $\widehat{W}_{k-1}$ any points on higher than $(k-1)$-dimensional sets of $\mathrm{Var}(P, f)$. This can be done either by homotopy membership tests for each $\widehat{W}_j$, $j > k-1$, or it can be done using a local dimension test [1].

Return $\widehat{W} = \{\widehat{W}_N, \ldots, \widehat{W}_0\}$**:**

Suppose we wish to find a witness set for $\mathrm{Var}(f_1, \ldots, f_n)\backslash Y$ starting from scratch. Letting $\mathcal{W}_k$ denote a witness set for $\mathrm{Var}(f_1, \ldots, f_k)\backslash Y$, we seek $\mathcal{W}_n$. With this notation, $\mathcal{W}_0$ is a witness set for $\mathbb{C}^N\backslash Y$ which has a single witness point $\mathrm{Var}(h_1, \ldots, h_N)$. To find the witness set $\mathcal{W}_n$, one may proceed by using the recursion

$$\mathcal{W}_k = \textbf{RegenWitness}(\mathcal{W}_{k-1}, f_k, Y) \quad \text{for} \quad k = 1, \ldots, n.$$

Problem 2 is solved similarly, beginning with $\mathcal{W}_0$ as the witness set for $Z \setminus Y$.

The justification of this procedure is similar to that for procedure **Regenerate**. One significant difference is that there is not an option to choose only components of multiplicity one. This is because the multiplicity of a component can decrease as new polynomials are introduced. In **Regenerate** this was not a problem, because we considered only the case $n = N$, which means that to get an isolated point at the end of the procedure, we only need the isolated points at the end of each stage. (Each new polynomial can either reduce multiplicity or reduce the dimension of a component, but not both, so if $n = N$, a component of multiplicity greater than 1 at an early stage can only lead to final components that are either multiplicity greater than 1 or dimension greater than zero.) Here, we are keeping track of solution sets at every dimension.

A second difference between **RegenWitness** and **Regenerate** is Steps 4(a)(ii)–(iii) in **RegenWitness**, which avoid the homotopies of Step 4(a)(iv) in certain circumstances. Step 4(a)(ii) recognizes when a component that a witness point lies on is contained in the hypersurface $\mathrm{Var}(f)$, and so the component will not be altered by intersection with $\mathrm{Var}(f)$. Hence, it can be sent to $\widehat{W}_k$ without alteration. This action corresponds to Step 6(i) of **Regenerate**, where such points are expunged because they cannot lead to isolated roots in the final output. Step 4(a)(iii) of **RegenWitness** eliminates any isolated point of Z that does not lie on $\mathrm{Var}(f)$. This is not relevant to the case $n = N$ in **Regenerate**, where isolated solutions cannot appear until the last equation is already used.

Finally, in Step 4(a)(iv)(D) and Step 4(b), we know the dimension of the sets that are relevant and remove points of the wrong dimension. For Step 4(a)(iv)(D), we only want isolated endpoints in $T_{m,\widehat{m}}$ (see Lemma 6.4). For Step 4(b), $\widehat{W}_{k-1}$ is the witness set for dimension $k-1$. Points eliminated in Step 4(b) are known as "junk points" [33, §13.6].

8. Diagonal homotopy

Like regeneration, an existing method, diagonal homotopy, can also solve systems incrementally. In §9, we compare regeneration to a slightly improved version of diagonal homotopy, which we describe here.

The extrinsic diagonal homotopy in [29] and its intrinsic reformulation in [30] compute the intersection of two components given by witness sets. Either of these

can be used as the core computational step in an equation-by-equation approach to solving Problems 1 and 2 [31]. In brief, to find $\mathrm{Var}(f) \cap \mathrm{Var}(g)$, diagonal homotopy finds a witness set for the composite system $\{f(x), g(y), x - y\}$. In the equation-by-equation approach, g is always a single polynomial, the one for the new equation being introduced.

The previous formulations of diagonal homotopy had an unnecessary randomization which adds cost to their implementation. Since we wish to compare regeneration to the best possible formulation of the diagonal homotopy, we give a revised formulation here that eliminates the unneeded randomization. We only present the equation-by-equation case, as that is the most relevant for the present comparison.

As in procedure **RegenWitness**, the main loop for equation-by-equation diagonal homotopy begins with $W_k = \{P, \{h_1, \ldots, h_k\}, X\}$, a witness set for a k-dimensional algebraic set $Z \subset \mathbb{C}^N$. If polynomial system P has cardinality more than $N - k$, we can randomly square it to $N - k$ polynomials without changing Z or W_k. Consequently, without loss of generality we may assume that P has exactly $N - k$ polynomials, say $P = \{p_1, \ldots, p_{N-k}\}$. Let Z be a component of $\mathrm{Var}(P)$ and let X be the set of isolated points in $Z \cap \mathrm{Var}(h_1, \ldots, h_k)$. We wish to introduce a single new polynomial f on $\mathbb{C}^N$ and find the witness set for $Z \cap \mathrm{Var}(f)$. As in **RegenWitness**, witness points $w \in W_k$ such that $f(w) = 0$ go directly to the output set $\widehat{W_k}$ with f added to their list of functions, while the others enter into a stage of diagonal homotopy. The output of the diagonal homotopy will be witness points for the next dimension down, W_{k-1}, cut out by $\{h_1, \ldots, h_{k-1}\}$.

To start the diagonal homotopy, we begin by finding a witness set R for $\mathrm{Var}(f)$, which consists of the points $\mathrm{Var}(f, h_1, \ldots, h_{N-1})$. Notice that R, W_k, and W_{k-1} all lie in the linear space $\Phi = \mathrm{Var}(h_1, \ldots, h_{k-1})$, which is $(N-k+1)$-dimensional. An extrinsic form of diagonal homotopy works on a doubled set of variables $(x, y) \in \mathbb{C}^N \times \mathbb{C}^N$, which are forced to satisfy the conditions $P(x) = 0$, $f(y) = 0$, and $x - y = 0$. In a manner similar to §6.3, we can keep the number of variables smaller by using $N-k+1$ intrinsic variables for Φ, restricting our computations to $(x, y) \in \Phi \times \Phi$. Since there is an independent set of intrinsic variables for each copy of Φ, this would result in $2(N - k + 1)$ variables altogether. We call a formulation of the diagonal homotopy that works on $\Phi \times \Phi$ "semi-intrinsic". An even more efficient approach is to work fully intrinsically on a linear space of just dimension $N - k + 1$, as we describe next.

The fully intrinsic approach works along the lines of [30], which also takes advantage of the linearity of the diagonal equations $x - y = 0$. For convenience, let $\kappa = N - k$. We compute a basis for Φ, say $\phi(u) = Au + b$, where $u \in \mathbb{C}^{\kappa+1}$ is a set of intrinsic variables and $A \in \mathbb{C}^{N \times (\kappa+1)}$ and $b \in \mathbb{C}^N$ are constants. Let $\widehat{\phi} : \mathbb{C}^{\kappa+1} \to \mathbb{C}^{2N}$ be defined as $\widehat{\phi}(u) = (\phi(u), \phi(u))$. Thus $\widehat{\phi}(u)$ is a parameterization of the target linear space of the diagonal homotopy. At the beginning of the diagonal homotopy, the witness points for Z and $\mathrm{Var}(f)$ are cut out by smaller linear spaces of dimensions κ and 1, respectively. We use parameterizations of these as well. Specifically, let $\psi_1(u) = A_1 u_{1:\kappa} + b_1$ parameterize $\Psi_1 = \mathrm{Var}(h_1, \ldots, h_k)$ and let $\psi_2(u) = A_2 u_{\kappa+1} + b_2$ parameterize $\Psi_2 = \mathrm{Var}(h_1, \ldots, h_{N-1})$, where $A_1 \in \mathbb{C}^{N \times \kappa}$, $A_2 \in \mathbb{C}^{N \times 1}$, and $b_1, b_2 \in \mathbb{C}^N$. From these, we form the linear map $\psi : \mathbb{C}^{\kappa+1} \to \mathbb{C}^{2N}$ as $\psi(u) = (\psi_1(u), \psi_2(u))$. Note that any point $(x, y) \in \Psi_1 \times \Psi_2$ has a unique preimage $u = \psi^{-1}(x, y) \in \mathbb{C}^{\kappa+1}$. Finally, let $\pi_1, \pi_2 : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}^N$ be projections defined as $\pi_1 : (u, v) \mapsto u$ and $\pi_2 : (u, v) \mapsto v$. Then, a fully intrinsic

homotopy function $H_2 : \mathbb{C}^{\kappa+1} \times \mathbb{C} \to \mathbb{C}^{\kappa+1}$ is

$$H_2(u,t) = \{P\big(\pi_1[(1-t)\widehat{\phi}(u) + \gamma t\psi(u)]\big), \\ f\big(\pi_2[(1-t)\widehat{\phi}(u) + \gamma t\psi(u)]\big)\} = 0, \tag{8.1}$$

where $\gamma \in \mathbb{S}^1$ is chosen generically. The set of starting points is $S' = \psi^{-1}(R \times W_k)$ and among the path endpoints, each isolated point of $H_2(u,0)$, say u^*, gives a witness point $\{\{P,f\},\{h_1,\ldots,h_{k-1}\},\phi(u^*)\}$ that is appended to $\widehat{W}_{k-1}$. This is the version of diagonal homotopy that we use in our experiments below.

As in the procedure **RegenWitness**, the singular endpoints of the diagonal homotopy paths must be checked for possible membership in higher-dimensional components. Since we work down dimension-by-dimension, we always have on hand witness sets for the higher-dimensional components, so it is feasible to perform homotopy membership tests. A local dimension test can be used as an alternative.

9. Computational experiments

Regeneration is implemented in the software package Bertini [3]. All the examples discussed here were run on an 2.4 GHz Opteron 250 processor with 64-bit Linux. The parallel examples were run on a cluster consisting of a manager that uses one core of a Xeon 5410 processor and 8 computing nodes, each containing two 2.33 GHz quad-core Xeon 5410 processors running 64-bit Linux, i.e., one manager and 64 workers. PHCpack v2.3.39 [34] and HOM4PS-2.0.15 [14] were used in the examples described below.

9.1. **Illustrative example.** To demonstrate the regeneration approach described in this paper to find nonsingular isolated solutions, consider the following system used in [25, 31]:

$$f(x,y,z) = \begin{bmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{bmatrix} = \begin{bmatrix} (y-x^2)(x^2+y^2+z^2-1)(x-0.5) \\ (z-x^3)(x^2+y^2+z^2-1)(y-0.5) \\ (y-x^2)(z-x^3)(x^2+y^2+z^2-1)(z-0.5) \end{bmatrix}.$$

As this is a square system and we seek only nonsingular isolated solutions, we apply algorithm **Regenerate**. Since f_1 has degree 5, 5 paths are tracked using $H_{0,1}^{\text{prod}}$ defined by Eq. (6.2) to compute solutions of F_1 defined by Eq. (6.4). All of these solutions are nonsingular, but 4 satisfy either f_2 or f_3 and are removed at Step 6(i), leaving only one point, which corresponds to $x = 0.5$.

Since f_2 has degree 6, we take the linear slice $h_2 = l_{2,1}$ and regenerate to the linear slices $l_{2,2},\ldots,l_{2,6}$ by tracking 1 path for each $H_{1,2,i}^{\text{parm}}$ defined by Eq. (6.1), $i = 2,\ldots,6$. This creates 6 paths that need to be tracked using $H_{1,2}^{\text{prod}}$ to compute solutions of F_2. Of these paths, 2 diverge, 3 have endpoints that satisfy f_3, and the remaining endpoint corresponds to $x = y = 0.5$.

Since f_3 has degree 8, we take the linear slice $h_3 = l_{3,1}$ and regenerate to the linear slices $l_{3,2},\ldots,l_{3,8}$ by tracking 1 path for each $H_{2,3,i}^{\text{parm}}$, $i = 2,\ldots,8$. This creates 8 paths that need to be tracked using $H_{2,3}^{\text{prod}}$ to compute solutions of $F_3 = f$. Of these paths, 4 diverge, 3 lead to singular endpoints which lie on positive dimensional components, and the remaining endpoint corresponds to the only isolated solution $(0.5, 0.5, 0.5)$.

TABLE 1. Comparison for solving the general 6R, serial-link robot system using different solving methods

	total degree		diagonal		regeneration (generic lin. prod.)		
setup	paths tracked	time, s	paths tracked	time, s	H^{prod} paths tracked	H^{parm} paths tracked	time, s
order A	1024	92.23	649	78.84	628	313	20.29
order B	1024	92.23	949	90.48	928	463	22.97

9.2. **A comparison of the methods.** To compare the diagonal and regeneration equation-by-equation approaches, consider a polynomial system arising from the inverse kinematics problem of general six-revolute, serial-link robots described in [38], [33, §9.4]. The polynomial system, available at [3], consists of 2 linear and 10 quadratic polynomials in 12 variables. The system was set up using random parameter values and has 16 nonsingular finite isolated solutions.

As discussed in §6.5, the ordering of the quadratic polynomials can affect the total number of paths that need to be tracked. In this problem, the 12 variables correspond to the entries of 4 vectors in $\mathbb{C}^3$. Four of the quadratics correspond to normalizing each of these vectors to unit length. The other six quadratics provide conditions on the interaction between two or more vectors. In ordering the equations, it is clear that the two linear equations should come first, but after that the best ordering of the quadratics may not be obvious. To illustrate the impact of different orderings, we tried two of them: in "order A", the four normalizing quadratics were placed last; in "order B", they were placed immediately after the linear equations. We experimented with other orderings of the quadratics and found none whose number of paths was outside the range established by orders A and B.

These two orderings of the polynomial system were solved using a generic total degree homotopy, the diagonal approach, and the regeneration approach using a generic linear product decomposition; i.e., each linear factor has a support set consisting of the 12 variables and 1. Since the diagonal approach intersects witness sets which utilize generic linear spaces, it is natural to compare it with the regeneration approach using a generic linear product decomposition. These methods were run using adaptive precision [4, 5] with tracking tolerance of 10^{-6} and a final tolerance of 10^{-10}. The results are summarized in Table 1, which shows a difference in both time and number of paths tracked between the two orderings. However, the choice of method is seen to be more significant than the choice of ordering.

Table 2 presents the regeneration approach using a linear product decomposition consisting of minimal support hyperplanes using the two orderings. For "order A", the minimal support linear product significantly reduces the number of paths tracked and the time needed compared with the generic linear product. For "order B", even though the same number of paths are tracked by both linear product decompositions, the sparseness in the linear algebra computations using the minimal support accounts for the slight reduction in time compared with the generic linear product.

The problems were first run using the diagonal approach implemented in PHCpack [34]. Various settings were tried, including the default settings, changing the tracking tolerances and changing the order of the endgame, but PHCpack was only able to produce at most 15 of the 16 solutions, taking at least 3 minutes for each

TABLE 2. Summary for solving the general 6R, serial-link robot system using regeneration with a minimal support linear product decomposition

	regeneration (minimal lin. prod.)		
setup	H^{prod} paths tracked	H^{parm} paths tracked	time, s
order A	272	135	8.49
order B	928	463	21.61

TABLE 3. Comparison of regeneration and polyhedral homotopy for the nine-point path synthesis problem

method	regeneration		polyhedral
paths tracked	H^{prod}	136,296	87,639
	H^{parm}	66,888	
time, hrs	8.065		11.656

attempt. The diagonal implementation in Bertini consistently found all 16 solutions and always ran faster than PHCpack. The times reported in Table 1 are those for the Bertini implementation.

9.3. **A multivariate system from robotics.** One benefit of equation-by-equation methods is their ability to numerically discover structure in a problem to reduce the number of paths that need to be tracked as demonstrated in [31] and the examples above. One advantage of regeneration over the diagonal approach is the ability to easily incorporate the known structure of the problem into the regeneration homotopies to help further reduce the total number of paths that need to be tracked.

Consider the nine-point path synthesis problem for four-bar linkages. The original formulation of Roth and Freudenstein [24] consists of 8 seven-degree polynomials in 8 variables with a natural two-fold symmetry. Utilizing the symmetry, the results of [39] show that this system has 4326 nondegenerate solutions appearing in 1442 cognate triples.

The Bézout count for the Roth and Freudenstein system utilizing the 4-homogeneous structure and two-fold symmetry is $322,560$. Regeneration is easily set up to exploit both the symmetry and multi-homogeneous structure. The regeneration method finds the 4326 nondegenerate solutions by tracking $136,296$ paths using H^{prod} homotopies and tracking $66,888$ paths using H^{parm} homotopies. Table 3 compares solving this system with regeneration utilizing adaptive precision [4, 5] and a polyhedral homotopy using HOM4PS-2.0 [14].

We also used our parallel processor to solve this problem using regeneration. This decreased the computation time to 7.785 minutes.

Regeneration could also be adapted to use the product decomposition structure of this system that is described in [20]. Although we did not test this option, it promises to further reduce both the total number of paths tracked and the time.

9.4. **A large sparse polynomial system.** Equation-by-equation methods can be used to solve large polynomial systems when other methods are impractical.

TABLE 4. Comparison of various methods for solving systems related to the Lotka-Volterra population model

	total degree	2-homogeneous	polyhedral	regeneration	
n	paths	paths	paths	H^{prod} paths	H^{parm} paths
1	256	70	16	60	42
2	65,536	12,870	256	1020	762
3	16,777,216	2,704,156	4096	16,380	12,282
4	4,294,967,296	601,080,390	65,536	262,140	196,602
5	1,099,511,627,776	137,846,528,820	1,048,576	4,194,300	3,145,722

To illustrate this, consider a sparse polynomial system arising from ongoing research by the first two authors and Bei Hu (University of Notre Dame) related to the discretization of the stationary Lotka-Volterra population model with diffusion [16, 37].

Let $n \in \mathbb{N}$. For $1 \le i \le n$ and $1 \le j \le 4$, define

$$\begin{aligned} f_{ij} &= \frac{1}{25}\left(u_{i+1,j} - 2u_{i,j} + u_{i-1,j}\right) \\ &\quad + \frac{1}{(n+1)^2}\left(u_{i,j+1} - 2u_{i,j} + u_{i,j-1}\right) + \frac{1}{25(n+1)^2}u_{i,j}\left(1 - v_{i,j}\right), \\ g_{ij} &= \frac{1}{25}\left(v_{i+1,j} - 2v_{i,j} + v_{i-1,j}\right) \\ &\quad + \frac{1}{(n+1)^2}\left(v_{i,j+1} - 2v_{i,j} + v_{i,j-1}\right) + \frac{1}{25(n+1)^2}v_{i,j}\left(u_{i,j} - 1\right) \end{aligned} \tag{9.1}$$

with $u_{0,j} = v_{0,j} = u_{n+1,j} = v_{n+1,j} = u_{i,0} = v_{i,0} = u_{i,5} = v_{i,5} = 0$.

These systems consist of $8n$ quadratic polynomials in $8n$ variables and have 2^{4n} nonsingular isolated solutions. The system has a natural 2-homogeneous structure, with each polynomial being of type $(1,1)$, but as shown in Table 4, the number of paths grows too quickly with n to consider using a traditional 2-homogeneous homotopy for n larger than 3. The mixed volume of the system is the same as the number of solutions, 2^{4n}, but even so, current implementations of the polyhedral method failed to solve the system in less than 45 days for $n = 5$.

To solve the system using regeneration, we used the natural ordering of the equations and introduced the equation two at a time as suggested by Eq. (9.1). The linear product decomposition of the polynomials used was:

$$\begin{aligned} f_{ij} &\in \langle\{1, u_{i+1,j}, u_{i,j}, u_{i-1,j}, u_{i,j+1}, u_{i,j-1}, v_{i,j}\} \times \{1, v_{i,j}\}\rangle, \\ g_{ij} &\in \langle\{1, v_{i+1,j}, v_{i,j}, v_{i-1,j}, v_{i,j+1}, v_{i,j-1}, u_{i,j}\} \times \{1, u_{i,j}\}\rangle. \end{aligned} \tag{9.2}$$

This decomposition was used so that the elements of the first vector space in f_{ij} form its support base, and similarly in g_{ij}. Furthermore, in **Regenerate**, the generic linear selected in Step 5 was taken as the first linear selected in Step 4. This means that in Step 6(a) of **Regenerate**, one-fourth of the H^{parm} homotopies are trivial as they start and end at the same linear coefficients. Thus, these homotopies require no computation.

With the above choices, regeneration tracks roughly 4 times as many paths as the number of solutions. Table 4 compares the number of paths for various methods and Table 5 contains timings for the various software packages. For $n \le 4$, regeneration can solve the system using only double precision, and for $n = 5$, regeneration utilized adaptive precision tracking [4, 5] to track the paths since double precision was not

TABLE 5. Single-processor timings for the polyhedral method and regeneration for solving systems related to the Lotka-Volterra population model

n	PHC polyhedral	HOM4PS-2.0 polyhedral	Bertini regeneration
1	0.56s	0.06s	0.34s
2	4m57s	7.33s	17.30s
3	18d10h18m56s	9m32s	10m3s
4	-	3d8h28m30s	5h5m50s
5	-	-	6d10h32m12s

adequate for some of the paths. Using the current implementation of regeneration in parallel, it took 7.28 minutes for $n = 4$ and 3.63 hours for $n = 5$.

10. Conclusions

Regeneration builds up the solution set of a polynomial system equation-by-equation or subsystem-by-subsystem using a sequence of parameter and linear product homotopies. An existing method, diagonal homotopy, also allows this sort of incremental solution of a system. By revealing the structure of the solution sets of subsets of the polynomials in the system, these incremental methods eliminate paths in the later, more expensive stages of homotopy. This tends to save overall computation. We compare the new regeneration algorithm with both the diagonal homotopy and with polyhedral homotopy, considered the most efficient nonincremental way to solve sparse polynomial systems. Our tests show that regeneration is on average better than diagonal homotopy. For small systems, polyhedral homotopy is often the best, but for large systems, regeneration takes fewer computations. The mixed volume computations used to create the start system for a polyhedral homotopy are combinatorial in nature. It appears that regeneration reveals much of the same sparse structure without a mixed volume computation, giving it the edge in large problems.

Appendix A. Theory

In this appendix we finish the proof of Theorem 5.3 by proving the remaining case when $k = 1$.

Before we start the proof of this result, we need to recall some standard definitions and notation. We refer to [33] for more details about algebraic sets.

Recall that a quasiprojective algebraic set is a Zariski open set U of a projective algebraic set X. All algebraic sets are complex and quasiprojective. An algebraic set in which all of whose irreducible components are dimension one is called an algebraic curve. By an algebraic function on U, we mean a rational function which is holomorphic, e.g., if U is a closed algebraic subset of $\mathbb{C}^N$, these are restrictions of polynomials from $\mathbb{C}^N$.

We denote the singular set of the reduction of an algebraic set X by $\mathrm{Sing}(X)$.

Given a vector space V of algebraic functions on an algebraic set X, the base locus of V (denoted $\mathrm{Bs}(V)$) is the set of common zeros of V on X. $\mathrm{Bs}(V)$ is an algebraic subset of X. If $\mathrm{Bs}(V)$ is the empty set, we say that V is basepoint free. Note that basepoint freeness is exactly the condition that we have assumed for the V_i on U in Theorem 5.2.1 and Lemma 5.2.2. Bertini's Theorem, e.g., see [33] for an

extensive discussion, guarantees that the solution set $\mathrm{Var}(g)$ of a general element g of V has various strong properties outside $\mathrm{Bs}(V)$; e.g., the intersection of $\mathrm{Var}(g)$ with the reduction of $X \setminus (\mathrm{Bs}(V) \cup \mathrm{Sing}(X))$ is smooth.

Now let us turn to Theorem 5.3 in the case $k = 1$. In that case we showed that it was sufficient to prove the result to compute the isolated singular points in the case when f_1 is generic. As noted in the discussion of Theorem 5.3, the result for nonsingular isolated solutions is the main theorem in [20].

If $N = 1$, the result is trivial. Indeed, we have a polynomial f which is a random sum of products of polynomials chosen from vector spaces of polynomials $V_{1,1}, \ldots, V_{1_{d_1}}$. Any general product g_1 of general elements $p_{1,j}$ will have degree equal to the degree of f, which is the easiest case of the most classical homotopy solution result for solving a polynomial system.

The difficulty when $N > 1$ comes from having to deal with the possible presence of positive dimensional components of the solution set of $f(x) = 0$ and the possibility of solution components at infinity.

The key is to think of $\mathrm{Var}(f)$ as $\mathrm{Var}(f_1(x)) \cap \mathrm{Var}(\{f_2(x), \ldots, f_N(x)\})$. For any irreducible component Z of $\mathrm{Var}(\{f_2(x), \ldots, f_N(x)\})$, it follows (see [33, Theorem 12.2.2]) that $\dim Z \geq 1$ and that either $Z \subset \mathrm{Var}(f_1(x))$ or $\dim Z \cap \mathrm{Var}(f_1(x)) = \dim Z - 1$. From this we conclude that the only components Z that can lead to isolated solutions of $\mathrm{Var}(f)$ are those of dimension one.

Except for some technical details, restricting to components Z of dimension one is pretty much the same as the case when $N = 1$. Indeed, the degree computation when $N = 1$ translates directly into the basic statement that the degree of a tensor product $L_1 \otimes \cdots \otimes L_m$ of line bundles on a smooth compact Riemann surface equals the sum of the degrees $\deg(L_1) + \cdots + \deg(L_m)$. The technical details are that Z might be singular, it is noncompact, and there might be some points in the complement of Z in its closure in projective space, where all the functions in one or more of the V_i of Theorem 5.3 are identically zero.

Since we need to show that given the generic product $g_1(x)$,

$$\#_{\mathrm{isol}}(g_1, U) \geq \#_{\mathrm{isol}}(f, U),$$

we can replace Z with $Z \cap U$ with no loss of generality. Moreover since we are counting isolated points without multiplicities, the nonreduced structure of Z is irrelevant; i.e., we may work on the reduction of Z.

Thus we have reduced to the following theorem.

Theorem A.1. *Let Z be a reduced algebraic curve. For i from 1 to m, let V_i be a finite vector space of algebraic functions on Z. Assume that each V_i is basepoint free. Let $f \in V$, where V is the image of $V_1 \otimes \cdots \otimes V_m$ in the space of algebraic functions on Z. Choose general elements $g_i \in V_i$ and let g denote the element in V that is the image of $g_1 \otimes \cdots \otimes g_m$. Then $\#_{\mathrm{isol}}(g, Z) \geq \#_{\mathrm{isol}}(f, Z)$.*

Proof. Without loss of generality we may assume that Z is smooth. To see this, let $\pi : Z' \to Z$ denote the desingularization of Z; i.e., π is a holomorphic finite-to-one map from a nonsingular algebraic curve Z' onto Z that maps $Z' \setminus \pi^{-1}(\mathrm{Sing}(Z))$ isomorphically onto $Z \setminus \mathrm{Sing}(Z)$. Let V' denote the vector space of algebraic functions induced by composition of functions in V with π. For each i from 1 to m, let V_i' denote the vector space of algebraic functions induced by composition of functions in V_i with π. The space of functions V' is isomorphic to the image of $V_1' \otimes \cdots \otimes V_m'$

in the space of algebraic functions on Z'. Note by Bertini's Theorem that the solutions of the g_i and g in Z are nonsingular and thus miss $\mathrm{Sing}(Z)$. Therefore we may identify the solutions of g (respectively g_i for i from 1 to m) on Z with the solutions of $g' = g \circ \pi$ (respectively $g'_i = g_i \circ \pi$ for i from 1 to m) on Z', and these points are also nonsingular. Thus we have $\#_{\mathrm{isol}}(g, Z) = \#_{\mathrm{isol}}(g', Z')$. Also since $\#_{\mathrm{isol}}(f, Z) \leq \#_{\mathrm{isol}}(f', Z)$, it follows that $\#_{\mathrm{isol}}(g', Z) \geq \#_{\mathrm{isol}}(f', Z)$ implies that $\#_{\mathrm{isol}}(g, Z) \geq \#_{\mathrm{isol}}(f, Z)$. Thus Z may be assumed to be smooth.

Let K be the unique smooth compact algebraic curve that contains Z as a Zariski open dense set. All the functions in V_i for i from 1 to m extend to rational functions on K. We regard V_i as a space of rational functions, which are holomorphic on Z. For each i, all of the functions in a nonempty Zariski open set of V_i have the same pole set D_i with multiplicities, i.e., $D_i = \sum_{j \in \mathcal{J}_i} \mu_{i,j} k_{i,j}$ with all of the points $k_{i,j}$ corresponding to poles lying in the finite set $K \setminus Z$ and with $\mu_{i,j}$ being the order of the pole $k_{i,j}$. Finite formal sums of integers times points of a curve are called divisors. Let L_i denote the algebraic line bundle associated to the divisor D_i, and let p_i denote the tautological section of L_i (unique up to multiplication by a nonzero constant) vanishing precisely at D_i. Sending $h \in V_i$ to $h \cdot p_i$ gives an isomorphism of V_i with a vector space of algebraic sections of L_i which we also denote by V_i. Similarly, sending the functions $h \in V$ to $h \cdot p_1 \cdots p_m$ gives an isomorphism of V with a vector space of algebraic sections of $L = L_1 \otimes \cdots \otimes L_m$, which we also denote by V.

All algebraic sections of an algebraic line bundle $\mathcal{L}$ on a compact smooth curve have the same number (counting multiplicities) of zeroes, and this number is denoted $\deg(\mathcal{L})$. We have $\deg(L) = \deg(L_1) + \cdots + \deg(L_m)$.

Let B_i denote the base locus of the elements of V_i; i.e., let B_i be the divisor $n_{i,1} z_{i,1} + \cdots + n_{i,k_i} z_{i,k_i}$, where $z_{i,1}, \ldots, z_{i,k_i}$ are the points on K where all the elements v_i of V_i are zero and where for each j, $n_{i,j}$ is the minimum over the $v_i \in V_i$ of the multiplicities $z_{i,j}$ as a solution of $v_i(x) = 0$. Note that a general element $g_i \in V_i$ vanishes at the $z_{i,j}$ with multiplicity exactly $n_{i,j}$ for each j from 1 to k_i. Thus by Bertini's Theorem, a general g_i has zero set consisting of $n_{i,1} z_{i,1} + \cdots + n_{i,k_i} z_{i,k_i}$ plus a set of nonsingular points with empty intersection with the set $(K \setminus Z) \cup \mathrm{Var}(f)$. In particular, for a general product g,

$$\#_{\mathrm{isol}}(g, Z) = \deg(L) - \sum_{i=1}^{m} \sum_{j=1}^{k_i} n_{i,j}.$$

Since f is in V, it vanishes on the points $z_{i,j}$ to multiplicities greater than or equal to the multiplicities of g at the points. Thus

$$\#_{\mathrm{isol}}(f, Z) \leq \deg(L) - \sum_{i=1}^{m} \sum_{j=1}^{k_i} n_{i,j}.$$

Putting these inequalities together we have $\#_{\mathrm{isol}}(g, Z) \geq \#_{\mathrm{isol}}(f, Z)$, which proves the theorem. □

References

1. D.J. Bates, J.D. Hauenstein, C. Peterson, and A.J. Sommese, A numerical local dimensions test for points on the solution set of a system of polynomial equations, *SIAM J. Numer. Anal.*, 47 (2009), pp. 3608–3623. MR2576513

2. D.J. Bates, C. Peterson, and A.J. Sommese, A numerical-symbolic algorithm for computing the multiplicity of a component of an algebraic set, *J. Complexity*, 22 (2006), pp. 475–489. MR2246892 (2007c:14067)
3. D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler, Bertini: Software for Numerical Algebraic Geometry. Available at www.nd.edu/∼sommese/bertini.
4. D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler, Adaptive multiprecision path tracking, *SIAM J. Numer. Anal.*, 46 (2008), pp. 722–746. MR2383209 (2008m:65147)
5. D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler, Stepsize control for path tracking, in *Interactions of Classical and Numerical Algebraic Geometry*, D. Bates, G. Besana, S. Di Rocco, and C. Wampler (eds.), *Contemporary Mathematics* 496 (2009), 21–31. MR2555948
6. M. Beltrametti and A.J. Sommese, *The adjunction theory of complex projective varieties*, Expositions in Mathematics, 16, Walter De Gruyter, Berlin, 1995. MR1318687 (96f:14004)
7. B. Dayton and Z. Zeng, Computing the multiplicity structure in solving polynomial systems, in Proceedings of ISSAC 2005, ACM, New York, 2005, pp. 116–123. MR2280537
8. W. Fulton, *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*, Springer-Verlag, Berlin, second edition, 1998. MR1644323 (99d:14003)
9. B. Huber and B. Sturmfels, A polyhedral method for solving sparse polynomial systems, *Math. Comp.*, 64(212) (1995), pp. 1541–1555. MR1297471 (95m:65100)
10. G. Lecerf, Quadratic Newton iteration for systems with multiplicity, *Found. Comput. Math.*, 2(3) (2002), pp. 247–293. MR1907381 (2003f:65090)
11. A. Leykin, Numerical primary decomposition, in Symbolic and Algebraic Computation, Proceedings of ISSAC 2008, Linz/Hagenberg, Austria, July 20-23, 2008. Edited by J.R. Sendra and L. González-Vega, ACM, 2008. MR2513502
12. A. Leykin, J. Verschelde, and A. Zhao, Newton's method with deflation for isolated singularities of polynomial systems, *Theor. Comp. Sci.* 359 (2006), pp. 111–122. MR2251604 (2007k:65083)
13. A. Leykin, J. Verschelde, and A. Zhao, Higher-order deflation for polynomial systems with isolated singular solutions, in *IMA Volume 146: Algorithms in Algebraic Geometry*, A. Dickenstein, F.-O. Schreyer, and A.J. Sommese, eds., Springer, New York, 2008, pp. 79–97. MR2397938 (2009f:65130)
14. T.-L. Lee, T.Y. Li, and C.-H. Tsai, HOM4PS-2.0, Solving Polynomial Systems by the Polyhedral Homotopy Method. Software available at www.math.msu.edu/∼li.
15. T.Y. Li, Numerical solution of polynomial systems by homotopy continuation methods, in *Handbook of Numerical Analysis. Volume XI. Special Volume: Found. Comp. Math.*, edited by F. Cucker, North-Holland, 2003, pp. 209–304. MR2009773 (2004k:65089)
16. A.J. Lotka, Undamped oscillations derived from the laws of mass action, *J. Amer. Chem. Soc.* 42 (1920), pp. 1595–1599.
17. A.P. Morgan, A transformation to avoid solutions at infinity for polynomial systems, *Appl. Math. Comput.*, 18(1) (1986), pp. 77–86. MR815773 (87c:90193)
18. A.P. Morgan and A.J. Sommese, A homotopy for solving general polynomial systems that respects m-homogeneous structures, *Appl. Math. Comput.*, 24(2) (1987), pp. 101–113. MR914806 (88j:65110)
19. A.P. Morgan and A.J. Sommese, Coefficient-parameter polynomial continuation, *Appl. Math. Comput.*, 29(2) (1989), pp. 123–160. Errata: *Appl. Math. Comput.*, 51 (1992), p. 207. MR977815 (90f:58018); MR 1180600
20. A.P. Morgan, A.J. Sommese, and C.W. Wampler, A product-decomposition bound for Bezout numbers, *SIAM J. Numer. Anal.* 32(4) (1995), pp. 1308–1325. MR1342295 (96i:65040)
21. D. Mumford, *Algebraic geometry I: Complex projective varieties*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1976 edition. MR1344216 (96d:14001)
22. T. Ojika, Modified deflation algorithm for the solution of singular problems. I. A system of nonlinear algebraic equations, *J. Math. Anal. Appl.* 123 (1987), pp. 199–221. MR881541 (88f:65085)
23. 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. MR719330 (85a:65083)

24. B. Roth and F. Freudenstein, Synthesis of Path-Generating Mechanisms by Numerical Methods, *ASME J. Engrg. Industry*, 85B(3) (1963), pp. 298–306.
25. A.J. Sommese, J. Verschelde and C.W. Wampler, Numerical decomposition of the solution sets of polynomial systems into irreducible components, *SIAM J. Numer. Anal.* 38(6) (2001), pp. 2022–2046. MR1856241 (2002g:65064)
26. A.J. Sommese, J. Verschelde, and C.W. Wampler, Using monodromy to decompose solution sets of polynomial systems into irreducible components. In *Application of Algebraic Geometry to Coding Theory, Physics and Computation*, edited by C. Ciliberto, F. Hirzebruch, R. Miranda, and M. Teicher. Proceedings of a NATO Conference, February 25 - March 1, 2001, Eilat, Israel, pages 297–315, Kluwer Academic Publishers. MR1866906 (2002k:65087)
27. A.J. Sommese, J. Verschelde and C.W. Wampler, Symmetric functions applied to decomposing solution sets of polynomial systems, *SIAM J. Numer. Anal.* 40(6) (2002), pp. 2026–2046. MR1974173 (2004m:65069)
28. A.J. Sommese, J. Verschelde, and C.W. Wampler, Numerical irreducible decomposition using PHCpack, in *Algebra, Geometry, and Software Systems*, M. Joswig and N. Takayama, eds., Springer-Verlag, 2003, pp. 109–130. MR2011756 (2004i:65047)
29. A.J. Sommese, J. Verschelde, and C.W. Wampler, Homotopies for Intersecting Solution Components of Polynomial Systems, *SIAM J. Numer. Anal.* 42(4) (2004), pp. 1552–1571. MR2114290 (2005m:65111)
30. A.J. Sommese, J. Verschelde, and C.W. Wampler, An intrinsic homotopy for intersecting algebraic varieties, *J. Complexity*, 21 (2005), pp. 593–608. MR2152724 (2006f:65057)
31. A.J. Sommese, J. Verschelde, and C.W. Wampler, Solving Polynomial Systems Equation by Equation, in *IMA Volume 146: Algorithms in Algebraic Geometry*, A. Dickenstein, F.-O. Schreyer, and A.J. Sommese, eds., Springer-Verlag, 2007, pp. 133–152. MR2397941 (2009g:65075)
32. A.J. Sommese and C.W. Wampler, Numerical algebraic geometry. in *The Mathematics of Numerical Analysis*, edited by J. Renegar, M. Shub, and S. Smale, volume 32 of *Lectures in Applied Mathematics*, 1996, pp. 749–763. Proceedings of the AMS-SIAM Summer Seminar in Applied Mathematics, Park City, Utah, July 17-August 11, 1995, Park City, Utah. MR1421365 (98d:14079)
33. A.J. Sommese and C.W. Wampler, *The numerical solution of systems of polynomials arising in engineering and science*, World Scientific Press, Singapore, 2005. MR2160078 (2007a:14065)
34. J. Verschelde, Algorithm 795: PHCpack: a general-purpose solver for polynomial systems by homotopy continuation, *ACM Trans. Math. Software* 25(2) (1999), pp. 251–276. Software available at www.math.uic.edu/∼jan.
35. J. Verschelde and R. Cools, Symbolic homotopy construction, *Appl. Algebra Engrg. Comm. Comput.*, 4 (1993), pp. 169–183. MR1226323 (94c:68102)
36. J. Verschelde, P. Verlinden, and R. Cools, Homotopies exploiting Newton polytopes for solving sparse polynomial systems, *SIAM J. Numer. Anal.*, 31(3) (1994), pp. 915–930. MR1275120 (94m:65084)
37. V. Volterra, Variazionie fluttuazioni del numero d'individui in specie animali convivent, *Mem. Acad. Lincei.*, 2 (1926), pp. 31–113.
38. C.W. Wampler and A.P. Morgan, Solving the kinematics of general 6R manipulators using polynomial continuation, Robotics: Applied Mathematics and Computational Aspects, K. Warwick, ed., Clarendon Press, Oxford, 1993, pp. 57–69. MR1250881
39. C.W. Wampler, A. Morgan, and A.J. Sommese, Complete solution of the nine-point path synthesis problem for four-bar linkages, *ASME J. Mech. Des.* 114(1) (1992), pp. 153–159.
40. B. Yu and B. Dong, A hybrid polynomial system solving method for mixed trigonometric polynomial systems, *SIAM J. Numer. Anal.*, 46(3) (2008), pp. 1503–1518. MR2391004 (2009b:65133)
41. Z. Zeng, The closedness subspace method for computing the multiplicity structure of a polynomial system, in *Interactions of Classical and Numerical Algebraic Geometry*, D. Bates, G. Besana, S. Di Rocco, and C. Wampler (eds.), *Contemporary Mathematics* 496 (2009), pp. 347–362. MR2555964

Department of Mathematics, Mailstop 3368, Texas A&M University, College Station, Texas 77843
E-mail address: jhauenst@math.tamu.edu

Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, Indiana 46556
E-mail address: sommese@nd.edu

General Motors Research and Development, Mail Code 480-106-359, 30500 Mound Road, Warren, Michigan 48090
E-mail address: charles.w.wampler@gm.com