Homotopy Continuation Methods for Solving Polynomial Systems

by Tien-Yien Li[*†]

1. Introduction

Let $P(\mathbf{x}) = \mathbf{0}$ be a system of n polynomial equations in n unknowns. Denoting $P = (p_1, \ldots, p_n)$, we want to find all isolated solutions of

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ p_n(x_1, \ldots, x_n) &= 0 \end{aligned} \tag{1}$$

for $\mathbf{x} = (x_1, \ldots, x_n)$. This problem is very common in many fields of science and engineering, such as formula construction, geometric intersection problems, inverse kinematics, power flow problems with PQ-specified bases, computation of equilibrium states, etc. Many of those applications has been well documented in [40]. Elimination theory based methods, most notably the Buchberger algorithm [7] for constructing Gröbner bases, are the classical approach to solving (1), but their reliance on symbolic manipulation makes those methods seem somewhat limited to relatively small problems.

Solving polynomial systems is an area where numerical computations arise almost naturally. Given the complexity of the problem, we must use standard machine arithmetic to obtain efficient programs. Moreover, by Galois theory explicit formulas for the solutions are unlikely to exist. We are concerned with the robustness of our methods and want to be sure that *all* isolated solutions are obtained, i.e., we want exhaustive methods. In 1977, Garcia and Zangwill [12] and Drexler [11] independently presented theorems suggesting that homotopy continuation could be used to find the full set of isolated solutions of (1) numerically. During the last few decades, this method has been successfully developed and proved to be a reliable and efficient numerical algorithm for approximating all isolated zeros of polynomial systems.

In recent years, a new research area called *Numerical Algebraic Geometry* has emerged. In which, developed homotopy continuation methods in finding all the isolated zeros have been employed to identify irreducible solution components of polynomial systems, finding their dimensions and their degrees. Several software packages, such as: PHCpack [42], PHoM [14], HOM4PS [9, 22, 26], Bertini [4], have been produced for those purposes. In this short survey, we only focus on the fundamental ideas of the methods in finding all the isolated zeros of polynomial systems.

2. Deficient Polynomial Systems

In the early stage, the homotopy continuation method for solving (1) is to define a trivial system $Q(\mathbf{x}) = (q_1(\mathbf{x}), \ldots, q_n(\mathbf{x})) = \mathbf{0}$ and then follow the curves in the real variable t which make up the solution set of

$$\mathbf{0} = H(\mathbf{x}, t) = (1-t)Q(\mathbf{x}) + tP(\mathbf{x}). \tag{2}$$

More precisely, if $Q(\mathbf{x}) = \mathbf{0}$ is chosen correctly, the following three properties hold:

Property 0 (*Triviality*). The solutions of $Q(\mathbf{x}) = \mathbf{0}$ are known.

[*] Department of Mathematics, Michigan State University, East Lansing, Michigan, U.S.A.
E-mail: li@math.msu.edu
[†] Research supported in part by NSF under Grant DMS 11-15587.

Property 1 (*Smoothness).* The solution set of $H(\mathbf{x},t) = \mathbf{0}$ for $0 \le t < 1$ consists of a finite number of smooth paths, each parametrized by t in $[0,1)$.

Property 2 (*Accessibility).* Every isolated solution of $H(\mathbf{x},1) = P(\mathbf{x}) = \mathbf{0}$ can be reached by some path originating at $t = 0$. It follows that this path starts at a solution of $H(\mathbf{x},0) = Q(\mathbf{x}) = \mathbf{0}$.

When the three properties hold, the solution paths can be followed from the initial points (known because of **Property 0**) at $t = 0$ to all solutions of the original problem $P(\mathbf{x}) = \mathbf{0}$ at $t = 1$ using standard numerical techniques ([2]).

Several authors have suggested choices of Q that satisfy the three properties (cf. [10, 24, 31, 44, 45] for a partial list). A typical suggestion is

$$\begin{aligned} q_1(\mathbf{x}) &= a_1 x_1^{d_1} - b_1, \\ &\vdots \\ q_n(\mathbf{x}) &= a_n x_n^{d_n} - b_n, \end{aligned} \tag{3}$$

where $d_1,\ldots,d_n$ are the degrees of $p_1(\mathbf{x}),\ldots,p_n(\mathbf{x})$ respectively and a_j, b_j are random complex numbers (and therefore nonzero, with probability one). So in one sense, the original problem we posed is solved. All solutions of $P(\mathbf{x}) = \mathbf{0}$ are found at the end of the $d_1 \times \cdots \times d_n$, called total degree, paths that make up the solution set of $H(\mathbf{x},t) = \mathbf{0}$, $0 \le t < 1$.

The book by A. Morgan [32] detailed many aspects of the above approach. A major part of this article will focus on the development afterwards that makes this method more convenient to apply.

The reason the problem is not satisfactorily solved by the above considerations is the existence of *extraneous paths*. Although the above method produces $d := d_1 \times \cdots \times d_n$ paths, the system $P(\mathbf{x}) = \mathbf{0}$ may have fewer than d solutions. We call such a system *deficient*. In this case, some of the paths produced by the above method will be extraneous paths.

More precisely, even though **Properties 0–2** imply that each solution of $P(\mathbf{x}) = \mathbf{0}$ will lie at the end of a solution path, it is also consistent with these properties that some of the paths may diverge to infinity as the parameter t approaches 1 (the smoothness property rules this out for $t \to t_0 < 1$). In other words, it is quite possible for $Q(\mathbf{x}) = \mathbf{0}$ to have more solutions than $P(\mathbf{x}) = \mathbf{0}$. In this case, some of the paths leading from roots of $Q(\mathbf{x}) = \mathbf{0}$ are extraneous, and diverge to infinity when $t \to 1$.

Empirically, we find that most systems arising in applications are deficient. A great majority of the systems have fewer than, and in some cases only a small fraction of, the "expected number" of solutions. For a typical example of this sort, let's look at the following Cassou-Nogues system [40]:

(4)

$$\begin{aligned} p_1 &= 15b^4cd^2 + 6b^4c^3 + 21b^4c^2d - 144b^2c - 8b^2c^2e \\ &\quad - 28b^2cde - 648b^2d + 36b^2d^2e + 9b^4d^3 - 120, \\ p_2 &= 30b^4c^3d - 32cde^2 - 720b^2cd - 24b^2c^3e - 432b^2c^2 \\ &\quad + 576ce - 576de + 16b^2cd^2e + 16d^2e^2 + 16c^2e^2 + 9b^4c^4 \\ &\quad + 39b^4c^2d^2 + 18b^4cd^3 - 432b^2d^2 + 24b^2d^3e - 16b^2c^2de \\ &\quad - 240c + 5184, \\ p_3 &= 216b^2cd - 162b^2d^2 - 81b^2c^2 + 1008ce - 1008de \\ &\quad + 15b^2c^2de - 15b^2c^3e - 80cde^2 + 40d^2e^2 \\ &\quad + 40c^2e^2 + 5184, \\ p_4 &= 4b^2cd - 3b^2d^2 - 4b^2c^2 + 22ce - 22de + 261. \end{aligned}$$

Since $d_1 = 7$, $d_2 = 8$, $d_3 = 6$ and $d_4 = 4$ for this system, the system $Q(\mathbf{x})$ in (3) yields $d_1 \times d_2 \times d_3 \times d_4 = 7 \times 8 \times 6 \times 4 = 1344$ paths for the homotopy in (2). However, the system (4) has only 16 isolated zeros. Consequently, a major fraction of the paths are extraneous. Sending out 1344 paths in search of 16 solutions is a highly wasteful computation.

The choice of $Q(\mathbf{x})$ in (3) to solve the system $P(\mathbf{x}) = \mathbf{0}$ requires an amount of computational effort proportional to $d_1 \times \cdots \times d_n$ and roughly, proportional to the size of the system. We would like to derive methods for solving deficient systems for which the computational effort is instead proportional to the actual number of solutions.

3. Linear Homotopy

3.1 m-Homogeneous Structure

The solution paths of (2) which do not proceed to a solution of $P(\mathbf{x}) = \mathbf{0}$ in $\mathbb{C}^n$ diverge to infinity. If the system (2) is viewed in projective space

$$\mathbb{P}^n = \{(x_0,\ldots,x_n) \in \mathbb{C}^{n+1} \backslash (0,\ldots,0)\} / \sim$$

where the equivalent relation "$\sim$" is given by $\mathbf{x} \sim \mathbf{y}$ if $\mathbf{x} = c\mathbf{y}$ for some nonzero $c \in \mathbb{C}$, the diverging paths simply proceed to a "point at infinity" in $\mathbb{P}^n$.

For a polynomial $f(x_1,\ldots,x_n)$ of degree d, denote the associated homogeneous polynomial by

$$\widetilde{f}(x_0,x_1,\ldots,x_n) = x_0^d f\left(\frac{x_1}{x_0},\ldots,\frac{x_n}{x_0}\right).$$

The solutions of $f(\mathbf{x}) = 0$ at infinity are those zeros of $\widetilde{f}$ in $\mathbb{P}^n$ with $x_0 = 0$ and the remaining zeros of $\widetilde{f}$ with $x_0 \neq 0$ are the solutions of $f(\mathbf{x}) = 0$ in $\mathbb{C}^n$ when x_0 is set to be 1.

In [33], an interesting approach to reduce the number of extraneous paths is developed, using the concept of m-homogeneous structure.

The complex n-space $\mathbb{C}^n$ can be naturally embedded in the projective space $\mathbb{P}^n$. Similarly, the space

$\mathbb{C}^{k_1} \times \cdots \times \mathbb{C}^{k_m}$ can be naturally embedded in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$. A point $(\mathbf{y}_1, \ldots, \mathbf{y}_m)$ in $\mathbb{C}^{k_1} \times \cdots \times \mathbb{C}^{k_m}$ with $\mathbf{y}_j = (y_1^{(j)}, \ldots, y_{k_j}^{(j)})$, $j = 1, \ldots, m$ corresponds to a point $(\mathbf{z}_1, \ldots, \mathbf{z}_m)$ in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ with $\mathbf{z}_j = (z_0^{(j)}, \ldots, z_{k_j}^{(j)})$ and $z_0^{(j)} = 1$, $j = 1, \ldots, m$. The set of such points in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ is usually called the *affine space* in this setting. The points in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ with at least one $z_0^{(j)} = 0$ are called the *points at infinity*.

Let f be a polynomial in the n variables $x_1, \ldots, x_n$. If we partition the variables into m groups $\mathbf{y}_1 = (x_1^{(1)}, \ldots, x_{k_1}^{(1)}), \mathbf{y}_2 = (x_1^{(2)}, \ldots, x_{k_2}^{(2)}), \ldots, \mathbf{y}_m = (x_1^{(m)}, \ldots, x_{k_m}^{(m)})$ with $k_1 + \cdots + k_m = n$ and let d_i be the degree of f with respect to $\mathbf{y}_i$ (more precisely, to the variables in $\mathbf{y}_i$), then we can define its m-homogenization as

$$\widetilde{f}(\mathbf{z}_1, \ldots, \mathbf{z}_m) = (z_0^{(1)})^{d_1} \times \cdots \times (z_0^{(m)})^{d_m} f(\mathbf{y}_1 / z_0^{(1)}, \ldots, \mathbf{y}_m / z_0^{(m)}).$$

This polynomial is homogeneous with respect to each $\mathbf{z}_j = (z_0^{(j)}, \ldots, z_{k_j}^{(j)})$, $j = 1, \ldots, m$. Here $z_i^{(j)} = x_i^{(j)}$, for $i \neq 0$. Such a polynomial is said to be *m-homogeneous*, and $(d_1, \ldots, d_m)$ is called the m-homogeneous degree of f.

In general, for an m-homogeneous system

$$\begin{aligned} \widetilde{p}_1(\mathbf{z}_1, \ldots, \mathbf{z}_m) &= 0, \\ &\vdots \\ \widetilde{p}_n(\mathbf{z}_1, \ldots, \mathbf{z}_m) &= 0, \end{aligned} \tag{5}$$

in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ with $\widetilde{p}_j$ having m-homogeneous degree $(d_1^{(j)}, \ldots, d_m^{(j)})$, $j = 1, \ldots, n$, with respect to $(\mathbf{z}_1, \ldots, \mathbf{z}_m)$, then the m-homogeneous Bézout number d of the system with respect to $(\mathbf{z}_1, \ldots, \mathbf{z}_m)$ is the coefficient of $\alpha_1^{k_1} \times \cdots \times \alpha_m^{k_m}$ in the product

$$\begin{aligned} &(d_1^{(1)} \alpha_1 + \cdots + d_m^{(1)} \alpha_m) \\ &\quad \times (d_1^{(2)} \alpha_1 + \cdots + d_m^{(2)} \alpha_m) \cdots (d_1^{(n)} \alpha_1 + \cdots + d_m^{(n)} \alpha_m) \end{aligned} \tag{6}$$

[37]. The classical Bézout Theorem says the system (5) has no more than d isolated solutions, counting multiplicities, in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$. The main theorem stated in [33] is,

Theorem 3.1. Let $Q(\mathbf{x})$ be a system of polynomials chosen to have the same m-homogeneous form as $P(\mathbf{x})$ with respect to certain partition of the variables $(x_1, \ldots, x_n)$. Assume $Q(\mathbf{x}) = \mathbf{0}$ has exactly the Bézout number of nonsingular solutions with respect to this partition, and let

$$H(\mathbf{x}, t) = (1 - t) c Q(\mathbf{x}) + t P(\mathbf{x})$$

where $t \in [0, 1]$ and $c \in \mathbb{C}^* = \mathbb{C} \backslash \{0\}$. If $c = re^{i\theta}$ for some positive $r \in \mathbb{R}$, then for all but finitely many θ, **Properties 1** and **2** hold.

For **Property 0**, it is easy to construct linear product forms according to the m-homogeneous structure that is easy to solve. Notice that when the number of isolated zeros of $Q(\mathbf{x})$, having the same m-homogeneous structure of $P(\mathbf{x})$ with respect to a given partition of variables $(x_1, \ldots, x_n)$, reaches the corresponding Bézout number, then no other solutions of $Q(\mathbf{x}) = \mathbf{0}$ exist at infinity.

The m-homogeneous Bézout number is apparently highly sensitive to the chosen partition: different ways of partitioning the variables produce different Bézout numbers. By using Theorem 3.1, we usually follow the Bézout number (with respect to the chosen partition of variables) of paths to obtain all the isolated zeros of $P(\mathbf{x})$. In order to minimize the number of paths need to be followed and hence avoid more extraneous paths, it's critically important to find a partition which gives the lowest Bézout number possible. In [43], an organized exhaustively searching algorithm for this purpose was designed. By using this algorithm, the lowest possible Bézout number for the system can be determined.

When exhaustively searching for the partitioning of the variables, there are several ways to speed up the process. For instance, as we sequentially test partitioning in search for minimal Bézout numbers, we can use the smallest one found so far to cut short unfavorable partitioning. Since the degrees are all nonnegative, the Bézout number is a sum of nonnegative degree products. If at any time the running subtotal exceeds the current minimal Bézout number, the calculation can be aborted and testing of the next partitioning can proceed. This can save a substantial amount of computation during an exhaustive search.

While the number of partitioning to be tested grows rapidly with the number of variables, the exhaustive search can be easily parallelized by subdividing the tree of partitioning and distributing these branches to multiple processors for examination. Thus, continuing advances in both raw computer speed and in parallel machines will make progressively larger problems feasible.

3.2 Cheater's Homotopy

A method called the *cheater's homotopy* has been developed, see [25] (a similar procedure can be found in [34]), to deal with the problem when the system $P(\mathbf{c}, \mathbf{x}) = \mathbf{0}$ is asked to be solved for several different values of the coefficients $\mathbf{c} = (c_1, \ldots, c_M)$.

The idea of the method is to theoretically establish **Properties 1** and **2** by deforming a sufficiently generic system and then to "cheat" on **Property 0** by using a preprocessing step. The amount of computation of preprocessing step may be large, but is amortized among the several solving characteristics of the problem.

The classical homotopy using the start system $Q(\mathbf{x}) = \mathbf{0}$ in (3) produces total degree d paths, be-

ginning at d trivial starting points. This may produce a big amount of extraneous paths. The cheater's homotopy continuation approach begins by solving $P(\bar{\mathbf{c}},\mathbf{x}) = \mathbf{0}$ with *randomly-chosen* complex coefficients $\bar{\mathbf{c}} = (\bar{c}_1,\ldots,\bar{c}_M)$; let X^* be the set of solutions obtained. No work is saved there, since d paths need to be followed. However, the elements of the set X^* are the seeds for the remainder of the process. In the future, for each choice of coefficients $\mathbf{c} = (c_1,\ldots,c_M)$ for which the system $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$ needs to be solved, we use the homotopy continuation method to follow a straight-line homotopy from the system with coefficient $\bar{\mathbf{c}}$ to the system with coefficient $\mathbf{c}$. We follow the paths beginning at the elements of X^*. Thus **Property 0**, that of having trivial-available starting points, is satisfied. The fact that **Properties 1** and **2** are also satisfied is the content of Theorem 3.2 below. Thus for each fixed $\mathbf{c}$, all isolated solutions of $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$ lie at the end of X^* smooth homotopy paths beginning at the seeds in X^*. After the foundational step of finding the seeds, the complexity of all further solving of $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$ is proportional to the number of solutions X^*, rather than the total degree d.

Furthermore, this method requires no a priori analysis of the system. The first preprocessing step of finding the seeds establishes a sharp theoretical upper bound on the number of isolated solutions as a by-product of the computation; further solving of the system uses the optimal number of paths to be followed.

We earlier characterized a successful homotopy continuation method as having three properties: triviality, smoothness, and accessibility. Given an arbitrary system of polynomial equations, it is not too hard (through generic perturbations) to find a family of systems with the last two properties. The problem is that one member of the family must be trivial to solve, or the path-following cannot get started. The idea of the cheater's homotopy is simply to "cheat" on this part of the problem, and run a preprocessing step (the computation of the seeds X^*) which gives us the triviality property in a roundabout way. Thus the name, the "cheater's homotopy".

A statement of the theoretical result we need follows. Let

$$\begin{aligned} p_1(c_1,\ldots,c_M,x_1,\ldots,x_n) &= 0,\\ &\vdots \\ p_n(c_1,\ldots,c_M,x_1,\ldots,x_n) &= 0, \end{aligned} \tag{7}$$

be a system of polynomial equations in the variables $c_1,\ldots,c_M,x_1,\ldots,x_n$. Write $P(\mathbf{c},\mathbf{x}) = (p_1(\mathbf{c},\mathbf{x}),\ldots,p_n(\mathbf{c},\mathbf{x}))$. For each choice of $\mathbf{c} = (c_1,\ldots,c_M)$ in $\mathbb{C}^M$, this is a system of polynomial equations in the variables $x_1,\ldots,x_n$. Let d be the total degree of the system for a generic choice of $\mathbf{c}$.

Theorem 3.2. Let $\mathbf{c}$ belong to $\mathbb{C}^M$. There exists an open dense full-measure subset U of $\mathbb{C}^{n+M}$ such that for $(b_1^*,\ldots,b_n^*,c_1^*,\ldots,c_M^*) \in U$, the following holds:

(a) The set X^* of solutions $\mathbf{x} = (x_1,\ldots,x_n)$ of

$$\begin{aligned} q_1(x_1,\ldots,x_n) &= p_1(c_1^*,\ldots,c_M^*,x_1,\ldots,x_n) + b_1^* = 0\\ &\vdots \\ q_n(x_1,\ldots,x_n) &= p_n(c_1^*,\ldots,c_M^*,x_1,\ldots,x_n) + b_n^* = 0 \end{aligned} \tag{8}$$

consists of d_0 isolated points, for some $d_0 \le d$.

(b) The smoothness and accessibility properties hold for the homotopy

$$\begin{aligned} H(\mathbf{x},t) = P((1-t)c_1^* + tc_1,\ldots,(1-t)c_M^* + tc_M,x_1,\ldots,x_n)\\ + (1-t)b^* \end{aligned} \tag{9}$$

where $\mathbf{b}^* = (b_1^*,\ldots,b_n^*)$. It follows that every solution of $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$ is reached by a path beginning at a point of X^*.

A proof of Theorem 3.2 can be found in [25]. The theorem is used as part of the following procedure. Let $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$ as in (7) denote the system to be solved for various values of the coefficients $\mathbf{c}$.

Cheater's Homotopy Procedure:

(1) Choose complex number $(b_1^*,\ldots,b_n^*,c_1^*,\ldots,c_M^*)$ at random, and use the classical homotopy continuation method to solve $Q(\mathbf{x}) = \mathbf{0}$ in (8). Let d_0 denote the number of solutions found (This number is bounded above by the total degree d). Let X^* denote the set of d_0 solutions.

(2) For each new choice of coefficients $c = (c_1,\ldots,c_M)$, follow the d_0 paths defined by $H(\mathbf{x},t) = \mathbf{0}$ in (9), beginning at the points in X^*, to find all solutions of $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$.

4. Nonlinear Homotopy

Let the polynomial system $P(\mathbf{c},\mathbf{x}) = (p_1(\mathbf{c},\mathbf{x}),\ldots,p_n(\mathbf{c},\mathbf{x}))$ be given with

$$\begin{aligned} p_1(\mathbf{c},\mathbf{x}) &= \sum_{\mathbf{a}\in S_1} c_{1,\mathbf{a}}\mathbf{x}^{\mathbf{a}},\\ &\vdots \\ p_n(\mathbf{c},\mathbf{x}) &= \sum_{\mathbf{a}\in S_n} c_{n,\mathbf{a}}\mathbf{x}^{\mathbf{a}}, \end{aligned} \tag{10}$$

where for each $j = 1,\ldots,n$, S_j is a fixed subset of $\mathbb{N}^n$ with cardinals $k_j = \#S_j$, and $c_{j,\mathbf{a}} \in \mathbb{C}^*$ for $\mathbf{a} \in S_j$. Here $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$. The coefficients $\mathbf{c} = (c_{j,\mathbf{a}})$ can be taken as a set of $m \equiv k_1 + \cdots + k_n$ variables.

To solve $P(\mathbf{c},\mathbf{x}) = \mathbf{0}$ in (10) by nonlinear homotopies was originally suggested by S. T. Yau. However, a major breakthrough for this method appeared in

the article by Hubert and Sturmfels [20], based on an alternative proof of Bernshteín's Theorem from combinatorial point of view. The method is therefore known as the *polyhedral* homotopy method.

In algebraic geometry, one can show that when S_j, $j=1,\dots,n$ in (10) are fixed, then there is an open dense set U in $\mathbb{C}^n$ with full measure such that if the coefficients $\mathbf{c}=(c_{j,\mathbf{a}})$ of (10) is in U ($P(\mathbf{c},\mathbf{x})$ is called *in general position* in this situation) then the corresponding $P(\mathbf{c},\mathbf{x})$ has the same number of isolated zeros, counting multiplicities, in $(\mathbb{C}^*)^n$ [19]. Otherwise, the number of isolated zeros of the corresponding system in $(\mathbb{C}^*)^n$ is bounded above by this number. Bernshteín proved [5] this number is the *mixed volume* of $(Q_1,\dots,Q_n)$ where $Q_j=conv\,S_j$, for $j=1,\dots,n$. In this note, we shall sketch the nonlinear method in pursuing all isolated zeros of $P(\mathbf{c},\mathbf{x})$ in $(\mathbb{C}^*)^n$ only. This can easily be converted to finding isolated zeros of $P(\mathbf{c},\mathbf{x})$ in $\mathbb{C}^n$ by theorems proved in [28, 36].

Since the coefficients $\mathbf{c}=(c_{j,\mathbf{a}})$ of a target system $P(\mathbf{c},\mathbf{x})=(p_1(\mathbf{c},\mathbf{x}),\dots,p_n(\mathbf{c},\mathbf{x}))$ in (10) are given, we simply write $P(\mathbf{x})=(p_1(\mathbf{x}),\dots,p_n(\mathbf{x}))$ for this system. Let $Q(\mathbf{x})=(q_1(\mathbf{x}),\dots,q_n(\mathbf{x}))$ where

$$Q(\mathbf{x})=\begin{cases} q_1(\mathbf{x}) &= \sum_{\mathbf{a}\in S_1}\bar{c}_{1,\mathbf{a}}\mathbf{x}^{\mathbf{a}},\\ &\vdots\\ q_n(\mathbf{x}) &= \sum_{\mathbf{a}\in S_n}\bar{c}_{n,\mathbf{a}}\mathbf{x}^{\mathbf{a}}, \end{cases} \tag{11}$$

be a system in (10) in which all the coefficients $\bar{\mathbf{c}}=(\bar{c}_{j,\mathbf{a}})$ are choosing at random. Since $U\subset\mathbb{C}^n$ is open and dense with full measure, this system is in general position with probability one. We will solve this system first, and then consider the linear homotopy

$$H(\mathbf{x},t)=(1-t)cQ(\mathbf{x})+tP(\mathbf{x})=\mathbf{0}\quad\text{for generic}\quad c\in\mathbb{C}^*. \tag{12}$$

Recall that, by Theorem 3.2, **Properties 1** and **2** (Smoothness and Accessibility) hold for this homotopy, and because all the isolated solutions of $Q(\mathbf{x})=\mathbf{0}$ are known, **Property 0** also holds. Therefore, every isolated zero of $P(\mathbf{x})$ lies at the end of a homotopy path of $H(\mathbf{x},t)=\mathbf{0}$, emanating from an isolated solution of $Q(\mathbf{x})=\mathbf{0}$.

To solve $Q(\mathbf{x})=\mathbf{0}$ in (11), let t be a new complex variable and consider the polynomial system $\hat{Q}(\mathbf{x},t)=(\hat{q}_1(\mathbf{x},t),\dots,\hat{q}_n(\mathbf{x},t))$ in the $n+1$ variables $(\mathbf{x},t)$ given by

$$\hat{Q}(\mathbf{x},t)=\begin{cases} \hat{q}_1(\mathbf{x},t) &= \sum_{\mathbf{a}\in S_1}\bar{c}_{1,\mathbf{a}}\mathbf{x}^{\mathbf{a}}t^{w_1(\mathbf{a})},\\ &\vdots\\ \hat{q}_n(\mathbf{x},t) &= \sum_{\mathbf{a}\in S_n}\bar{c}_{n,\mathbf{a}}\mathbf{x}^{\mathbf{a}}t^{w_n(\mathbf{a})}, \end{cases} \tag{13}$$

where each $w_j:S_j\to\mathbb{R}$ for $j=1,\dots,n$ is a generic function in the sense that its images are generically chosen numbers. The function $w=(w_1,\dots,w_n)$ may be considered as a *generic lifting* which lifts S_j to its graph

$$\hat{S}_j=\{\hat{\mathbf{a}}=(\mathbf{a},w_j(\mathbf{a}))\,|\,\mathbf{a}\in S_j\},\quad j=1,\dots,n.$$

We now regard $\hat{Q}(\mathbf{x},t)=\mathbf{0}$ as a homotopy, known as the *polyhedral homotopy*, defined on $(\mathbb{C}^*)^n\times[0,1]$ with target system $\hat{Q}(\mathbf{x},1)=Q(\mathbf{x})$, and the zero set of this homotopy is made up of k, the mixed volume of $(\text{conv}\,S_1,\dots,\text{conv}\,S_n)$, homotopy paths, say, $\mathbf{x}^{(1)}(t),\dots,\mathbf{x}^{(k)}(t)$. Since the coefficients $\bar{\mathbf{c}}=(\bar{c}_{j,\mathbf{a}})$ of $Q(\mathbf{x})$ are choosing at random, by a standard application of generalized Sard's Theorem [1], all those homotopy paths are smooth with no bifurcations. Therefore, both **Property 1** (Smoothness) and **Property 2** (Accessibility) given in Section 1 hold for this homotopy. However, at $t=0$, the solutions of $\hat{Q}(\mathbf{x},0)=\mathbf{0}$ are not known, so those homotopy paths can not get started because their starting points $\mathbf{x}^{(1)}(0),\dots,\mathbf{x}^{(k)}(0)$ can not be identified. This obstacle can be resolved by the following device.

Let $\hat{\alpha}=(\alpha,1)$ with $\alpha=(\alpha_1,\dots,\alpha_n)\in\mathbb{R}^n$ satisfy the following condition:

(A) *There exists a collection of pairs* $\{\mathbf{a}_1,\mathbf{a}_1'\}\subset S_1,\dots,\{\mathbf{a}_n,\mathbf{a}_n'\}\subset S_n$, *where* $\{\mathbf{a}_1-\mathbf{a}_1',\dots,\mathbf{a}_n-\mathbf{a}_n'\}$ *is linearly independent, such that*

$$\begin{aligned}&\langle\hat{\mathbf{a}}_j,\hat{\alpha}\rangle=\langle\hat{\mathbf{a}}_j',\hat{\alpha}\rangle\\ &\langle\hat{\mathbf{a}},\hat{\alpha}\rangle>\langle\hat{\mathbf{a}}_j,\hat{\alpha}\rangle\quad\text{for}\quad\mathbf{a}\in S_j\backslash\{\mathbf{a}_j,\mathbf{a}_j'\}.\end{aligned}$$

Here, $\langle\cdot,\cdot\rangle$ stands for the usual inner product in the Euclidean space. For such $\hat{\alpha}=(\alpha,1)$, let

$$\begin{aligned} y_1 &= t^{-\alpha_1}x_1,\\ &\vdots\\ y_n &= t^{-\alpha_n}x_n. \end{aligned} \tag{14}$$

In short, we write $\mathbf{y}=t^{-\alpha}\mathbf{x}$ where $\mathbf{y}=(y_1,\dots,y_n)$ and $\mathbf{y}t^{\alpha}=\mathbf{x}$. Notice that when $t=1$, $\mathbf{y}=\mathbf{x}$. With this transformation and $\mathbf{a}=(a_1,\dots,a_n)\in\mathbb{N}^n$,

$$\begin{aligned}\mathbf{x}^{\mathbf{a}} &= x_1^{a_1}\cdots x_n^{a_n},\\ &= (y_1t^{\alpha_1})^{a_1}\cdots(y_nt^{\alpha_n})^{a_n}\\ &= y_1^{a_1}\cdots y_n^{a_n}\,t^{\alpha_1a_n+\cdots+\alpha_na_n}\\ &= \mathbf{y}^{\mathbf{a}}t^{\langle\mathbf{a},\alpha\rangle},\end{aligned} \tag{15}$$

and $\hat{q}_j(\mathbf{x},t)$ of $\hat{Q}(\mathbf{x},t)$ in (13) becomes

$$\begin{aligned}\hat{q}_j(\mathbf{y}t^{\alpha},t) &= \sum_{\mathbf{a}\in S_j}\bar{c}_{j,\mathbf{a}}\mathbf{y}^{\mathbf{a}}t^{\langle\mathbf{a},\alpha\rangle}t^{w_j(\mathbf{a})}\\ &= \sum_{\mathbf{a}\in S_j}\bar{c}_{j,\mathbf{a}}\mathbf{y}^{\mathbf{a}}t^{\langle(\mathbf{a},w_j(\mathbf{a}),(\alpha,1))\rangle}\\ &= \sum_{\mathbf{a}\in S_j}\bar{c}_{j,\mathbf{a}}\mathbf{y}^{\mathbf{a}}t^{\langle\hat{\mathbf{a}},\hat{\alpha}\rangle}.\end{aligned} \tag{16}$$

Let

$$\beta_j=\min_{\mathbf{a}\in S_j}\langle\hat{\mathbf{a}},\hat{\alpha}\rangle,\quad j=1,\dots,n, \tag{17}$$

and consider the homotopy

$$H^\alpha(\mathbf{y},t) = (h_1^\alpha(\mathbf{y},t),\ldots,h_n^\alpha(\mathbf{y},t)) = \mathbf{0} \tag{18}$$

on $(\mathbb{C}^*)^n \times [0,1]$ where for $j = 1,\ldots,n$

$$\begin{aligned} h_j^\alpha(\mathbf{y},t) &= t^{-\beta_j}\hat{q}_j(\mathbf{y}t^\alpha,t) = \sum_{\mathbf{a}\in S_j} \bar{c}_{j,\mathbf{a}}\mathbf{y}^{\mathbf{a}}t^{\langle\hat{\mathbf{a}},\hat{\alpha}\rangle-\beta_j} \\ &= \sum_{\substack{\mathbf{a}\in S_j \\ \langle\hat{\mathbf{a}},\hat{\alpha}\rangle=\beta_j}} \bar{c}_{j,\mathbf{a}}\mathbf{y}^{\mathbf{a}} + \sum_{\substack{\mathbf{a}\in S_j \\ \langle\hat{\mathbf{a}},\hat{\alpha}\rangle>\beta_j}} \bar{c}_{j,\mathbf{a}}\mathbf{y}^{\mathbf{a}}t^{\langle\hat{\mathbf{a}},\hat{\alpha}\rangle-\beta_j}. \end{aligned} \tag{19}$$

This homotopy retains most of the properties of the homotopy $\hat{Q}(\mathbf{x},t) = \mathbf{0}$; in particular, both **Properties 1** (Smoothness) and **Property 2** (Accessibility) remain valid and

$$H^\alpha(\mathbf{y},1) = \hat{Q}(\mathbf{y},1) = Q(\mathbf{y}) = Q(\mathbf{x}), \tag{20}$$

since when $t = 1, \mathbf{y} = \mathbf{x}$. From condition (A), for each $j = 1,\ldots,n$, $\langle\hat{\mathbf{a}}_j,\hat{\alpha}\rangle = \langle\hat{\mathbf{a}}'_j,\hat{\alpha}\rangle = \beta_j$ and $\langle\hat{\mathbf{a}},\hat{\alpha}\rangle > \beta_j$ for $\mathbf{a} \in S_j\backslash\{\mathbf{a}_j,\mathbf{a}'_j\}$, hence, the start system $H^\alpha(\mathbf{y},0) = \mathbf{0}$ of this homotopy equals

$$H^\alpha(\mathbf{y},0) = \begin{cases} h_1^\alpha(\mathbf{y},0) = \displaystyle\sum_{\substack{\mathbf{a}\in S_1 \\ \langle\hat{\mathbf{a}},\hat{\alpha}\rangle=\beta_1}} \bar{c}_{1,\mathbf{a}}\mathbf{y}^{\mathbf{a}} = \bar{c}_{1,\mathbf{a}_1}y^{\mathbf{a}_1} + c_{1,\mathbf{a}'_1}y^{\mathbf{a}'_1} = 0, \\ \quad\vdots \\ h_n^\alpha(\mathbf{y},0) = \displaystyle\sum_{\substack{\mathbf{a}\in S_n \\ \langle\hat{a},\hat{\alpha}\rangle=\beta_n}} \bar{c}_{n,\mathbf{a}}\mathbf{y}^{\mathbf{a}} = \bar{c}_{n,\mathbf{a}_n}y^{\mathbf{a}_n} + c_{n,\mathbf{a}'_n}y^{\mathbf{a}'_n} = 0. \end{cases} \tag{21}$$

Such system is known as the *binomial system*, and its isolated solutions in $(\mathbb{C}^*)^n$ are constructively available [8].

Proposition 4.1. The above binomial system

$$\begin{aligned} \bar{c}_{1,\mathbf{a}_1}\mathbf{y}^{\mathbf{a}_1} + \bar{c}_{1,\mathbf{a}'_1}\mathbf{y}^{\mathbf{a}'_1} &= 0, \\ &\vdots \\ \bar{c}_{n,\mathbf{a}_n}\mathbf{y}^{\mathbf{a}_n} + \bar{c}_{n,\mathbf{a}'_n}y^{\mathbf{a}'_n} &= 0, \end{aligned} \tag{22}$$

has

$$k_\alpha := \left|\det\begin{pmatrix} \mathbf{a}_1 - \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}_n - \mathbf{a}'_n \end{pmatrix}\right| \tag{23}$$

nonsingular isolated solutions in $(\mathbb{C}^*)^n$.

Now, by (20), following paths $\mathbf{y}(t)$ of the homotopy $H^\alpha(\mathbf{y},t) = \mathbf{0}$ in (18) that emanate from k_α, as in (23), isolated zeros in $(\mathbb{C}^*)^n$ of the binomial start system $H^\alpha(\mathbf{y},0) = \mathbf{0}$ in (21), yields k_α isolated zeros of the system $Q(\mathbf{x})$ in (11) when $t = 1$. Moreover, different $\hat{\alpha} = (\alpha,1) \in \mathbb{R}^{n+1}$ that satisfy condition (A) will induce different homotopies $H^\alpha(\mathbf{y},t) = \mathbf{0}$ in (19) and following corresponding paths of those different homotopies will reach different sets of isolated zeros of $Q(\mathbf{x})$. One can show that those different sets of isolated zeros of $Q(\mathbf{x})$ are actually disjoint from each other, and they hence provide $\sum_\alpha k_\alpha$ isolated zeros of $Q(\mathbf{x})$ in total. On the other hand, with rather complicated scheme, one can also show that $\sum_\alpha k_\alpha$ is actually equal to the total number of isolated zero of $Q(\mathbf{x})$ in $(\mathbb{C}^*)^n$.

Remark 1. A key step in the procedure described above for solving system $Q(\mathbf{x})$ is the search for all those vectors $\hat{\alpha} = (\alpha,1) \in \mathbb{R}^{n+1}$ and their associated collection of pairs $\{\mathbf{a}_1,\mathbf{a}'_1\} \subset S_1,\ldots,\{\mathbf{a}_n,\mathbf{a}'_n\} \subset S_n$ that satisfy condition (A). This step turns out to be the main bottleneck in the polyhedral homotopy method. A solution to this problem is the so called *mixed cell enumeration* in which mixed volume of the system becomes a by-product [13, 21, 30].

Remark 2. Actually, those two steps, solving $Q(\mathbf{x}) = \mathbf{0}$ by the polyhedral homotopy in the first place and then using the homotopy (12) to solve $P(\mathbf{x}) = \mathbf{0}$, can be combined in one step (see [20] for details).

5. Solutions of Positive Dimension

For polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}),\ldots,p_n(\mathbf{x}))$, positive dimensional components of the solution set of $P(\mathbf{x}) = \mathbf{0}$ are a common occurrence. Sometimes they are an unpleasant side show [38] that happens with a system generated using a model for which only the isolated nonsingular solutions are of interest; and sometimes, the positive dimensional solution components are of primary interest. In either case, dealing with positive dimensional solution components, is usually computationally difficult.

Solution set of $P(\mathbf{x}) = \mathbf{0}$ in $\mathbb{C}^n$ is known as an algebraic set. An algebraic set can always be decomposed into a finite union of *irreducible components* which are algebraic sets that cannot be further decomposed into nontrivial union of algebraic sets. Putting aside formal definition with technical terms, by a *generic point* of an irreducible component Y of the solution set of $P(\mathbf{x}) = \mathbf{0}$, it usually means a point of Y which has no special properties not possessed by the whole component Y. Numerically, it is modeled by a point in Y with random coordinates.

By Noether's normalization theorem combined with Bertini's theorem, it can be shown that if X is of pure k-dimensional, *i.e.*, any irreducible component of X is k-dimensional, then a generic $(n-k)$-dimensional affine subspace of $\mathbb{C}^n$ (which can be identified with $\mathbb{C}^{n-k}$) meets X at a set χ of d isolated points. Those are generic points of X. The set χ is refereed to as a set of **witness points** for X. The number d is called the *degree* of X and denoted deg X. A generic $(n-k)$-dimensional affine subspace $L^{(n-k)} \approx \mathbb{C}^{n-k}$, called *slicing plane*, can be given by

$$\begin{aligned} \lambda_{11}x_1 + \cdots + \lambda_{1n}x_n &= \lambda_1 \\ \vdots \qquad\qquad &\vdots\ \ \vdots \\ \lambda_{k1}x_1 + \cdots + \lambda_{kn}x_n &= \lambda_k \end{aligned}$$

with all the λ's being random complex numbers. Thus, the existence of isolated solutions of the system

$$\begin{aligned} p_1(x_1,\ldots,x_n) &= 0 \\ &\vdots \qquad\qquad \vdots \\ p_n(x_1,\ldots,x_n) &= 0 \\ \lambda_{11}x_1+\cdots+\lambda_{1n}x_n &= \lambda_1 \\ \vdots \qquad\qquad &\vdots \quad \vdots \\ \lambda_{k1}x_1+\cdots+\lambda_{kn}x_n &= \lambda_k \end{aligned} \tag{24}$$

warrants the existence of k-dimensional components of the solution set of the original system $P(\mathbf{x}) = (p_1(\mathbf{x}),\ldots,p_n(\mathbf{x})) = \mathbf{0}$. Furthermore, the set of isolated solutions of (24) contains at least one generic point of each irreducible component of dimension k of X.

System (24) is overdetermined. With extra parameters $z_1,\ldots,z_k$, one may consider the square system

$$\mathcal{E}_k(\mathbf{x},z_1,\ldots,z_k) = \begin{cases} p_1(\mathbf{x})+\lambda_{11}z_1+\cdots+\lambda_{1j}z_k = 0 \\ \qquad\qquad\qquad\qquad \vdots \\ p_n(\mathbf{x})+\lambda_{n1}z_1+\cdots+\lambda_{nj}z_k = 0 \\ \\ z_1+\lambda_{11}x_1+\cdots+\lambda_{1n}x_n = \lambda_1 \\ \qquad\vdots \qquad\qquad \vdots \quad \vdots \\ z_k+\lambda_{k1}x_1+\cdots+\lambda_{kn}x_n = \lambda_k \end{cases}$$

and search for isolated solutions with $z_1 = z_2 = \cdots = z_k = 0$. When X is a pure k-dimensional *reduced* component of system $P(\mathbf{x}) = (p_1(\mathbf{x}),\ldots,p_n(\mathbf{x})) = \mathbf{0}$, then the witness points in χ are nonsingular roots of $P(\mathbf{x})$ restricted to $L^{(n-k)}$. (When X is not reduced, a slightly richer structure is needed.) The data structure $W := (\chi, L^{(n-k)}, P(\mathbf{x}))$ is called a **witness set** for X, and when X is an irreducible component, then W is recognized as an **irreducible witness** set.

Using witness sets, one can make numerical sense out of what it means to find the solution set of a system of polynomials $P(\mathbf{x}) = (p_1(\mathbf{x}),\ldots,p_n(\mathbf{x})) = \mathbf{0}$. Let the decomposition of the entire solution set V of $P(\mathbf{x}) = \mathbf{0}$ be the nested union:

$$V := \bigcup_{j=0}^{j_0} V_j := \bigcup_{j=0}^{j_0} \bigcup_{i\in I_j} V_{ji} \tag{25}$$

where j_0 is the dimension of the solution set V, the V_j is the union of all j-dimensional components, the V_{ji} are the irreducible components of dimension j, and the index sets I_j are finite and possibly empty. We wish to find a *numerical irreducible decomposition* that mirrors the irreducible decomposition of (25). This means finding a collection of witness sets W_j for the j-dimensional components V_j, which are themselves decomposed into *irreducible witness sets* W_{ji} for the irreducible components V_{ji}, namely,

$$W := \bigcup_{j=0}^{j_0} W_j := \bigcup_{j=0}^{j_0} \bigcup_{i\in I_j} W_{ji}. \tag{26}$$

Encapsulating the whole process (for details see [39]) is the foundation of the subject *Numerical Algebraic Geometry* which was started in [38], and its name coined.

6. Parallel Computing

Modern scientific computing is marked by the advent of vector and parallel computers and search for algorithms that are to a large extend parallel in nature. A great advantage of the homotopy continuation algorithm for solving polynomial systems is that it is to a large degree parallel, in the sense that each isolated zero can be computed independently. In this respect, it stands in contrast to the highly serial algebraic elimination methods, which use resultants or Gröbner bases. On the other hand, to attain more computing resources for solving larger polynomial systems, the parallelization of the homotopy method becomes inevitably essential.

The landscape of computation hardware has seen extremely active development in recent years making available a wide spectrum of exciting new technologies. First, developments in new processor design and network technology have allowed supercomputers and computer clusters to grow larger and faster than ever. Second, new ideas such as cycle-scavenging and grid computing has led to the creation of virtual supercomputers out of large numbers of individual computers around the globe. Another exciting development is the advent of parallel computing on GPUs (Graphical Processing Units). While originally designed to handle 2D and 3D graphics rendering only, over the years GPUs has become sufficiently sophisticated to handle a much wider range of problems. Highly parallel by design, GPUs are more efficient than general purpose CPUs in carrying out a range of complex algorithms. Living in such interesting times is exciting and daunting. We must rise up to the challenge, fully incorporate all these cutting-edge parallel computing technology, and solve larger and larger polynomial systems.

As mentioned in **Remark 1** above, a key step when the polyhedral homotopy is employed for solving system $Q(\mathbf{x})$ in (11) is the search for all those vectors $\hat{\alpha} = (\alpha, 1) \in \mathbb{R}^{n+1}$ and their associated collection of pairs $\{\mathbf{a}_1, \mathbf{a}_1'\} \subset S_1, \ldots, \{\mathbf{a}_n, \mathbf{a}_n'\} \subset S_n$ that satisfy condition (A). This step of *mixed cell enumeration* is seemly

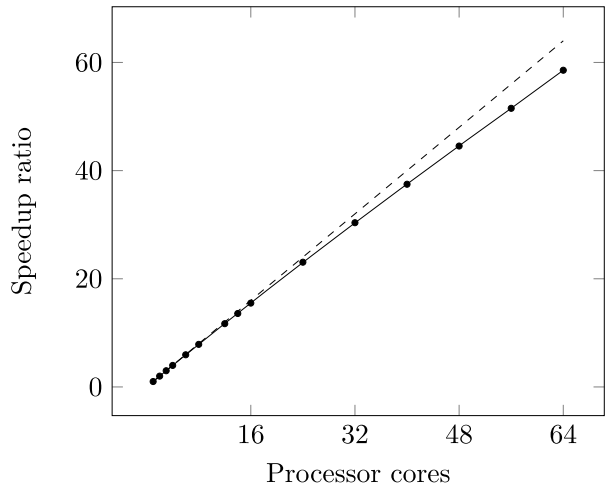

Figure 1. Speedup ratio achieved on a 64 core system (AMD Opteron with 512 GB memory) for the cyclic-15 [6] problem showing close to n-fold linear speedups for up to 64 cores. The speedup is computed in comparison with the fastest serial implementations published: MixedVol-2.0 [21] and DEMiCs [29].

quite serial. While the "path tracking" part of the polyhedral homotopy continuation method is *pleasantly parallel*, the mixed cell enumeration process is potentially a major bottleneck in terms of parallel scalability. Based on the idea of reformulating the problem into a graph-theoretic search problem, a fully parallel mixed cell enumeration algorithm that is efficient, robust, and highly scalable has been developed in Hom4PS-3 [9].

On multi-core systems, the implementation of the parallel mixed cell enumeration algorithm in Hom4PS-3, based on Intel TBB, has achieved remarkable efficiency and scalability. Nearly n-fold linear speedups scalable up to 64 processor cores have been observed in experiments on standard test suite problems. Figure 1 shows the speedup ratio observed on the standard benchmark problem cyclic-15 [6].

Modern shared-memory systems with a large number of processor cores usually adopt a Non-Uniform Memory Access [18], or NUMA, architecture in which each processor core can access all the available memory but potentially at different speeds depending on the relative closeness between the core and memory. Figure 2 shows the "memory-processor topology" of a NUMA system that consists of 8 nodes. Each node contains 4 processor cores as well as their "local" memory which they can access at full speed. The edges between nodes indicate the direct connectedness between nodes and determines the speed at which processor cores on one node can access memory on other nodes. For instance a processor on node 1 can access the memory on node 2 at a slower rate than it could access its local memory on node 1. The same core can access memory on node 3 at a even slower rate due to the minimum two jumps required (through node 2 or 4). Similarly there are at least three jumps between node 1 and node 7. Consequently, that processor core would have the slowest memory access to memory on node 7.

Table 1. The memory access speedup (with errors within ±0.05) and overall reductions in run time over the basic algorithm observed in experiments with a few large systems in standard test suites on a NUMA system consisting of 8 node each having 8 quad-core AMD Opteron processor with a total of 256 cores. Experiments marked by '' only used 4 out of 8 nodes (128 of the 256 cores) due to smaller size*

System	Mem. acc. speedup	Overall reduction
cyclic-14*	1.40	4.9%
cyclic-15*	2.40	8.2%
cyclic-16*	4.50	9.5%
fivebody	17.55	33.2%
vortex-6	19.95	34.5%

The technique developed in Hom4PS-3 substantially improves the memory access time on NUMA systems. As shown in Table 1, in experiments using standard benchmark problems: the cyclic family [6], the five-body central configuration problem (fivebody) [3, 16, 23], and the 6-vortex problem [15, 17], approximately 1.5× to 20× speedup in memory access time[1] have been observed which resulted in 5% to 35% overall speedups.

While the above referenced NUMA architecture allows shared-memory systems to scale to tens or even more than 100 processor cores, their scalability is still limited by the inherently high cost. Larger systems that contain several hundreds or even thousands of cores generally take the form of distributed-memory systems in which nodes do not directly share memory spaces but instead communicate with one another by passing messages. The parallel algorithm can be extended to distributed-memory systems including *computer clusters* in which nodes are connected by dedicated high speed network. In such distributed-memory systems, a *master-worker* model is chosen to extend the parallel algorithm. In this model, the "master" runs on a single node in the system. It first populates its own task pool. This initial task pool is then divided into equal portions and sent to each of the remaining nodes, called "workers", as seeds for exploration until its task pool becomes empty. At the

[1] The memory access time is approximated by using the "memory access latency" provided by the Intel VTune software which closely correlates to the actual memory access time which is generally difficult to measure. For best accuracy, all CPU caches were disabled when measuring memory access latency.

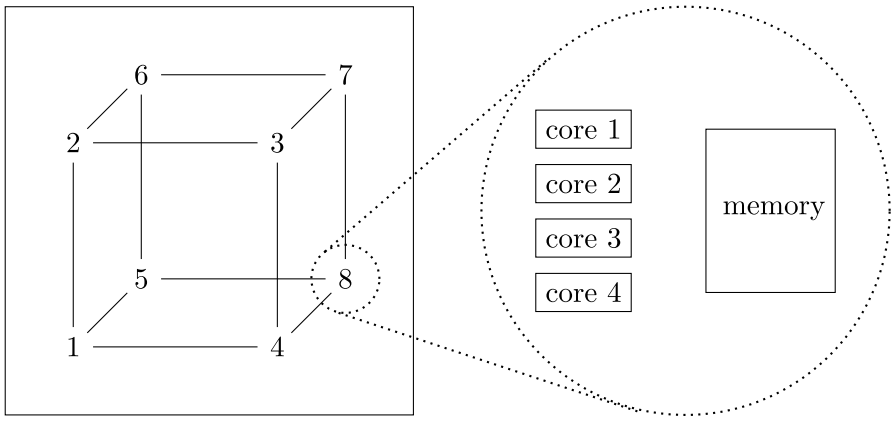

Figure 2. An example of a NUMA node structure.

end each worker would have a collection of mutually exclusive mixed cells. These are then passed back to the master to form a final set of mixed cells.

The Message Passing Interface, or MPI, is a specification that allows nodes to communicate with each other in a cluster. Though not sanctioned by any major standards body, MPI has became a *de facto* standard for scientific computation on computer clusters. In Hom4PS-3 [9], this protocol is used for the communication between the master and workers. The implementation exhibits a great scalability on clusters having between 32 and 200 nodes. It is expected that the speedup ratio cannot get close to those achieved on a multi-core system (as shown in Figure 1) due to the inherently higher cost in communication. However it is possible to scale to many more processors cores than on multi-core or NUMA system. For example, the speedup ratios achieved using multiple nodes in a cluster for the fivebody (five body central configuration) problem [3, 16, 23] is shown in Figure 3.

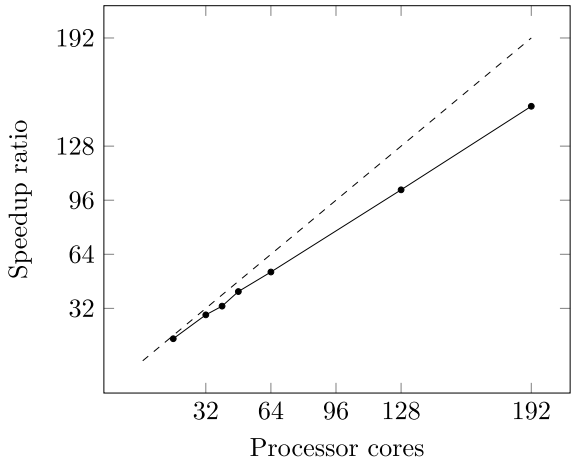

Figure 3. Speedup ratios achieved by the distributed-memory variation of the parallel algorithm for mixed cell enumeration over the fastest serial implementations MixedVol-2.0 [21] and DEMiCs [29]. Measurements are done in a cluster containing up to 192 processor cores.

As one tries to solve larger and more challenging systems of polynomial equations, the computational power needed may very well exceed that of a single computer or cluster. The distributed computing technology (the *cloud computing*) has allowed a very loosely coupled set of computers connected to each other via slower and less reliable networks, such as the Internet, to collaborate with each other and act as a single supercomputer. Adopting a distributed model, the computation will not be affected by the sudden failure of a single node, and nodes can join or leave the computation process at almost any point of time. An experimental version of the mixed cell enumeration algorithm was implemented in Hom4PS-3 using a client/server model based on TCP/IP protocol [41]. This implementation consists of a client side program and a server side program. One server is running at any time, and it is responsible for dividing problems into groups of tasks and sends the groups to clients via the network upon request. Multiple client programs can run simultaneously. A client can run on any computer that is connected to the Internet. Each client will continuously request tasks from the server via the network until the server can no longer reply with new tasks. For each task received, the client performs and then sends the result back to the server. This approach has been used on the VortexAC6 [15, 17] problem which is notorious for its sheer size (30 equations of 30 variables with a total degree of 2^{30}). While the computation would take a single CPU core several months to 1 year based on our estimate, with around 300 individual nodes provided by Amazon EC2 with a total of 1200 CPU cores, the computation can be finished in just a few hours! A great potential exists for other

parallel computing techniques. We remain very optimistic that this amazing advantage can be carried over to more general problems with new knowledge on those new techniques.

References

[1] R. Abraham and J. Robbin (1967), *Transversal mapping and flows*, W. A. Benjamin (New York, Amsterdam).

[2] E. L. Allgower and K. Georg (1990), *Numerical Continuation Methods, an Introduction*, Springer Series in Comput. Math., Vol. 13, Springer-Verlag (Berlin Heidelberg, New York).

[3] A. Albouy and V. Kaloshin (2012), "Finiteness of central configurations of five bodies in the plane", *Ann. Math.*, **176**(1), 535–588.

[4] D. J. Bates, J. D. Hauenstein, A. J. Sommese and C. W. Wampler (2013), "Numerically solving polynomial systems with Bertini", SIAM (Philadephia, PA).

[5] D. N. Bernshteín (1975), "The number of roots of a system of equations", *Functional Analysis and Appl.*, **9**(3), 183–185. Translated from *Funktsional. Anal. i Prilozhen.*, **9**(3), 1–4.

[6] G. Björck and R. Fröberg (1991), "A faster way to count the solutions of inhomogeneous systems of algebraic equations with applications to cyclic n-roots", *J. of Symbolic Computation,* **12**(3), 329–336.

[7] B. Buchberger (1985), "Gröbner basis: An algorithmic method in polynomial ideal theory", In *Multidimensional System Theory* (N. K. Bose, ed.), D. Reidel Publishing Company (Dordrecht Boston Lancaster), 184–232.

[8] T. Chen and T. Y. Li (2014), "Solutions to systems of binomial equations", *Annales Mathematicae Silesianae.*, **28**, 7–34.

[9] T. Chen, T.-L. Lee, and T. Y. Li (2014), "Hom4PS-3: A parallel numerical solver for systems of polynomial equations based on polyhedral homotopy continuation methods", In *Mathematical Software – ICMS 2014*, (H. Hong and C. Yap, eds.), Lecture Notes in Computer Science, Vol. 8592, Springer (Berlin Heidelberg), 183–190.

[10] S. N. Chow, J. Mallet-Paret, and J. A. Yorke (1979), "Homotopy method for locating all zeros of a system of polynomials", In *Functional differential equations and approximation of fixed points* (H. O. Peitgen and H. O. Walther, eds.), Lecture Notes in Mathematics, Vol. 730, Springer-Verlag (Berlin Heidelberg), 77–88.

[11] F. J. Drexler (1977), "Eine Methode zur Berechnung sämtlicher Lösungen von Polynomgleichungssystemen", *Numer. Math.*, **29**, 45–58.

[12] C. B. Garcia and W. I. Zangwill (1979), "Finding all solutions to polynomial systems and other systems of equations", *Math. Programming*, **16**, 159–176.

[13] T. Gao and T. Y. Li (2003), "Mixed volume computation for semi-mixed systems", *Discrete Comput. Geom.*, **29**(2), 257–277.

[14] T. Gunji, S. Kim, M. Kojima, A. Takeda, K. Fujisawa and T. Mizutani, (2004), "PHoM – a Polyhedral homotopy Continuation Method", *Computing*, **73**, 53–77.

[15] M. Hampton and R. Moeckel (2009), "Finiteness of stationary configurations of the four-vortex problem", *Transactions of the American Mathematical Society*, **361**(3), 1317–1332.

[16] M. Hampton and A. Jensen (2011), "Finiteness of spatial central configurations in the five-body problem". *Celestial Mechanics and Dynamical Astronomy*, **109**(4), 321–332.

[17] H. von Helmholtz (1858), "Uber integrale der hydrodynamischen gleichungen welche den wirbelbewegungen entsprechen", *Crelle's Journal fur Mathematic*, **55**, 25–55.

[18] M. Herlihy and N. Shavit (2012), "The art of multiprocessor programming", Elsevier.

[19] B. Huber (1996), *Solving sparse polynomial systems*, Ph. D. thesis, Department of Mathematics, Cornell University.

[20] B. Huber and B. Sturmfels (1995), "A polyhedral method for solving sparse polynomial systems", *Math. Comp.*, **64**, 1541–1555.

[21] T. L. Lee and T. Y. Li (2011), "Mixed volume computation in solving polynomial systems", *Contemporary Math*, **556**, 97–112.

[22] T. L. Lee, T. Y. Li, and C. H. Tsai (2008), "HOM4PS-2.0 : a software package for solving polynomial systems by the polyhedral homotopy continuation method", *Computing*, **83** (2), 109–133.

[23] T.-L. Lee and M. Santoprete (2009), "Central configurations of the five-body problem with equal masses" *Celestial Mechanics and Dynamical Astronomy*, **104**(4), 369–381.

[24] T. Y. Li (1983), "On Chow, Mallet-Paret and Yorke homotopy for solving systems of polynomials", *Bull. Inst. Math. Acad. Sinica*, **11**, 433–437.

[25] T. Y. Li, T. Sauer and J. A. Yorke (1989), "The cheater's homotopy: an efficient procedure for solving systems of polynomial equations", *SIAM J. Numer. Anal.*, **26**, 1241–1251.

[26] T. Y. Li, and C. H. Tsai (2009), "HOM4PS-2.0para: Parallelization of HOM4PS-2.0 for solving polynomial systems", *Parallel Computing*, **35** (4), 226–238,

[27] T. Y. Li and X. Wang (1994), "Higher order turning points", *Appl. Math. Comput.*, **64**, 155–166.

[28] T. Y. Li and X. Wang (1997), "The BKK root count in $\mathbb{C}^n$", *Math. Comp.*, **65**, 1477–1484.

[29] T. Mizutani and A. Takeda (2008), "DEMiCs: A Software Package for Computing the Mixed Volume Via Dynamic Enumeration of all Mixed Cells", In *Software for Algebraic Geometry* (M. Stillman, J. Verschelde, and N. Takayama, eds.), The IMA Volumes in Mathematics and its Applications, Vol. 148, Springer, 59–79.

[30] T. Mizutani, A. Takeda and M. Kojima (2007), "Dynamic enumeration of all mixed cells", *Discrete Comput. Geom.*, **37**, 351–367.

[31] A. P. Morgan (1986), "A homotopy for solving polynomial systems", *Appl. Math. Comput.*, **18**, 173–177.

[32] A. P. Morgan (1987), *Solving polynomial systems using continuation for engineering and scientific problems*, Prentice-Hall (Englewood Cliffs, N. J.).

[33] A. P. Morgan and A. J. Sommese (1987), "A homotopy for solving general polynomial systems that respect m-homogeneous structures", *Appl. Math. Comput.*, **24**, 101–113.

[34] A. P. Morgan and A. J. Sommese (1989), "Coefficient-parameter polynomial continuation", *Appl. Math. Comput.*, **29**, 123–160. Errata: *Appl. Math. Comput.*, **51**, 207 (1992).

[35] B. Mourrain, *The handbook of polynomial systems*, available at http://www.inria.fr/safir/POL/index.html.

[36] J. M. Rojas and X. Wang (1996), "Counting affine roots of polynomial systems via pointed Newton polytopes", *J. of Complexity*, **12**(2), 116–133.

[37] I. R. Shafarevich (1977), *Basic Algebraic Geometry*, Springer-Verlag (New York).

[38] A. J. Sommese and C. W. Wampler (1996), "Numerical algebraic geometry", In *The Mathematics of Numerical*

Analysis (J. Renegar, M. Shub, and S. Smale, eds.), Lectures in Applied Mathematics, Vol. 32, 749–763. Proceedings of the 1995 AMS-SIAM Summer Seminar in Applied Mathematics, Park City, Utah.

[39] A. J. Sommese and C. W. Wampler (2005), *The numerical solution of systems of polynomials arising in engineering and science*, World Scientific Publishing Co. (Singapore).

[40] C. Traverso (1997), "The PoSSo test suite examples", Availabe at: http://www.fr/safir/POL/index.html.

[41] W. R. Stevens (1994), "*TCP/IP Illustrated, Vol. 1: The Protocols*", Addison-Wesley Professional.

[42] J. Verschelde (1999), "PHCPACK: A general-purpose solver for polynomial systems by homotopy continuation", *ACM Tran. Math. Softw.*, **25**, 251–276.

[43] C. W. Wampler (1992), "Bézout number calculations for multi-homogeneous polynomial systems", *Appl. Math. Comput.*, **51**, 143–157.

[44] A. H. Wright (1985), "Finding all solutions to a system of polynomial equations", *Math. Comp.*, **44**, 125–133.

[45] W. Zulener (1988), "A simple homotopy method for determining all isolated solutions to polynomial systems", *Math. Comp.*, **50**, 167–177.