

Regeneration Homotopies for Solving Systems of Polynomials

Jonathan D. Hauenstein ^{*} Andrew J. Sommese [†]

Charles W. Wampler II [‡]

August 15, 2008

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.

Key words and phrases. Algebraic set, component of solutions, diagonal homotopy, embedding, equation-by-equation solver, generic point, homotopy

^{*}Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556 (jhauenst@nd.edu, www.nd.edu/~jhauenst). This author was supported by the Duncan Chair of the University of Notre Dame, the University of Notre Dame, the Center for Applied Mathematics at the University of Notre Dame, the NSF grant DMS-0410047, and the Institute for Mathematics and its Applications in Minneapolis (IMA)

[†]Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556 (sommese@nd.edu, www.nd.edu/~sommese). This author was supported by the Duncan Chair of the University of Notre Dame, the University of Notre Dame, NSF grant DMS-0410047, and the Institute for Mathematics and its Applications in Minneapolis (IMA)

[‡]General Motors Research and Development, Mail Code 480-106-359, 30500 Mound Road, Warren, MI 48090 (Charles.W.Wampler@gm.com, www.nd.edu/~cwampler1) This author was supported by NSF grant DMS-0410047 and the Institute for Mathematics and its Applications in Minneapolis (IMA)

continuation, irreducible component, numerical irreducible decomposition, numerical algebraic geometry, path following, polynomial system, witness point, witness set.

2000 Mathematics Subject Classification. Primary 65H10; Secondary 13P05, 14Q99, 68W30.

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 [7, §11.4], underlie most of the traditional homotopies used to compute numerical solutions of polynomial systems. For example, consider a system

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix} = 0. \quad (1)$$

of n polynomials of degrees d_1, \dots, d_n , respectively, on complex Euclidean space, \mathbb{C}^N , and let $\mathcal{V}(f)$ 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 $\mathcal{V}(f_i)$ of the i -th equation to a union of i linear hyperplanes, for $i = 1, \dots, N$. Assuming some genericity, the intersection of the hyperplanes gives the total degree, $D = d_1 \cdots d_N$, 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 solutions all lie in the closure of $\mathcal{V}(f)$ 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 $B = \mathcal{V}(f_1, \dots, f_k)$, in terms of a witness set (explained in more detail below) and seeks the isolated points in $\mathcal{V}(f)$ by deforming from $A' \cap B$ to $A \cap B$, where $A = \mathcal{V}(f_{k+1}, \dots, f_N)$. 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 [17], set structures [34], product decomposition [19], and polyhedral homotopies [8, 35, 14]. These homotopies deform all of the polynomials in f simultaneously to find its isolated solutions in one

stage. The method in [39] 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 $\mathcal{V}(f)$, but does so in several successive stages. Instead of deforming the polynomials all at once, we replace some of the $\mathcal{V}(f_i)$ 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 $\mathcal{V}(f)$ after n stages of regeneration. We call this working “equation by equation.”

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 [24, 32]. In this approach, a k -dimensional solution component, say $Z_k \subset \mathcal{V}(f)$, 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 set of points that includes $Z_k \cap L$. Regeneration can be applied to find witness supersets, after which other methods remove junk points to obtain the true witness set for each dimension and then break these into irreducible components using monodromy [25] verified by linear traces [26]. Specifically, an equation-by-equation form of regeneration builds a witness set for $\mathcal{V}(f_1, \dots, f_{k+1})$ out of one previously found for $\mathcal{V}(f_1, \dots, f_k)$, proceeding in this fashion to ultimately find a witness set for $\mathcal{V}(f_1, \dots, f_n)$.

An existing method, diagonal homotopy [28, 29], can also find witness sets working equation by equation [30]. 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*. Section 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 diagonal homotopy and with polyhedral homotopy on some test examples.

2 Background

The book [32] overviews the entire field of numerical algebraic geometry, while the survey [14] 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 $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$: here we mean that the property holds true except for a finite set of $\gamma \in S^1$.

Genericity, e.g., generic points for a property of an irreducible algebraic set, and Bertini Theorems are discussed thoroughly in [32], starting with [32, 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 $\mathcal{V}(f)$, read as “variety of f ,” to mean

$$\mathcal{V}(f) = \{x \in \mathbb{C}^N \mid f(x) = 0\}. \quad (2)$$

Hence, $\mathcal{V}(f)$ is just a set of points that carries no multiplicity information. We use $f^{-1}(0)$ to denote $\mathcal{V}(f)$ with its natural structure of a possibly nonreduced scheme. Though multiplicities may be defined, e.g., [7, §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 $\mathcal{V}(f)$. 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, 6].

For an irreducible k -dimensional component Z of $\mathcal{V}(f)$, 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, \dots, L_{N-k} on \mathbb{C}^N ;
2. choose a point $x^* \in Z \cap \mathcal{V}(L_1, \dots, L_{N-k})$;
3. compute the multiplicity of x^* as a component of $(f \cup \{L_1, \dots, L_{N-k}\})^{-1}(0)$.

2.3 Numerical Irreducible Decomposition

For a system of polynomials f on \mathbb{C}^N , the set $Z := \mathcal{V}(f)$ 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 $Z_{i,j}$:

$$Z = \bigcup_{i=1}^{\dim Z} Z_i = \bigcup_{i=1}^{\dim Z} \left(\bigcup_{j \in \mathcal{I}_i} Z_{i,j} \right), \quad (3)$$

where

1. for each i , $Z_i := \cup_{j \in \mathcal{I}_i} Z_{i,j}$;
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 $\mathcal{V}(f)$ 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 $S \subset \mathcal{V}(f) \cap L$ 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 $\mathcal{V}(f)$. At present, the only proven way to determine which points are junk is to find all components of $\mathcal{V}(f)$ 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 $\mathcal{V}(f)$ 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 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 $\mathcal{V}(\xi)$ 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 $\mathcal{V}(\xi)$ 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 [25, 32]. 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 polynomials $f = \{f_1, \dots, f_n\}$. A randomization of f to size $m \leq n$ is a new polynomial system, say g , of the form $g = Af$, where $A \in \mathbb{C}^{m \times n}$ is a generic matrix and f and g are treated as column matrices whose entries are polynomials. For any nonsingular $m \times m$ matrix B , $\mathcal{V}(Bg) = \mathcal{V}(g)$. Since A is generic, the leading $m \times m$ block has an inverse. Premultiplying by this inverse puts g in the form:

$$g = \begin{bmatrix} g_1 \\ \vdots \\ g_m \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} + R \begin{bmatrix} f_{m+1} \\ \vdots \\ f_n \end{bmatrix}, \quad (4)$$

where $R \in \mathbb{C}^{m \times (n-m)}$. Thus, any full randomization $g = Af$ has a special randomization of the form of Eq. 4 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 is favored for requiring fewer operations to evaluate.

Randomization's usefulness derives from the following properties [32, Theorem 13.5.1]. Assume that $A \subset \mathbb{C}^N$ is an irreducible algebraic set. Then, for a generic randomization g of $f = \{f_1, \dots, f_n\}$ to size $m < n$:

1. if $\dim A > N - m$, then A is an irreducible component of $\mathcal{V}(f)$ iff it is an irreducible component of g ;
2. if $\dim A = N - m$, then A is an irreducible component of $\mathcal{V}(f)$ implies that it is also an irreducible component of g ; and
3. if A is an irreducible component of $\mathcal{V}(f)$ with $\dim A \geq N - m$, its multiplicity as a component of $\mathcal{V}(g)$ is greater than or equal to its multiplicity as a component of $\mathcal{V}(f)$ 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 $\mathcal{V}(g)$ include the isolated (respectively, nonsingular) points in $\mathcal{V}(f)$ [31, 32]. When g is "square," that is, when

$m = N$, some of the isolated points in $\mathcal{V}(g)$ may be extraneous, that is, not in $\mathcal{V}(f)$; 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 $\mathcal{V}(f)$, we may “square up” f to form g , a generic randomization f to size N and find all isolated (resp. nonsingular) points in $\mathcal{V}(g)$. 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, $f = \{f_1, \dots, f_N\}$, where each f_i is a polynomial 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 $\mathcal{V}(f) \setminus Y$. Given f and a membership test for Y , we identify two problems:*

- (a) *find all nonsingular points in Z_0 ; and*
- (b) *find a finite set $S \subset \mathcal{V}(f)$ such that $Z_0 \subset S$.*

When seeking only the nonsingular isolated roots in Problem 1a, one may short-cut the complexities of working with sets of multiplicity greater than one. Moreover, in practical problems, the nonsingular roots are often the ones of highest interest. In contrast, Problem 1b 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 $\mathcal{V}(f)$) is a post-processing step not addressed here.

As noted in the final paragraph of § 2.5, the equivalent of Problem 1 for a nonsquare system $f = \{f_1, \dots, f_n\}$, $n > N$, can be treated by squaring up f to size N . For a nonsquare system with $n < N$ polynomials, 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 $\mathcal{V}(f)$, 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 $(Z \cap \mathcal{V}(f)) \setminus Y$.*

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 $\mathcal{V}(f)$, 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 $\mathcal{V}(f)$ 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 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

$$F(x, q) : \mathbb{C}^N \times \mathbb{C}^M \rightarrow \mathbb{C}^n \quad (5)$$

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, but when x^* is singular, these methods fail. One approach to handling singular points by making use of deflation [9, 11, 21, 22] is presented in [32, §15.2.2, §15.3.3]. This approach, outlined below, works if x^* has the same multiplicity structure as the other points in its neighborhood, a condition that is satisfied wherever deflation is applied in this article due to the genericity of the points in question.

Deflation is a regularization operation for polynomial systems in several variables. Introduced by Ojika, Watanabe, and Mitsui [22], it was refined by Ojika [21], and brought to its current form for isolated roots, which includes a proof of termination, by Leykin, Verschelde, and Zhao [11] (see also [9]).

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

Sommese and Wampler [32, §15.2.2, §15.3.3] observed that given a k -dimensional irreducible solution component, the deflation procedure may be done for the component as a whole. Given a system $F(x, q) = 0$ of the form in Eq. 5; a point $q^* \in \mathbb{C}^M$; and an isolated solution x^* of $F(x, q^*) = 0$, it follows, e.g., [32, Corollary A.4.19], that each irreducible solution component Z of $F(x, q^*) = 0$ containing (x^*, q^*) has dimension M and there is a neighborhood $V \subset \mathbb{C}^M$ of q^* such that the map $Z \cap (\mathbb{C}^N \times V) \rightarrow V$ is finite, proper, and onto. The procedure that deflates $F(x, q^*)$ canonically gives rise to a system $\hat{F}(x, z, q) = 0$ of $N + N'$ equations on $\mathbb{C}^N \times \mathbb{C}^{N'} \times \mathbb{C}^M$ with $F(x, z, q^*) = 0$ the system that deflates $F(x, q^*)$. If q^* was a general point of \mathbb{C}^M , then

1. there is a single irreducible solution component Z of $F(x, q) = 0$ containing (x^*, q^*) ;
2. if (x^*, z^*, q^*) is the unique solution of the deflated system $\widehat{F}(x, z, q) = 0$ lying over (x^*, q^*) , then there is a unique irreducible solution component Z' of $\widehat{F}(x, z, q) = 0$ that contains (x^*, z^*, q^*) ;
3. under the projection $(x, z, q) \rightarrow (x, q)$, Z' maps onto Z : this map from Z' to Z is generically one-to-one and gives an isomorphism of a neighborhood of (x^*, z^*, q^*) on Z' onto a neighborhood of (x^*, q^*) on Z .

In this article the genericity assumption on q^* is always satisfied.

Procedures for deflating an isolated point can be found in [11, 12]. These involve forming certain derivative expressions and determining the rank of a sequence of matrices formed from these expressions. A deflation for a component is carried out by finding a deflation for a generic point on the component (isolated by intersecting the component with a generic slicing linear space). Since the point is generic, the same rank sequence holds on a Zariski-open subset of the component, and hence the deflation function that works at the generic point may be used on the whole Zariski-open subset of the component.

4.2 Trackable Paths

Definition 1 (Trackable Path) *Let $H(x, t) : \mathbb{C}^N \times \mathbb{C} \rightarrow \mathbb{C}^N$ 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 $\psi_{x^*} : (0, 1] \rightarrow \mathbb{C}^N$ such that $\psi_{x^*}(1) = x^*$ and $\psi_{x^*}(t)$ is a nonsingular isolated solution of $H(x, t) = 0$ for $t \in (0, 1]$; and*
2. *when x^* is singular, letting $\widehat{H}(x, z, t) = 0$ denote the system that arises through deflation, and letting (x^*, z^*) denote the nonsingular isolated solution of $\widehat{H}(x, z, 1) = 0$ over x^* , we can track the nonsingular solution (x^*, z^*) of $\widehat{H}(x, z, 1)$ for $t \in (0, 1]$ from $t = 1$ to $t = 0$, i.e., there is a smooth map $\psi_{x^*} : (0, 1] \rightarrow \mathbb{C}^N \times \mathbb{C}^{N'}$ such that $\psi_{x^*}(1) = (x^*, z^*)$ and $\psi_{x^*}(t)$ is a nonsingular isolated solution of $H(x, z, t) = 0$ 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 $\lim_{t \rightarrow 0} \psi_{x^*}(t)$ in case (1) and the x coordinates of $\lim_{t \rightarrow 0} \psi_{x^*}(t)$ 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 1, 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 $\lim_{t \rightarrow 0} \psi_{x^*}(t)$ of the paths overcome this difficulty [32, Chapter 10].

4.3 Endpoints at Infinity

The endpoints of some paths in a homotopy may diverge to infinity. The such paths are numerically difficult to track and they are infinitely long. Fortunately, there is a simple maneuver that tames these paths [16]: 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 2 (Complete Homotopy) *Let $H(x, t) : \mathbb{C}^N \times \mathbb{C} \rightarrow \mathbb{C}^N$ be polynomial in x and complex analytic in t . Let S be a finite set of points in $\mathcal{V}(H(x, 1))$. 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 1 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

Suppose that $f(x, q) : \mathbb{C}^N \times \mathbb{C}^M \rightarrow \mathbb{C}^N$ is polynomial in x and complex analytic in q . Let S be the set of isolated (respectively, nonsingular) solutions to $f(x, q_1) = 0$ for generic $q_1 \in \mathbb{C}^M$. Then, the theory of parameter continuation says that for any $q_0 \in \mathbb{C}^M$, the homotopy function $H(x, t) = f(x, tq_1 + (1-t)q_0)$ with start points S is a complete homotopy for finding the isolated (respectively, nonsingular) solutions of $f(x, q_0) = 0$. The nonsingular case was proven in [18] and the full isolated case is proven in [32].

We need a slightly stronger statement of parameter continuation. Suppose that Y is a k -dimensional irreducible algebraic set in \mathbb{C}^N and Y^* a proper algebraic subset of Y . Then, $X = Y \setminus Y^*$ is called a quasiprojective algebraic set.

Theorem 5.1 (Parameter Continuation) *Let X_k be a k dimensional quasiprojective algebraic set. Suppose $f(x, q) : \mathbb{C}^N \times \mathbb{C}^M \rightarrow \mathbb{C}^k$ is polynomial in x and complex analytic in q . Let S be the set of isolated (respectively, nonsingular) points in $\mathcal{V}(f(x, q_1)) \cap X_k$ for generic $q_1 \in \mathbb{C}^M$. Then, $H(x, t) = f(x, tq_1 + (1-t)q_0)$ with start points S is a complete homotopy for finding the isolated (respectively, nonsingular) points in $\mathcal{V}(f(x, q_0)) \cap X_k$.*

This theorem follows from [32, § A.14].

5.2 Product Decomposition

The idea of a homotopy based on a product decomposition was introduced in [19] with related ideas in [34]. 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 $V_1 = \langle \alpha_1, \dots, \alpha_k \rangle$, where the α_i are the basis polynomials for V_1 . Similarly, let's write $V_2 = \langle \beta_1, \dots, \beta_\ell \rangle$. 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 a polynomial f on \mathbb{C}^N , we mean a list $\mathbb{V} := \{V, V_1, \dots, V_m\}$ of vector spaces V, V_1, \dots, V_m of polynomials on \mathbb{C}^N such that f is in the image V of $V_1 \otimes \dots \otimes V_m$ in the space of polynomials. Suppose we form another 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 . 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 5.2 (Product Decomposition) *Suppose that X_k is a k dimensional quasiprojective algebraic set in \mathbb{C}^N . Let $f(x) = \{f_1(x), \dots, f_k(x)\}$ be a polynomial system on \mathbb{C}^N and let $\mathbb{V}_i = \{V_i, V_{i,1}, \dots, V_{i,d_i}\}$ be a product decomposition for f_i , $i = 1, \dots, k$. For each i , $i = 1, \dots, k$, choose a generic product member g_i of V_i . Let S be the set of isolated (respectively, nonsingular isolated) points in $\mathcal{V}(\{g_1, \dots, g_k\}) \cap X_k$. Then, for generic choice of $\gamma \in S^1$, the homotopy*

$$H(x, t) = \{(1-t)f_1(x) + \gamma t g_1(x), \dots, (1-t)f_k(x) + \gamma t g_k(x)\} = 0 \quad (6)$$

with start set S is a complete homotopy for the isolated (respectively, nonsingular isolated) points in $\mathcal{V}(f) \cap X_k$.

Remark 5.3 *If we choose g_i using random complex coefficients in the vector spaces of the product decomposition, then it suffices to choose $\gamma = 1$ in Theorem 5.2. This way of choosing g_i is our usual procedure.*

The theory presented in [19] covers the case where we seek only nonsingular isolated solutions and where $X = \mathbb{C}^N \setminus A$, A a closed algebraic variety. The generalization to the full set of isolated solutions and an arbitrary quasiprojective set X is proven in Appendix A below.

In the regeneration method as described below, we will use linear products in the product decomposition phase. That is, each vector space forming the product has a basis that is a subset of $\{1, x_1, \dots, x_N\}$. It should be noted that such linear products are nearly identical to the set structures described in [34], although the theory presented there only covers nonsingular solutions on $X = \mathbb{C}^N$ and each set must contain 1.

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 $f(x) = \{f_1(x), \dots, f_n(x)\}$, it often happens that not all variables $\{x_1, \dots, x_N\}$ appear in all the polynomials. This leads to the following definition.

Definition 3 *Let $g(x_1, \dots, x_N)$ 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.*

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 A is the support of g , B is the support base of g , and d is the degree of g , then

$$A \subset \underbrace{B \otimes \dots \otimes B}_{d \text{ times}}.$$

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

Definition 4 *For a polynomial g on \mathbb{C}^N , i.e., for $g \in \mathbb{C}[x_1, \dots, x_N]$, a linear support set for g is any subset of $\{1, x_1, \dots, x_N\}$ that includes the support base of g . Associated to a linear support set, say S , is a linear support vector space, say $V = \langle S \rangle$, i.e., the basis elements of V are the monomials in S . A support linear for g is a linear function in a linear support vector space for g , 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 is one whose*

monomials are exactly the support base of g , and its zero set is called a minimal support hyperplane.

As an example, suppose $g = 2x_1^3 + 3x_1x_3 + 1.2$. Then, the support of g is $A = \{x_1^3, x_1x_3, 1\}$, the support base of g is $B = \{x_1, x_3, 1\}$, a minimal support hyperplane is $\mathcal{V}(\alpha x_1 + \beta x_3 + \gamma)$ for any constants $(\alpha, \beta, \gamma) \in \mathbb{C}^3$, and one may easily confirm that $A \subset B \otimes B \otimes B$. A second example where g is homogeneous is $g' = x_1^3 + 2x_1x_3^2 + x_3^2x_4$, which has support $\{x_1^3, x_1x_3^2, x_3^2x_4\}$, support base $\{x_1, x_3, x_4\}$, and whose minimal support hyperplanes are of the form $\mathcal{V}(\alpha x_1 + \beta x_3 + \gamma x_4)$. Assuming we are working on $\mathbb{C}[x_1, x_2, x_3, x_4]$, then $\mathcal{V}(\alpha x_1 + \beta x_2 + \gamma x_3 + \delta x_4 + \zeta)$ is a support hyperplane of g' 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: we use a parameter continuation to get the start points of a product decomposition homotopy that completes the 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.

In the following, it should be understood that a sequence w_i, \dots, w_j is empty if $i > j$.

Lemma 6.1 (Regeneration of a Linear Product) *Let $X_N \subset \mathbb{C}^M$ be an N -dimensional quasiprojective algebraic set, let f_1, \dots, f_m be polynomials on \mathbb{C}^M , and suppose that for $i = m+1, \dots, \widehat{m}$, $m < \widehat{m} \leq N$, we have $g_i = \ell_{i,1} \cdots \ell_{i,d_i}$, where each $\ell_{i,j}$ is a linear function on \mathbb{C}^M . Further, let S_m be the isolated (resp., nonsingular isolated) points of*

$$\mathcal{V}(f_1, \dots, f_m, h_{m+1}, \dots, h_N) \cap X_N,$$

where for $i = m+1, \dots, \widehat{m}$, h_i is a generic supporting linear for $\ell_{i,1}, \dots, \ell_{i,d_i}$. Let $T_{m,\widehat{m}}$ be the isolated (resp., nonsingular isolated) points of

$$\mathcal{V}(f_1, \dots, f_m, g_{m+1}, \dots, g_{\widehat{m}}, h_{\widehat{m}+1}, \dots, h_N) \cap X_N.$$

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 $a = (a_{m+1}, \dots, a_{\widehat{m}}) \in \mathcal{I}_{m,\widehat{m}}$, the start points S_m , and the homotopy function

$$\begin{aligned} H_{m,\widehat{m},a}^{\text{parm}}(x, t) = & \{f_1, \dots, f_m, \\ & (1-t)\ell_{m+1,a_{m+1}} + t h_{m+1}, \dots, (1-t)\ell_{\widehat{m},a_{\widehat{m}}} + t h_{\widehat{m}}, \\ & h_{\widehat{m}+1}, \dots, h_N\} = 0 \quad (7) \end{aligned}$$

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

$$\mathcal{V}(f_1, \dots, f_m, \ell_{m+1,a_{m+1}}, \dots, \ell_{\widehat{m},a_{\widehat{m}}}, h_{\widehat{m}+1}, \dots, h_N) \cap X_N.$$

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

The proof follows immediately from Theorem 5.1, since each homotopy at Eq. 7 is a parameter homotopy in the coefficients of the linear functions $h_{m+1}, \dots, h_{\widehat{m}}$. When applying Theorem 5.1 to the above situation, the k dimensional space X_k appearing in it is the $\widehat{m}-m$ dimensional component of $\mathcal{V}(f_1, \dots, f_m, h_{\widehat{m}+1}, \dots, h_N) \cap X_N$.

The procedure implied by Lemma 6.1 allows us to extend a solution for f_1, \dots, f_m into one for $f_1, \dots, f_{\widehat{m}}$, $\widehat{m} > m$. The following lemma establishes the secondary step of regeneration that accomplishes this.

Lemma 6.2 (Incremental Product Decomposition) *Adopt all the notations of Lemma 6.1. Further, let $\mathbb{V}_i := \{V_i, V_{i,1}, \dots, V_{i,d_i}\}$ be a linear product decomposition for f_i , $i = m+1, \dots, \widehat{m}$, and assume that each g_i , $i = m+1, \dots, \widehat{m}$, is a generic product member of V_i . Then, the start set $T_{m,\widehat{m}}$ with the homotopy function*

$$\begin{aligned} H_{m,\widehat{m}}^{\text{prod}}(x, t) = \{ & f_1, \dots, f_m, \\ & (1-t)f_{m+1} + tg_{m+1}, \dots, (1-t)f_{\widehat{m}} + tg_{\widehat{m}}, \\ & h_{\widehat{m}+1}, \dots, h_N \} = 0 \quad (8) \end{aligned}$$

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

This lemma follows immediately from Theorem 5.2. To apply the theorem, the k dimensional space X_k appearing in it is the $\widehat{m}-m$ dimensional component of $\mathcal{V}(f_1, \dots, f_m, h_{\widehat{m}+1}, \dots, h_N) \cap X_N$.

To apply Theorem 6.2, we need a linear product decomposition $V_{i,1} \otimes \dots \otimes V_{i,d_i}$ for each f_i , $i = m+1, \dots, \widehat{m}$. Here, the number of vector spaces, $d_i \geq \deg f_i$, where $\deg f_i$ means the degree of f_i . Usually we will choose a decomposition such that $d_i = \deg f_i$. 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 both Eq. 7 and Eq. 8, there are linear functions $h_{\widehat{m}+1}, \dots, h_N$ 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 $\mathcal{V}(h_{\widehat{m}+1}, \dots, h_N)$. That is, we use linear algebra once at the beginning of each

incremental stage to find $A \in \mathbb{C}^{N \times \hat{m}}, b \in \mathbb{C}^N$ such that $\text{rank} A = \hat{m}$ and $h_i(Au+b) = 0$ for all $u \in \mathbb{C}^{\hat{m}}, i = \hat{m} + 1, \dots, N$. Then, the homotopies $H(x, t) = 0$ (where H is either H^{param} 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 \hat{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 \hat{m} is not sufficiently small, the extra arithmetic in evaluating ϕ cancels out the savings of tracking on $\mathbb{C}^{\hat{m}}$ instead of \mathbb{C}^N , so it is better to work extrinsically. When \hat{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 6.1 and 6.2 in hand, it is straightforward to solve Problem 1. One merely specifies any set of strictly increasing integers ending at N , say $0 = m_0 < m_1 < \dots < m_r = N$. Then, one solves r incremental problems for $(m, \hat{m}) = (0, m_1), (m_1, m_2), \dots, (m_{r-1}, N)$, 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 6.3 (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 6.1 and 6.2. The only elaboration necessary is to observe that some of the endpoints of the homotopies Eq. 7 and Eq. 8 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}, \dots, h_N , the linear functions that form the linear products g_1, \dots, g_N , and generic choices required in any homotopy membership test.

Procedure $S = \mathbf{Regenerate}(F, Y, \sigma)$

Inputs A set $F = \{f_1, \dots, f_N\}$ of N polynomials on \mathbb{C}^N , a proper subset Y of \mathbb{C}^N in a form suitable for 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 $\mathcal{V}(F) \cap X$, where $X = \mathbb{C}^N \setminus Y$.

Begin 1. Reorder the polynomials f_1, \dots, 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 $0 = m_0 < m_1 < \dots < m_r = N$.
3. Specify a linear product decomposition $V_{i,1} \otimes \dots \otimes V_{i,d_i}$ for each f_i , $i = 1, \dots, N$. One alternative that always suffices is $d_i = \deg f_i$ with $V_{i,j}$, $j = 1, \dots, d_i$, generated by the support base of f_i .
4. Choose a generic product member $\ell_{i,j}$ in each $V_{i,j}$, $i = 1, \dots, N$, $j = 1, \dots, d_i$. This means $\ell_{i,j}$ is a linear function with generic coefficients. Let $g_i = \prod_{j=1}^{d_i} \ell_{i,j}$.
5. For $i = 1, \dots, N$, choose a generic linear h_i that supports all $\ell_{i,j}$, $j = 1, \dots, d_i$.
6. For $i = 1, \dots, r$, let $(m, \widehat{m}) = (m_{i-1}, m_i)$, let

$$G_{m, \widehat{m}} = \{f_1, \dots, f_m, g_{m+1}, \dots, g_{\widehat{m}}, h_{\widehat{m}+1}, \dots, h_N\} \quad (9)$$

$$F_{\widehat{m}} = \{f_1, \dots, f_m, f_{m+1}, \dots, f_{\widehat{m}}, h_{\widehat{m}+1}, \dots, h_N\}, \quad (10)$$

and do the following:

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

Case $m = 0$. Use numerical linear algebra to solve the initial system $\{g_1, \dots, g_{\widehat{m}}, h_{\widehat{m}+1}, \dots, h_N\}$. 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 $H_{m, \widehat{m}, a}^{\text{parm}}$ from Eq. 7 with start set S_m .

- (b) Use a membership test to expunge any points of $T_{m, \widehat{m}}$ that are in Y .
- (c) If $\sigma = \text{True}$, use a local dimension test to expunge any singular points that are not isolated.
- (d) If $\sigma = \text{False}$, 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 $\mathcal{V}(\{f_1, \dots, f_{\widehat{m}}, h_{\widehat{m}+1}, \dots, h_N\}) \cap X$ using the product homotopy $H_{m, \widehat{m}}^{\text{prod}}$ from Eq. 8 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 $\sigma = \text{True}$, use a local dimension test to expunge any singular points of $S_{\widehat{m}}$ that are not isolated.
- (h) If $\sigma = \text{False}$, eliminate any singular points from $S_{\widehat{m}}$.

Return $S = S_N$.

If we are seeking all isolated roots (i.e., if $\sigma = \text{True}$), then Steps 6c and 6g 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 $p = d_1 + d_1 d_2 + \dots + d_1 d_2 \dots d_N$. By reordering the functions so that $d_1 \leq d_2 \leq \dots \leq d_N$, 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 appear in some equations. A good strategy is to order the functions so that cumulative number of distinct variables that have been introduced at any stage is minimal.

When these two strategies are compatible, as in § 9.2, 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 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 $m_0, \dots, m_r = 0, 1, 2, \dots, N - 1, N$. 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. One may take advantage of multilinearity and other forms of sparseness here.

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 algebraic set $Z \setminus Y$, and we wish to update this to a witness set for $(Z \cap \mathcal{V}(f)) \setminus Y$. It is sufficient to be able to do this when f is a single polynomial, because to address a system $f = \{f_1, \dots, f_n\}$, one may repeat the procedure introducing one new equation each time. Let $Z_k = (Z \cap \mathcal{V}(f_1, \dots, f_k)) \setminus Y$. 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, $\mathcal{V}(f)$ is either \mathbb{C}^N (if f is trivial) or it is a hypersurface. If a generic point w of irreducible algebraic set X satisfies $f(w) = 0$, then $X \subset \mathcal{V}(f)$ and so $X \cap \mathcal{V}(f) = X$.

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. 3. 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, $L_k = \{h_1, \dots, h_k\}$ is a set of k generic linear functions on \mathbb{C}^N , and S is the finite set of points $Z_k \cap \mathcal{V}(L_k)$. In the following, we assume that the same linear functions h_i , $i = 1, \dots, N$, 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 $L_N = \{h_1, \dots, h_N\}$ 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 $S = \mathbf{RegenWitness}(W, f, Y)$

Inputs A witness set W for 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 $(Z \cap \mathcal{V}(f)) \setminus Y$.

- Begin**
1. Initialize \widehat{W}_k , $k = N, \dots, 0$, as empty.
 2. Specify a linear product decomposition $V_1 \otimes \dots \otimes V_d$ 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 product member ℓ_i in each V_i , $i = 1, \dots, d$. This means ℓ_i is a linear function with generic coefficients. Let $g = \prod_{i=1}^d \ell_i$.
 4. For $k = N, N-1, \dots, 0$, 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 $w^* \in \mathcal{V}(P, L_k)$) 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. 4). For simplicity, call the result P again.
 - B. For $i = 1, \dots, d$, start at w and track the continuation path of

$$H_1(x, t) = \{P(x), h_1(x), \dots, h_{k-1}(x), \\ th_k(x) + (1-t)\ell_i(x)\} = 0. \quad (11)$$

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 $\mathcal{V}(H_1(x, 0))$.
- E. Track all paths starting at T for the homotopy function

$$H_2(x, t) = \{P(x), h_1(x), \dots, h_{k-1}(x), \\ tg(x) + (1-t)f(x)\} = 0. \quad (12)$$

Call the set of endpoints S .

- F. Use a membership test to expunge any points of S that are in Y .
- G. Append $\{\{P, f\}, \{h_1, \dots, h_{k-1}\}, S\}$ to \widehat{W}_{k-1} .
- (b) Remove from \widehat{W}_{k-1} any points on higher than $k-1$ dimensional sets of $\mathcal{V}(P, f)$. 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 $\widehat{W} = \{\widehat{W}_N, \dots, \widehat{W}_0\}$.

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

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

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**. There are just two significant differences. First, 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. Here, we are keeping track of solution sets at every dimension.

8 Diagonal Homotopy

The extrinsic diagonal homotopy in [28] and its intrinsic reformulation in [29] 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 [30]. In brief, to find $\mathcal{V}(f) \cap \mathcal{V}(g)$, diagonal homotopy finds a witness set for the composite system $\{f(x), g(y), x - y\}$. In the equation-by-equation approach, f 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 $W_k = \{P, \{h_1, \dots, h_k\}, X\}$, 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 $P = \{p_1, \dots, p_{N-k}\}$. Set Z is a component of $\mathcal{V}(P)$ and X is the set of isolated points in $Z \cap \mathcal{V}(h_1, \dots, h_k)$. We wish to introduce a single new polynomial f on \mathbb{C}^N and find the witness set for $Z \cap \mathcal{V}(f)$. 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 $\{h_1, \dots, h_{k-1}\}$.

To start the diagonal homotopy, we begin by finding a witness set R for $\mathcal{V}(f)$, which consists of the points $\mathcal{V}(f, h_1, \dots, h_{N-1})$. Notice that R , W_k , and W_{k-1} all lie in the linear space $\Phi = \mathcal{V}(h_1, \dots, h_{k-1})$, which is $(N - k + 1)$ dimensional. A fully extrinsic form of diagonal homotopy works on a doubled set of variables $(x, y) \in \mathbb{C}^N \times \mathbb{C}^N$. In a manner similar to § 6.3, we can keep the number of variables smaller by restricting our computations to Φ . This is done by computing a basis for Φ , say $\phi(u) = Au + b$, where $u \in \mathbb{C}^{N-k+1}$ is a new set of intrinsic variables and $A \in \mathbb{C}^{N \times (N-k+1)}$ and $b \in \mathbb{C}^N$ are constants. Each point $x \in \Phi$ has a unique preimage $\phi^{-1}(x)$. Let $S = \phi^{-1}(R) \times \phi^{-1}(W_k) \subset \mathbb{C}^N \times \mathbb{C}^N$. We will call working on Φ as the “semi-intrinsic” form of diagonal homotopy to distinguish it from the fully intrinsic form to be discussed later.

For convenience, let $\kappa = N - k$. With this and the above notations, the semi-intrinsic form of the diagonal homotopy becomes $H : \mathbb{C}^{\kappa+1} \times \mathbb{C}^{\kappa+1} \times \mathbb{C} \rightarrow \mathbb{C}^{2(\kappa+1)}$ defined as

$$\begin{aligned} H_1(u, v, t) = & \{P(\phi(u)), f(\phi(v)), \\ & (1-t)(u_1 - v_1) + \gamma th_k(\phi(v)), \dots, (1-t)(u_\kappa - v_\kappa) + \gamma th_{N-1}(\phi(v)), \\ & (1-t)(u_{\kappa+1} - v_{\kappa+1}) + \gamma th_k(\phi(u))\} = 0. \end{aligned} \quad (13)$$

Starting at points S and tracking $H(u, v, t) = 0$ from $t = 1$ to $t = 0$ gives witness points of the form (u, u) , and for each such point, say (u^*, u^*) , that is an isolated solution of $H_1(u, v, 0) = 0$, we append

$$\{\{P, f\}, \{h_1, \dots, h_k\}, \phi(u^*)\}$$

to \widehat{W}_{k-1} .

The semi-intrinsic homotopy still involves a partial of doubling of the number of variables. The doubling can be avoided altogether by a fully intrinsic formulation along the lines of [29], which also takes advantage of the linearity of the diagonal equations $u - v = 0$ that appear in $H_1(u, v, 0)$. Let $\widehat{\phi} : \mathbb{C}^{\kappa+1} \rightarrow \mathbb{C}^{2N}$ 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 $\mathcal{V}(f)$ are cut out by smaller linear spaces, $N - k$ and 1 dimensional, respectively. Our fully intrinsic formulation uses parameterizations of these as well. Specifically, let $\psi_1(u) = A_1 u_{1:\kappa} + b_1$ parameterize $\Psi_1 = \mathcal{V}(h_1, \dots, h_k)$ and let $\psi_2(u) = A_2 u_{\kappa+1} + b_2$ parameterize $\Psi_2 = \mathcal{V}(h_1, \dots, h_{N-1})$, where $A_2 \in \mathbb{C}^{N \times \kappa}$, $A_1 \in \mathbb{C}^{N \times 1}$, and $b_1, b_2 \in \mathbb{C}^N$. From these, we form the linear map $\psi : \mathbb{C}^{\kappa+1} \rightarrow \mathbb{C}^{2N}$ 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 $\pi_1, \pi_2 : \mathbb{C}^N \times \mathbb{C}^N \rightarrow \mathbb{C}^N$ be projections defined as $\pi_1 : (u, v) \mapsto u$ and $\pi_2 : (u, v) \mapsto v$. Then, a fully intrinsic homotopy function $H_2 : \mathbb{C}^{\kappa+1} \times \mathbb{C} \rightarrow \mathbb{C}^{\kappa+1}$ is

$$\begin{aligned} H_2(u, t) = & \{P(\pi_1[(1-t)\widehat{\phi}(u) + \gamma t\psi(u)]), \\ & f(\pi_2[(1-t)\widehat{\phi}(u) + \gamma t\psi(u)])\} = 0, \end{aligned} \quad (14)$$

where $\gamma \in S^1$ 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 $\{\{P, f\}, \{h_1, \dots, h_k\}, \phi(u^*)\}$ that is appended to \widehat{W}_{k-1} .

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. If it is available, a local dimension test can be used as an alternative.

The semi-intrinsic formulation tracks $2(\kappa+1)$ variables whereas the fully intrinsic one tracks just $\kappa + 1$. Consequently, the fully intrinsic version is computationally less expensive.

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 [33] and HOM4PS-2.0.15 [13] were used in the examples described below.

9.1 Illustrative Example

To demonstrate both the diagonal and regeneration approaches described in this paper to find nonsingular isolated solutions, consider the following system used in [24, 30]:

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

In the diagonal approach, a witness point set W_j is found for each of the functions by tracking a total of 19 paths. When finding the isolated solutions, all but one point in each W_j can be removed, which correspond to $x = 0.5$, $y = 0.5$, and $z = 0.5$, respectively. The points removed satisfy another function, which means that these points cannot lead to isolated solutions.

On the first stage of the diagonal approach, a diagonal homotopy tracks one path that finds the witness point corresponding to $x = y = 0.5$. On the second stage of the diagonal approach, a diagonal homotopy tracks one path which leads to the only isolated solution $(0.5, 0.5, 0.5)$.

In the regeneration approach, 5 paths are tracked to solve F_1 defined by Eq. 10. Of these solutions, 4 satisfy another function and can be removed leaving the point that corresponds to $x = 0.5$. Since the second function is of degree 6, we take the linear slice $h_2 = l_{2,1}$ and regenerate to the linear slices $l_{2,2}, \dots, l_{2,6}$. This creates 6

paths that need to be tracked for the second stage regeneration homotopy. Of these paths, 2 diverge, 3 have endpoints that satisfy the third function and the remaining endpoint corresponds to $x = y = 0.5$. Since the third function is of degree 8, we take the linear slice $h_3 = l_{3,1}$ and regenerate to the linear slices $l_{3,2}, \dots, l_{3,8}$. This creates 8 paths that need to be tracked for the third stage regeneration homotopy. 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 $(0.5, 0.5, 0.5)$.

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 [37, 32]. The polynomial system, available at [3], consists of 2 linear and 10 quadratic polynomials in 12 variables. The system was setup using random parameter values and has 16 nonsingular finite isolated solutions.

As suggested in §6.5, the ordering of the quadratic polynomials that minimizes the total number of paths that need tracked is suggested by the setup. 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. This suggests that the four normalizing quadratics should be placed last. In this optimal setup, which we will call “order A,” the total number of paths tracked by the regeneration method is 628 and total number of slices moved is 313. For comparison, another ordering (“order B”) was setup by placing the four normalizing quadratics ahead of the other six. In order B, the number of paths tracked increased to 928 and the number of slices moved increased to 463.

As a comparison of the methods, these two orderings of the polynomial system were solved using a generic total degree homotopy, the diagonal approach and the regeneration approach. All 3 methods were run using adaptive precision [4] 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.

The problems were first run using the diagonal approach implemented in PHCpack [33]. 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 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.

setup	total degree		diagonal		regeneration		
	paths tracked	time, s	paths tracked	time, s	paths tracked	slices moved	time, s
order A	1024	92.23	649	78.84	628	313	21.07
order B	1024	92.23	949	90.48	928	463	28.89

Table 1: Summary of solving the general 6R, serial-link robot system

method	paths tracked	slices moved	time, hrs
regeneration	136,296	66,888	8.065
polyhedral	87,639		11.656

Table 2: Comparison of regeneration and polyhedral homotopy for the nine-point path synthesis problem

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 [30] and the examples above. One advantage of regeneration over the diagonal approach is the ability to easily incorporate 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 [23] consists of 8 seven degree polynomials in 8 variables with a natural two-fold symmetry. Utilizing the symmetry, the results of [38] show that this system has 4326 nondegenerate solutions appearing in 1442 cognate triples.

The Bezout count for the Roth and Freudenstein system utilizing the 4-homogeneous structure and two-fold symmetry is 322,560. Regeneration is easily setup to exploit both the symmetry and multi-homogeneous structure. The regeneration system is able to find the 4326 nondegenerate solutions by tracking 136,296 paths and moving 66,888 slices utilizing adaptive precision [4]. Table 2 compares solving this system with regeneration and a polyhedral homotopy using HOM4PS-2.0 [13].

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 [19]. Although we did not test this option, it promises to further reduce both the 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. To illustrate this, consider a sparse polynomial system arising from ongoing research by the first two authors and Bei Hu (Univer-

sity of Notre Dame) related to the discretization of the stationary Lotka-Volterra population model with diffusion [15, 36].

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

$$\begin{aligned} f_{ij} &= \frac{1}{25} (u_{i+1,j} - 2u_{i,j} + u_{i-1,j}) \\ &\quad + \frac{1}{(n+1)^2} (u_{i,j+1} - 2u_{i,j} + u_{i,j-1}) + \frac{1}{25(n+1)^2} u_{i,j} (1 - v_{i,j}) \\ g_{ij} &= \frac{1}{25} (v_{i+1,j} - 2v_{i,j} + v_{i-1,j}) \\ &\quad + \frac{1}{(n+1)^2} (v_{i,j+1} - 2v_{i,j} + v_{i,j-1}) + \frac{1}{25(n+1)^2} v_{i,j} (u_{i,j} - 1) \end{aligned} \tag{15}$$

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. Even though the system has a natural 2-homogeneous structure, with each polynomial being of type $(1, 1)$, it is not practical to use this as n increases. Also, even though the mixed volume of the system is the same as the number of solutions, 2^{4n} , 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. 15. The linear product decomposition of the polynomials used was:

$$\begin{aligned} f_{ij} &\in \langle \{1, u_{i+1,j}, u_{i,j}, u_{i-1,j}, u_{i,j+1}, u_{i,j-1}, v_{i,j}\} \times \{1, v_{i,j}\} \rangle, \\ g_{ij} &\in \langle \{1, v_{i+1,j}, v_{i,j}, v_{i-1,j}, v_{i,j+1}, v_{i,j-1}, u_{i,j}\} \times \{1, u_{i,j}\} \rangle. \end{aligned} \tag{16}$$

This decomposition was used so that a generic member of the first set of the product decomposition is a generic linear that supports the second linear. In particular, in **Regenerate**, the generic linear selected in Step 5 was taken as the first linear selected in Step 4.

With the above choices, regeneration tracks roughly 4 times as many paths as the number of solutions. Table 3 compares the number of paths for various methods and Table 4 contains timings for the various software packages. For $n \leq 4$, regeneration can solve the system using only double precision, and for $n = 5$, regeneration utilized adaptive precision tracking [4] to track the paths since double precision was not 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

	total degree	2-homogeneous	polyhedral	regeneration	
n	paths	paths	paths	paths	slices moved
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

Table 3: Comparison of various methods 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

Table 4: Single-processor timings for polyhedral method and regeneration for solving systems related to the Lotka-Volterra population model

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 non-incremental 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 less computation. The mixed volume computations used to create the start system for polyhedral homotopy is 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.

A Theory

In this section we prove the technical results underlying regeneration. They are a generalization of the results in the appendices of [19].

We refer to [32] for more details about algebraic sets.

Recall a quasiprojective algebraic set is a Zariski open set U of a projective algebraic set X . All algebraic sets are complex and quasiprojective. We allow X and U to be nonreduced. An algebraic set 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 say that an algebraic set X is generically reduced if the set where X is

nonreduced is the complement of a Zariski open dense subset of X . When X is a component of the solution set of a system of algebraic functions, this is the same as saying that X has multiplicity one.

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

Given a vector space V of algebraic functions on an algebraic set X , the base locus of V (denoted $\text{Bs}(V)$) is the set of common zeros of V on X . $\text{Bs}(V)$ is an algebraic subset of X . If $\text{Bs}(V)$ is the empty set, we say V is basepoint free. Bertini's theorem, e.g., see [32] for an extensive discussion, guarantees that the solution set $\mathcal{V}(g)$ of a general element g of V has various strong properties outside $\text{Bs}(V)$, e.g., the intersection of $\mathcal{V}(g)$ with the reduction of $X \setminus (\text{Bs}(V) \cup \text{Sing}(X))$ is smooth.

The following generalizes [19, Lemma 3 of A.3].

Theorem A.1 *Let W be a generically reduced algebraic curve. For i from 1 to m , let V_i be a finite vector space of algebraic functions on W . 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 W . 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 for a general $\gamma \in S^1$, it follows that for the homotopy*

$$H(w, t) = (1 - t)f(w) + \gamma tg(w)$$

and each element \mathfrak{g} of the set \mathcal{G} of nonsingular solutions of g on W , there is a differentiable function $s_{\mathfrak{g}}(t) : (0, 1] \rightarrow W$ satisfying $H(t, s_{\mathfrak{g}}(t)) = 0$ and $s_{\mathfrak{g}}(1) = \mathfrak{g}$. Every solution $w^ \in W$ of f is the limit at $t = 0$ of at least one of the paths $s_{\mathfrak{g}}(t)$.*

Proof. Without loss of generality we may assume that W is smooth. To see this, let \mathcal{S} denote the union of $\text{Sing}(W)$, the singular set of the reduction of W , plus the finite set of nonreduced points of W . Let $\pi : W' \rightarrow W$ denote the desingularization of W , i.e., π is a holomorphic finite-to-one map from a nonsingular algebraic curve W' onto W , which maps $W \setminus \pi^{-1}(\mathcal{S})$ isomorphically onto $W \setminus \mathcal{S}$. 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 W' . Note by Bertini's theorem the solutions of the g_i and g in W are nonsingular and thus miss \mathcal{S} . Therefore we may identify the solutions of g (respectively g_i for i from 1 to m) on W with the solutions of $g' := g \circ \pi$ (respectively $g'_i := g_i \circ \pi$ for i from 1 to m) on W' , which are also nonsingular. Note that the hypotheses of the Theorem hold for $W', V', V'_1, \dots, V'_m$. To see that the conclusion on W' implies the conclusion on W , it suffices to show that

$$\pi^{-1}(\mathcal{S}) \cap [\cup_{\mathfrak{g} \in \mathcal{G}} s_{\mathfrak{g}}((0, 1])] = \emptyset. \quad (17)$$

To see this consider the algebraic curves $S := \mathbb{P}^1 \times (W' \cap \pi^{-1}(\mathcal{S}))$ and

$$X := \{(w, [\lambda, \mu]) \in W \times \mathbb{P}^1 \mid \lambda f(w) + \mu g(w) = 0\}.$$

Since the solutions of g on W don't meet \mathcal{S} , the intersection $S \cap X$ is finite, and thus so is its image $\mathcal{D} \subset \mathbb{P}^1$. Since given any finite set of \mathbb{P}^1 and a generic $\gamma \in S^1$, the paths $[t, \gamma(1-t)]$ for $t \in (0, 1]$ miss the finite set, condition 17 must hold. Thus W may be assumed to be smooth.

Let K be the unique smooth compact algebraic curve, which contains W 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 W . For each i , a Zariski dense set of the functions in V_i have the same pole set with multiplicities $D_i := \sum_{j \in \mathcal{J}_i} \mu_{i,j} k_{i,j}$ with all of the points $k_{i,j}$ lying in the finite set $K \setminus W$. 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, we 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 .

Let B_i denote the base locus of the elements of V_i , i.e., let B_i denote the common solution set $n_{i,1}w_{i,1} + \cdots + n_{i,k_i}w_{i,k_i}$ of the elements of V_i . Here $n_{i,1}$ denotes the multiplicity of the solution $w_{i,1}$. Note that a general element $g_i \in V_i$ vanishes at the $v_{i,j}$ with multiplicity exactly $n_{i,j}$ for each j from 1 to k_i . The base locus of V on K is

$$B_1 + \cdots + B_m = \sum_{i=1}^m \sum_{j=1}^{k_i} n_{i,j} w_{i,j}.$$

Note that by hypothesis, $B \subset K \setminus W$. For each i , let \mathcal{B}_i denote the line bundle on K associated to B_i with s_i the tautological section of \mathcal{B}_i (unique up to multiplication by a nonzero constant) having B_i as its zero set counting multiplicities. Let s denote the section $s_1 \otimes \cdots \otimes s_m$ of $\mathcal{B} := \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_m$. Note that the zero set with multiplicities of s is B .

Now for each i , let $\widehat{L}_i := L_i \otimes \mathcal{B}_i^*$, and let $\widehat{L} := \widehat{L}_1 \otimes \cdots \otimes \widehat{L}_m$. For each i , dividing each element v_i of V_i by s_i gives rise to an algebraic section \widehat{v}_i of \widehat{L}_i . These \widehat{v}_i form a vector space that is sent bijectively to V_i by tensoring with s_i . Similarly using s , elements of V give rise to elements of a vector space \widehat{V} of sections of \widehat{L} .

Note that on $W \times \mathbb{C}$, the zero set of

$$\widehat{H}(w, t) = (1-t)\widehat{f}(w) + \gamma t \widehat{g}(w)$$

coincides with that of $H(w, t)$. Thus to prove the Theorem we may assume without loss of generality that \mathcal{B} and all the \mathcal{B}_i are empty. In this case the result follows by the same reasoning as the part of the proof of [19, Lemma 3 of A.3]. \square

Theorem A.2 *Let*

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_N(x) \end{bmatrix} = 0 \tag{18}$$

denote a system of polynomials on \mathbb{C}^N . Let V_1, \dots, V_m denote vector spaces of polynomials on \mathbb{C}^N . Assume that $f_k(x)$ is in the image V of $V_1 \otimes \dots \otimes V_m$ in the space of polynomials. Assume that U is a Zariski open dense set of \mathbb{C}^N over which each of the V_j is basepoint free. Choose generic $g_j \in V_j$ for $1 \leq j \leq m$. Let A_j for j from 1 to m denote the isolated solutions of

$$\begin{bmatrix} f_1(x) \\ \vdots \\ f_{k-1}(x) \\ g_j(x) \\ f_{k+1}(x) \\ \vdots \\ f_N(x) \end{bmatrix} = 0$$

and let $A := \cup_{j=1}^m A_j$. Then, letting $\gamma \in S^1$ be general, the homotopy

$$H(x, t) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_{k-1}(x) \\ (1-t)f_k(x) + \gamma t g_{k,1}(x) \cdots g_{k,m_k}(x) \\ f_{k+1}(x) \\ \vdots \\ f_N(x) \end{bmatrix} = 0$$

is a complete homotopy with respect to A in the sense of (4.4) for computing the isolated solutions of

$$\begin{bmatrix} f_1(x) \\ \vdots \\ f_N(x) \end{bmatrix} = 0.$$

contained in U as the limits of the tracking of the points in A as t goes to zero.

Proof. Let y be an isolated solution of the system (18) on U . By [32, Corollary A.4.19], y is contained in a one dimensional irreducible component $W \subset U$ of the solution set of

$$\begin{bmatrix} f_1(x) \\ \vdots \\ f_{k-1}(x) \\ f_{k+1}(x) \\ \vdots \\ f_N(x) \end{bmatrix} = 0.$$

The result now follows immediately from Theorem A.1. □

Corollary A.3 *Let*

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_N(x) \end{bmatrix} = 0 \quad (19)$$

denote a system of polynomials on \mathbb{C}^N . For i from 1 to N , let $V_{i,1}, \dots, V_{i,m_i}$ denote vector spaces of polynomials on \mathbb{C}^N . Assume that $f_k(x)$ is in the image V_i of $V_{i,1} \otimes \dots \otimes V_{i,m_i}$ in the space of polynomials. Assume that U is a Zariski open dense set of \mathbb{C}^N over which each of the $V_{i,j}$ is basepoint free. For each i choose generic $g_{i,j} \in V_{i,j}$ for $1 \leq j \leq m_i$. For any $k+1 \leq \ell \leq N$ and any choices of $1 \leq j_\ell \leq m_\ell$; let $A_{k,j}$ for j from 1 to m_k denote the isolated solutions of

$$\begin{bmatrix} f_1(x) \\ \vdots \\ f_{k-1}(x) \\ g_{k,j}(x) \\ g_{k+1,j_{k+1}}(x) \\ \vdots \\ g_{N,j_N}(x) \end{bmatrix} = 0;$$

and let $A_k := \cup_{j=1}^{m_k} A_{k,j}$. Then, letting $\gamma \in S^1$ be general, the homotopy

$$H(x, t) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_{k-1}(x) \\ (1-t)f_k(x) + \gamma t g_{k,1}(x) \cdots g_{k,m_k}(x) \\ g_{k+1,j_{k+1}}(x) \\ \vdots \\ g_{N,j_N}(x) \end{bmatrix} = 0$$

is a complete homotopy with respect to A in the sense of (4.4) for computing the isolated solutions of

$$\begin{bmatrix} f_1(x) \\ \vdots \\ f_k(x) \\ g_{k+1,j_{k+1}}(x) \\ \vdots \\ g_{N,j_N}(x) \end{bmatrix} = 0.$$

contained in U as the limits of the tracking of the points in A as t goes to zero.

Proof. Choose the $g_{i,j}$ inductively. First choose the $g_{1,j}$ generically given f_1, \dots, f_N . Assuming that the $g_{i,j}$ have been chosen for $i \leq k$, choose the $g_{k,j}$ generically given f_1, \dots, f_N and the $g_{i,j}$ have been chosen for $i \leq k$. The Corollary now follows by successive application of Theorem A.2. \square

References

- [1] D.J. Bates, J.D. Hauenstein, C. Peterson, and A.J. Sommese, A local dimension test for numerically approximated points on algebraic sets, preprint.
- [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.
- [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.
- [5] M. Beltrametti and A.J. Sommese, *The adjunction theory of complex projective varieties*. Expositions in Mathematics, 16, Walter De Gruyter, Berlin, 1995.
- [6] 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.
- [7] 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.
- [8] B. Huber and B. Sturmfels, A polyhedral method for solving sparse polynomial systems, *Math. Comp.*, 64(212) (1995), pp. 1541–1555.
- [9] G. Lecerf, Quadratic Newton iteration for systems with multiplicity, *Found. Comput. Math.*, 2(3) (2002), pp. 247–293.
- [10] G. Lecerf, Computing the equidimensional decomposition of an algebraic closed set by means of lifting fibers, *J. Complexity* 19 (2003), pp. 564–596.
- [11] A. Leykin, J. Verschelde, and A. Zhao, Newton’s method with deflation for isolated singularities of polynomial systems, *Theoretical Computer Science* 359 (2006), pp. 111–122.
- [12] 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.
- [13] 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.

- [14] T.Y. Li, Numerical solution of polynomial systems by homotopy continuation methods, in *Handbook of Numerical Analysis. Volume XI. Special Volume: Foundations of Computational Mathematics*, edited by F. Cucker, North-Holland, 2003, pp. 209–304.
- [15] A.J. Lotka, Undamped oscillations derived from the laws of mass action, *J. Amer. Chem. Soc.* 42 (1920), pp. 1595–1599.
- [16] A.P. Morgan A transformation to avoid solutions at infinity for polynomial systems, *Appl. Math. Comput.*, 18(1) (1986), pp. 77–86.
- [17] 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.
- [18] 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), pp. 207.
- [19] 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.
- [20] D. Mumford, *Algebraic geometry I: Complex projective varieties*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1976 edition.
- [21] 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.
- [22] 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.
- [23] B. Roth and F. Freudenstein, Synthesis of Path-Generating Mechanisms by Numerical Methods, *ASME Journal of Engineering for Industry*, 85B(3) (1963), pp. 298–306.
- [24] 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.
- [25] 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.

- [26] 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.
- [27] 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.
- [28] 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.
- [29] A.J. Sommese, J. Verschelde, and C.W. Wampler, An intrinsic homotopy for intersecting algebraic varieties, *Journal of Complexity*, 21 (2005), pp. 593–608.
- [30] 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.
- [31] 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.
- [32] A.J. Sommese and C.W. Wampler, *The Numerical solution of systems of polynomials arising in engineering and science*, World Scientific Press, Singapore, 2005.
- [33] 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.
- [34] J. Verschelde and R. Cools, Symbolic homotopy construction, *Appl. Algebra Engrg. Comm. Comput.*, 4 (1993), pp. 169–183.
- [35] 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.
- [36] V. Volterra, Variazioni e fluttuazioni del numero d'individui in specie animali conviventi, *Mem. Acad. Lincei.*, 2 (1926), pp. 31–113.
- [37] 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.

- [38] C.W. Wampler, A. Morgan, and A.J. Sommese, Complete solution of the nine-point path synthesis problem for four-bar linkages, *ASME Journal of Mechanical Design* 114(1) (1992), pp. 153–159.
- [39] 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.