J. Symbolic Computation (1999) **11**, 1–17

Toric Newton Method for Polynomial Homotopies

Jan Verschelde †

† *Michigan State University, Department of Mathematics*

East Lansing, MI 48824-1027, U.S.A.

(Received 24 September 1999)

This paper defines a generalization of Newton's method to deal with solution paths defined by polynomial homotopies that lead to extremal values. Embedding the solutions in a toric variety leads to explicit scaling relations between coefficients and solutions. Toric Newton is a symbolic-numeric algorithm where the symbolic pre-processing exploits the polyhedral structures. The numerical stage uses the additional variables introduced by the homogenization to scale the components of the solution vectors to the complex unit circle. Toric Newton generates appropriate affine charts and enables to approximate the magnitude of large solutions of polynomial systems.

AMS Subject Classification. 14M25, 52B55, 65H10, 68Q40.

Keywords. polynomial system, toric varieties, Newton's method, homotopy continuation, homogeneous coordinates.

1. Introduction

If a polynomial system has a large root, we first want to estimate its magnitude before the approximation of its actual value. In the context of polyhedral homotopies we can certify (Huber and Verschelde 1998) diverging solution paths and separate those from paths leading to extremal values. The purpose of this paper is to show how polyhedral methods can approximate the magnitude of large roots. In (Huber and Sturmfels 1995) and (Verschelde *et al.* 1994) polyhedral homotopies were introduced to exploit sparsity. The number of solution paths equals the bound of Bernshteĭn's theorem. First we give some notations to state Bernshteĭn's theorems.

Denote $\mathbf{x}^{\mathbf{a}} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$. The *support* of a polynomial $f(\mathbf{x}) = \sum_{\mathbf{a} \in A} c_{\mathbf{a}} \mathbf{x}^{\mathbf{a}}$ is the set $A = \{ \mathbf{a} \in \mathbb{Z}^n \mid c_{\mathbf{a}} \neq 0 \}$. The *Newton polytope* P of f is the convex hull of its support A. The structure of a polynomial system $F(\mathbf{x}) = \mathbf{0}$ is represented by the Newton polytopes $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ spanned by the supports $\mathcal{A} = (A_1, A_2, \ldots, A_n)$ of the polynomials in the system $F = (f_1, f_2, \ldots, f_n)$. Denote $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. In (Bernshteĭn 1975) Bernshteĭn

† This work was done while the author was post-doctor at the Mathematical Sciences Research Institute, 1000 Centennial Drive, Berkeley, CA 94720-507, USA. Supported in part by NSF under Grant DMS - 9804846 at MSU. E-mail addresses: jan@math.msu.edu or jan.verschelde@na-net.ornl.gov. URL: http://www.mth.msu.edu/~jan.

0747–7171/90/000000 + 00 $03.00/0 © 1999 Academic Press Limited

proved that the number of isolated solutions in $(\mathbb{C}^*)^n$ of a polynomial system with generic coefficients equals the mixed volume $V_n(\mathcal{P})$ of its Newton polytopes $\mathcal{P}$.

In the nongeneric case we examine faces of systems. A *face* $\partial_{\mathbf{v}}P$ of a polytope P is characterized by a direction $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and spanned by $\partial_{\mathbf{v}}A = \{\ \mathbf{a} \in A \mid \text{for all } \mathbf{b} \in A : \langle \mathbf{a}, \mathbf{v} \rangle \leq \langle \mathbf{b}, \mathbf{v} \rangle\ \}$. A *face polynomial of* f is $\partial_{\mathbf{v}}f(\mathbf{x}) = \sum_{\mathbf{a} \in \partial_{\mathbf{v}}A} c_{\mathbf{a}}\mathbf{x}^{\mathbf{a}}$. A *face system of* F is $\partial_{\mathbf{v}}F(\mathbf{x}) = \mathbf{0}$ with equations $\partial_{\mathbf{v}}f_i(\mathbf{x}) = 0$, $i = 1, 2, \ldots, n$. For $V_n(\mathcal{P}) > 0$, Bernshteĭn's second theorem (Bernshteĭn 1975) states that $F(\mathbf{x}) = \mathbf{0}$ has fewer than $V_n(\mathcal{P})$ isolated solutions in $(\mathbb{C}^*)^n$ if and only if there is some $\partial_{\mathbf{v}}F(\mathbf{x}) = \mathbf{0}$ that has a solution in $(\mathbb{C}^*)^n$.

A toric variety X is an algebraic variety with an embedded (dense and open) copy T of $(\mathbb{C}^*)^n$ such that the natural $(\mathbb{C}^*)^n$-action on T extends to all of X. In section 2 we make this definition from (Fulton 1993) concrete in applying the homogenization procedure described in (Cox *et al.* 1998). The group action is used in the third section to scale the solution components to the complex unit circle, narrowing the torus $(\mathbb{C}^*)^n$ to a numerically favorable region. See (Cox 1995), (Danilov 1978), (Fulton 1993), (Gel'fand *et al.* 1994), and (Oda 1988) for more on toric varieties.

Representing diverging and badly-scaled solutions is a common struggle in the daily life of a numerical analysist. In (Morgan 1987, Chapter 5), (Meintjes and Morgan 1987), (Watson *et al.* 1987, Watson *et al.* 1997) we find coefficient and equation scaling, combined with centering, an effective and useful tool. The classical homogenization (Blum *et al.* 1997), (Li 1997), (Morgan 1987) in projective space easily leads to clustering and fails to recognize the deficiency behavior when applied to sparse polynomial systems. In the third section of this paper we exemplify this situation.

In the fourth section we define the Toric Newton method and analyze the cost of its execution. This algorithm is a symbolic-numeric approach. The Toric Newton method consists of a symbolic pre-processing stage, where the homogeneous forms are determined and the normal fans are refined to be simplicial. This stage is performed only once per support structure. The numerical stage is executed iteratively while following the solution paths and takes place in the original space of dimension n.

Toric Newton augments the robustness of polyhedral end games (Huber and Verschelde 1998) as used in numerical homotopies (Li 1997). Convergence properties of Toric Newton are addressed in the fifth section. The complexity of Newton of a general homotopy method for the dense case† has been analyzed in (Blum *et al.* 1997) and has recently been partially extended to sparse complexity models (Dedieu 1997), multi-homogeneous (Dedieu and Shub 1999a) and over-determined systems (Dedieu and Shub 1999b). Since homotopy methods for the dense case ultimately work in projective space (either implicitly or explicitly), Toric Newton provides a generalization of these methods to toric varieties. Indeed, projective space (and products thereof) are special instances of toric varieties, and the first stage of Toric Newton finds the correct generalized homogeneous coordinates (i.e., toric variety) to use for the input problem.

The Toric Newton method also parallels similar developments in providing algorithms using torus embeddings to compute sparse resultants (Rojas 1999), residues (Cattani and Dickenstein 1997), (Cattani *et al.* 1996), and Bezoutians (Roy and Szpirglas 1998). All these methods rely on polytopes (Ziegler 1995). Here we give practical evidence for the performance of the polyhedral constructions.

† Dense means that all monomials up to a certain degree appear with nonzero coefficient.

2. Newton Polytopes and Homogeneous Coordinates

If the Newton polytope lies in a hyperplane, then the coordinates of the polynomial are quasi-homogeneous. A special case of this is the Newton polytope of a homogeneous polynomial, where the hyperplane turns out to be normal to $(1,1,\ldots,1)$. The exponents of the scaler λ are given by the components of the direction that is perpendicular to the hyperplane, as expressed formally in the following proposition.

PROPOSITION 2.1. (Gel'fand *et al.* 1994, page 193) *Consider* $f(\mathbf{x}) = \sum_{\mathbf{a}\in A} c_{\mathbf{a}}\mathbf{x}^{\mathbf{a}}$. *If there exists a* $\mathbf{v} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, *and a* $d \in \mathbb{Z}$ *such that for all* $\mathbf{a} \in A, \langle \mathbf{a}, \mathbf{v}\rangle = d$, *then* $f(\lambda^{v_1}x_1, \lambda^{v_2}x_2, \ldots, \lambda^{v_n}x_n) = \lambda^d f(\mathbf{x})$.

Proof. Denote $\lambda^{\mathbf{v}}\mathbf{x} = \lambda^{v_1}x_1\lambda^{v_2}x_2\cdots\lambda^{v_n}x_n$, then $(\lambda^{\mathbf{v}}\mathbf{x})^{\mathbf{a}} = (\lambda^{v_1}x_1)^{a_1}(\lambda^{v_2}x_2)^{a_2}\cdots(\lambda^{v_n}x_n)^{a_n}$ $= \lambda^{\langle \mathbf{a},\mathbf{v}\rangle}\mathbf{x}^{\mathbf{a}}$. So, $f(\lambda^{\mathbf{v}}\mathbf{x}) = \sum_{\mathbf{a}\in A} c_{\mathbf{a}}(\lambda^{\mathbf{v}}\mathbf{x})^{\mathbf{a}} = \sum_{\mathbf{a}\in A} c_{\mathbf{a}}\lambda^{\langle \mathbf{a},\mathbf{v}\rangle}\mathbf{x}^{\mathbf{a}} = \lambda^d f(\mathbf{x})$. □

The multilinearity of the inner product provides us with a whole range of homogeneous coordinates when the polytope lies in the intersection of many hyperplanes. In what follows, we will use $[x_1 : x_2 : \cdots : x_n]$ to denote a point in $\mathbb{P}^n$. This notation means that we identify any other $[x'_1 : x'_2 : \cdots : x'_n]$ with $[x_1 : x_2 : \cdots : x_n]$ iff there is a $\lambda \in \mathbb{C}^*$ with $x_i = \lambda x'_i$ for all i. What we illustrate in the following examples is how to generalize the preceding scaling to involve several λ's.

EXAMPLE 2.1. *(projective space)* Let $[x : y : z] \in \mathbb{P}^2$ and consider $f(x,y,z) = x + y + z$. The support of f is the set $A = \{(1,0,0),(0,1,0),(0,0,1)\}$ and lies in the hyperplane $\langle \mathbf{a}, \mathbf{v}\rangle = d$ with $\mathbf{v} = (1,1,1)$, $d = 1$. Note that given x, y, and z, we are thus identifying all triples of the form $(\lambda^1 x, \lambda^1 y, \lambda^1 z)$ to the same point. ■

EXAMPLE 2.2. *(multi-projective space)* Let $[x_0 : x_1] \times [y_0 : y_1] \in \mathbb{P}^1 \times \mathbb{P}^1$ and consider $f(x_0,x_1,y_0,y_1) = x_1y_1 - x_0y_0$. The support of f is the set $A = \{(0,1,0,1),(1,0,1,0)\}$ and lies in the hyperplanes $\langle \mathbf{a}, \mathbf{v}^i\rangle = d$ with $\mathbf{v}^1 = (1,1,0,0)$, $\mathbf{v}^2 = (0,0,1,1)$, $\mathbf{v}^3 = (1,0,0,1)$, and $d = 1$. Now note that we are identifying all quadruples of the form $(\lambda^1 x_0, \lambda^1 x_1, y_0, y_1)$, and this can be observed from the normal to our first hyperplane. Similarly, our second and third hyperplane tell us to "collapse" $\{(x_0, x_1, \lambda^1 y_0, \lambda^1 y_1) \mid \lambda \in \mathbb{C}^*\}$ into a single point, and to do the same for $\{(\lambda^1 x_0, x_1, y_0, \lambda^1 y_1) \mid \lambda \in \mathbb{C}^*\}$. ■

Examples 2.1 and 2.2 are particular cases of our approach to generalizing homogeneous coordinates. The general definition is given below. Our approach is similar to that of (Cox *et al.* 1998) in that both methods define maps from variables to monomials in new variables, much like our first examples. However, the method of (Cox *et al.* 1998) "forgets" the original variables whereas ours "remembers" them. By this adaption we recognize the structure of the polytope and have explicit scaling relations between the coefficients and solutions, as needed in the toric version of Newton's method.

DEFINITION 2.1. Consider $f(\mathbf{x}) = \sum_{\mathbf{a}\in A} c_{\mathbf{a}}\mathbf{x}^{\mathbf{a}}$, with $P = \mathrm{conv}(A)$. Assume $\mathrm{vol}(P) > 0$. Let P have N facets given by N inward pointing normals $\mathbf{v}^i \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ with all components relatively prime. For every $\mathbf{v}^i$, define $d_i = \min_{\mathbf{a}\in A}\langle \mathbf{a}, \mathbf{v}^i\rangle$. The facets are spanned by

$\partial_{\mathbf{v}^i} A = \{\mathbf{a} \in A | \langle \mathbf{a}, \mathbf{v}^i \rangle = d_i\}$. The *toric homogenization map* ϕ_A introduces for every $\mathbf{v}^i$ a new variable z_i, as follows:

$$\phi_A(\mathbf{x}^\mathbf{a}) = \mathbf{x}^\mathbf{a} \prod_{i=1}^{N} z_i^{\langle \mathbf{a}, \mathbf{v}^i \rangle - d_i} \quad \text{and} \quad \phi_A \left(\sum_{\mathbf{a} \in A} c_\mathbf{a} \mathbf{x}^\mathbf{a} \right) = \sum_{\mathbf{a} \in A} c_\mathbf{a} \phi_A(\mathbf{x}^\mathbf{a}). \tag{2.1}$$

If $\mathrm{vol}(P) = 0$, then A lies in some hyperplanes $\langle \mathbf{x}, \mathbf{v}^i \rangle = c_i$, and we add $\pm \mathbf{v}^i$ to the normals of the k-faces of P, where $k = \dim(A) - 1$.

The result of the homogenization is the polynomial function $\phi_A(f)$ which takes in its evaluation as input the original n variables x_i and N additional variables z_j. Formally we obtain $\phi_A(f) : (\mathbb{C}^*)^n \times \mathbb{C}^N \to \mathbb{C} : (\mathbf{x}, \mathbf{z}) \mapsto \phi_A(f)(\mathbf{x}, \mathbf{z})$. We may allow zero values for the z-variables, because of the following property.

PROPOSITION 2.2. *Let* $\deg(f, x)$ *denote the maximal exponent of the variable* x *in the polynomial* f. *Then,*

$$\deg\left(\phi_A \left(\sum_{\mathbf{a} \in A} c_\mathbf{a} \mathbf{x}^\mathbf{a} \right), z_i \right) \geq 0 \quad \text{and} \quad \deg(\phi_A(\mathbf{x}^\mathbf{a}), z_i) \begin{cases} = & 0 \quad \text{if } \mathbf{a} \in \partial_{\mathbf{v}^i} A \\ > & 0 \quad \text{if } \mathbf{a} \notin \partial_{\mathbf{v}^i} A \end{cases} . \tag{2.2}$$

Proof. The directions perpendicular to the facets are given by inward pointing normals, i.e.: for a facet normal $\mathbf{v}^i$: $d_i \leq \langle \mathbf{a}, \mathbf{v}^i \rangle$, for all $\mathbf{a} \in A$. Therefore all exponents of z_i of ϕ_A are nonnegative. □

This homogenization corresponds geometrically to letting the exponent of z_i be the lattice distance from $\mathbf{a}$ to the ith facet of P. If we collect for every vertex monomial the exponents of the new z-variables and store these vectors row-wise in a matrix, then we obtain the vertex-facet incidence matrix (Ziegler 1995) of the polytope P. The (inner) normal fan Σ_P collects all cones that are generated by (inner) normals to a face of P (Ziegler 1995).

EXAMPLE 2.3. *(running example)* Consider $f(x_1, x_2) = c_{10} x_1 + c_{01} x_2 + c_{20} x_1^2 + c_{12} x_1 x_2^2$. The Newton polytope has four facets, with inward pointing normals given by $\mathbf{v}^1 = (1, 1)$, $\mathbf{v}^2 = (0, 1)$, $\mathbf{v}^3 = (-2, -1)$, and $\mathbf{v}^4 = (1, -1)$. In Figure 1, the homogenizing along $\mathbf{v}^4$ is illustrated.

The homogenized polynomial is $\phi_A(f)(\mathbf{x}, \mathbf{z}) = c_{10} z_3^2 z_4^2 x_1 + c_{01} z_2 z_3^2 x_2 + c_{20} z_1 z_4^3 x_1^2 + c_{12} z_1^2 z_2^2 x_1 x_2^2$. ■

The inward pointing normals to the facets tell us how to identify the equivalence classes of points and lead to our new homogeneous coordinates. The point of Proposition 2.3 is that the homogenized polynomial $\phi_A(f)$ is multi-homogeneous relative to as many different $\mathbb{C}^*$ actions as there are facets of the polytope spanned by A.

PROPOSITION 2.3. *Consider the facet normal* $\mathbf{v}^i$ *with corresponding* d_i.
Denote $\lambda^{\mathbf{v}^i} \mathbf{xz} = (\lambda^{v_1^i} x_1, \lambda^{v_2^i} x_2, \ldots, \lambda^{v_n^i} x_n, z_1, \ldots, z_{i-1}, \lambda^{-1} z_i, z_{i+1}, \ldots, z_n)$.
Then $\phi_A(f)(\lambda^{\mathbf{v}^i} \mathbf{xz}) = \lambda^{d_i} \phi_A(f)(\mathbf{x}, \mathbf{z})$.

Proof. First consider one monomial $\mathbf{x}^\mathbf{a}$:

$$\phi_A(\mathbf{x}^\mathbf{a}) \mid_{(\mathbf{x}, \mathbf{z}) = \lambda^{\mathbf{v}^i} \mathbf{xz}} \tag{2.3}$$

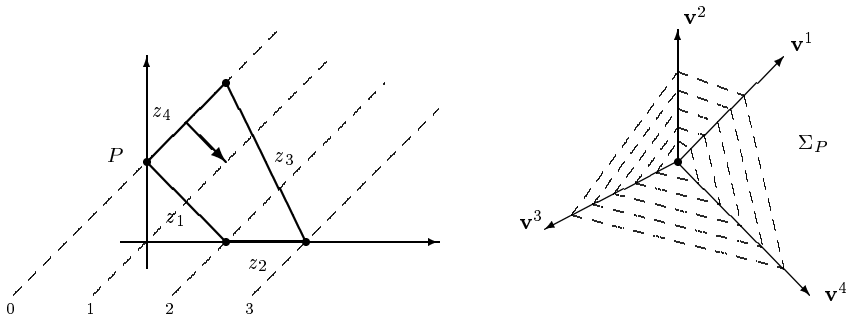


Figure 1. On the left we see that the exponents of z_4 are lattice distances relative to the fourth facet of P with normal $\mathbf{v}^4$. On the right the normal fan Σ_P of P is displayed.

$$= \mathbf{x}^{\mathbf{a}} \prod_{j=1}^{N} z_j^{\langle \mathbf{a}, \mathbf{v}^j \rangle - d_j} \,|_{(\mathbf{x},\mathbf{z}) = \lambda^{\mathbf{v}^i} \mathbf{x}\mathbf{z}} \tag{2.4}$$

$$= (\lambda^{v_1^i} x_1 \lambda^{v_2^i} x_2 \cdots \lambda^{v_n^i} x_n)^{\mathbf{a}} \left(\prod_{j=1, \neq i}^{N} z_j^{\langle \mathbf{a}, \mathbf{v}^j \rangle - d_j} \right) (\lambda^{-1} z_i)^{\langle \mathbf{a}, \mathbf{v}^i \rangle - d_i} \tag{2.5}$$

$$= \lambda^{\langle \mathbf{a}, \mathbf{v}^i \rangle} \mathbf{x}^{\mathbf{a}} \lambda^{-\langle \mathbf{a}, \mathbf{v}^i \rangle + d_i} \prod_{j=1}^{N} z_j^{\langle \mathbf{a}, \mathbf{v}^j \rangle - d_j} \tag{2.6}$$

$$= \lambda^{d_i} \phi_A(\mathbf{x}^{\mathbf{a}})(\mathbf{x}, \mathbf{z}) \tag{2.7}$$

The multilinearity of ϕ_A completes the proof. □

The transition from toric to affine space is immediate by taking $\lambda = z_i$ in every $\lambda^{\mathbf{v}^i} \mathbf{x}\mathbf{z}$.

EXAMPLE 2.4. *(Example* 2.3 *continued.)* The equivalence relations corresponding to the facets are given by

$$\begin{array}{ccccccccccccc}
(& \lambda^1 x_1 & , & \lambda^1 x_2 & , & \lambda^{-1} z_1 & , & z_2 & , & z_3 & , & z_4 &) \\
(& \lambda^0 x_1 & , & \lambda^1 x_2 & , & z_1 & , & \lambda^{-1} z_2 & , & z_3 & , & z_4 &) \\
(& \lambda^{-2} x_1 & , & \lambda^{-1} x_2 & , & z_1 & , & z_2 & , & \lambda^{-1} z_3 & , & z_4 &) \\
(& \lambda^1 x_1 & , & \lambda^{-1} x_2 & , & z_1 & , & z_2 & , & z_3 & , & \lambda^{-1} z_4 &)
\end{array} \tag{2.8}$$

for the respective normals $\mathbf{v}^i$, $i = 1, 2, 3, 4$. Observe that as $\lambda \to 0$, the x-values in affine space go either to zero or infinity. We have a multitude of equivalence relations by taking combinations, see Σ_P in Figure 1. Positive combinations of any two vectors normal to adjacent facets correspond to constructing a vector inside the cone $\sigma \in \Sigma_P$ that has these vectors among its generators.

The equivalence relations in (2.8) are not independent because the linear relations between the normals impose relations between them. The scaling actions are defined by the group $(\mathbb{C}^*)^N$ modulo these imposed relations, and this quotient group is isomorphic to $(\mathbb{C}^*)^n$. Intuitively we can see this isomorphism when we "reverse" the equivalence relations so that the i-th one has $\lambda^{\delta_{ij}} x_j$ as its j-th entry, with $\delta_{ij} = 1$ when $i = j$ and $\lambda_{ij} = 0$ when $i \neq j$. This reversal shows that there are only n independent relations. ■

So far we only dealt with one polynomial. For a polynomial system $F(\mathbf{x}) = \mathbf{0}$ we have a tuple of n Newton polytopes $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ spanned by the supports $\mathcal{A} = (A_1, A_2, \ldots, A_n)$ of the polynomials in $F = (f_1, f_2, \ldots, f_n)$. To homogenize $F(\mathbf{x}) = \mathbf{0}$ we extend the toric homogenization map of Definition 2.1 using n distinct sets of additional variables. Note that since every facet normal to the Minkowski sum of the polytopes P_i is a facet normal to one of the P_i's, we do not need to construct this Minkowski sum explicitly.

DEFINITION 2.2. Let $\mathcal{A}$ be the supports of $F(\mathbf{x}) = \mathbf{0}$. Suppose the polytope spanned by A_i has N_i facet normals. Abbreviate $\mathbf{z} = (\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^n)$. The *mixed toric homogenization map* $\phi_{\mathcal{A}}$ is

$$\phi_{\mathcal{A}}(F) : (\mathbb{C}^*)^n \times \mathbb{C}^{N_1} \times \mathbb{C}^{N_2} \times \cdots \times \mathbb{C}^{N_n} \to \mathbb{C}^n : (\mathbf{x}, \mathbf{z}) \mapsto \phi_{\mathcal{A}}(F)(\mathbf{x}, \mathbf{z}), \tag{2.9}$$

with $\phi_{\mathcal{A}}(F)(\mathbf{x}, \mathbf{z}) = (\phi_{A_1}(f_1)(\mathbf{x}, \mathbf{z}^1), \phi_{A_2}(f_2)(\mathbf{x}, \mathbf{z}^2), \ \ldots, \ \phi_{A_n}(f_n)(\mathbf{x}, \mathbf{z}^n))$ and $\phi_{A_i}(f_i)$ as in (2.1). When $\mathrm{vol}(P_i) = 0$ for some i we proceed as in Definition 2.1.

For several polytopes sharing the same facet normal, we can use the same corresponding z-variable. With this economical use of z-variables, Definition 2.2 yields the same homogenization map ϕ_A as in Definition 2.1 when applied to $\mathcal{A} = (A, A, \ldots, A)$. Since the general homogenization is also determined by inward pointing normals $\mathbf{v}^i$, Propositions 2.2 and 2.3 have a direct generalization to the case of different supports and are omitted. We close this section introducing nontrivial solutions and solutions at infinity.

PROPOSITION 2.4. *Suppose A is the support of $f(\mathbf{x})$ and consider the solution $(\mathbf{x}^*, \mathbf{z}^*) \in (\mathbb{C}^*)^n \times \mathbb{C}^N$: $\phi_A(f)(\mathbf{x}^*, \mathbf{z}^*) = 0$. Let*

$$I = \{ \ i \mid z_i^* = 0, \#(\cap \partial_{\mathbf{v}^i} A) \geq 2 \ \}, \ \textit{with } z_j^* \neq 0, \ \textit{for all } j \notin I, \ \textit{and } \mathbf{v} = \sum_{i \in I} \mathbf{v}^i. \tag{2.10}$$

Then $\partial_{\mathbf{v}} f(\mathbf{x}) = 0$ has a solution in $(\mathbb{C}^*)^n$ and the solutions $\mathbf{x}^*$ of $\partial_{\mathbf{v}} f$ correspond to solutions of f with components equal to 0 and/or ∞.

Proof. We first distinguish the case when $\#I = 1$:

1. Suppose there is only one $z_i^* = 0$ with corresponding facet normal $\mathbf{v}^i$. Then all monomials in $\phi_A(f)$ whose exponent vector does not lie on the ith facet vanish when $z_i = z_i^* = 0$. For $z_j^* \neq 0$, for all $j \neq i$, the homogeneous coordinates $\lambda^{\mathbf{v}^j} \mathbf{x}^* \mathbf{z}^*$ (as in Proposition 2.3) are used successively, each time setting $\lambda = z_j^*$, for all $j \neq i$, to obtain a solution in $(\mathbb{C}^*)^n$ of $\partial_{\mathbf{v}^i} f(\mathbf{x}) = \mathbf{0}$.
2. When several z_i^*'s are zero, more monomials in the system vanish. The condition on I guarantees that there are still enough monomials left in $\partial_{\mathbf{v}} f$ to admit a solution in $(\mathbb{C}^*)^n$. Successive application of Proposition 2.3 setting $\lambda = z_j^*$, for all $j \notin I$, yields a solution in $(\mathbb{C}^*)^n$ of $\partial_{\mathbf{v}} f(\mathbf{x}) = \mathbf{0}$.

The homogeneous coordinates contain as exponents for λ the components of the normals. Any normal $\mathbf{v}$ has at least one component different from zero, call it v_k. Setting $\lambda = z_i \to 0$ to transform to affine space will lead to $x_k \to \infty$, if $v_k < 0$, or $x_k \to 0$, in case $v_k > 0$. □

In (Cox *et al.* 1998), solutions that have no z-variables equal to zero, or only in cases covered by Proposition 2.4 are called *nontrivial*. Indeed, without the conditions on I in

Proposition 2.4, all monomials could vanish. The generalization of Proposition 2.4 to polynomial systems is immediate.

PROPOSITION 2.5. *Let $F(\mathbf{x}) = \mathbf{0}$ have supports $\mathcal{A}$. The nontrivial solutions of the homogeneous system $\phi_{\mathcal{A}}(F)(\mathbf{x},\mathbf{z}) = \mathbf{0}$ with some z-variables equal to zero correspond to solutions of face systems.*

3. Scaling Coefficients and Solutions

By the homogeneity properties of $\phi_A(f)(\mathbf{x},\mathbf{z}) = \mathbf{0}$ we can scale $(\mathbf{x},\mathbf{z})$ so that every component of $\mathbf{x}$ lies on the unit circle in $\mathbb{C}^*$ while keeping the z-variables bounded.

PROPOSITION 3.1. *Suppose $\mathbf{x} \in (\mathbb{C}^*)^n$ satisfies $\phi_A(f)(\mathbf{x},\mathbf{z}) = \mathbf{0}$ for fixed $\mathbf{z}$ with $0 < z_j \leq 1$, for all $j = 1,2,\ldots,N$. Let $\mathbf{v} = (-\log|x_1|, -\log|x_2|, \ldots, -\log|x_n|)$, with* log *of base 10. Then*

1. *there exist positive coefficients $\mathbf{c} \in \mathbb{R}^n_+$ and an index set $J = (j_1, j_2, \ldots, j_n)$, $j_i \in \{1,2,\ldots,N\}$ such that $\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}^{j_i}$.*
2. *The equivalence relations*

$$\left(\lambda^{\left(-\sum_{i=1}^{n} c_i \mathbf{v}^{j_i}\right)_k} x_k \Big|_{k=1}^{n}, \ldots, \lambda^{c_i} z_{j_i} \Big|_{i=1}^{n}, \ldots \right) \tag{3.1}$$

with $\lambda = 0.1$ scale $(\mathbf{x},\mathbf{z})$ to $(\widetilde{\mathbf{x}},\widetilde{\mathbf{z}})$ with $|\widetilde{x}_k| = 1$ and $0 < \widetilde{z}_j \leq 1$.

Proof. The above statements are true because:

1. The normal fan Σ_P of P is complete (Ziegler 1995). This means that any vector of $\mathbb{R}^n$ belongs to some cone of the fan. Since any vector $\mathbf{v}$ defines some face $\partial_{\mathbf{v}} P$ of P, this also holds in case $\mathrm{vol}(P) = 0$. Whence every vector $\mathbf{v}$ can be written as a positive combination of the generators of a cone $\sigma \in \Sigma_P$.
2. $\left(-\sum_{i=1}^{n} c_i \mathbf{v}^{j_i}\right)_k = \log|x_k|$, so $\widetilde{x}_k = (0.1)^{\log|x_k|} x_k = \frac{x_k}{|x_k|}$, whence $|\widetilde{x}_k| = 1$.
   $0 < z_{j_i} \leq 1 \Rightarrow 0 < \widetilde{z}_{j_i} = (0.1)^{c_i} z_{j_i} \leq 1$, as $c_i \geq 0$. Thus $0 < \widetilde{z}_j \leq 1$. □

Note that any base b can be used in the logarithm, it suffices to replace 0.1 in Proposition 3.1 by b^{-1}.

EXAMPLE 3.1. *(toric scaling analogue to projective space)* Suppose the Newton polytope is the triangle spanned by the points $(0,0)$, $(1,0)$ and $(0,1)$. The scaling relations are

$$\begin{array}{ccccccccc} (& \lambda^1 x_1 & , & \lambda^0 x_2 & , & \lambda^{-1} z_1 & , & z_2 & , & z_3 &) \\ (& \lambda^0 x_1 & , & \lambda^1 x_2 & , & z_1 & , & \lambda^{-1} z_2 & , & z_3 &) \\ (& \lambda^{-1} x_1 & , & \lambda^{-1} x_2 & , & z_1 & , & z_2 & , & \lambda^{-1} z_3 &) \end{array} \tag{3.2}$$

for the respective facet normals $\mathbf{v}^1 = (1,0)$, $\mathbf{v}^2 = (0,1)$, and $\mathbf{v}^3 = (-1,-1)$. Since this is a *toric* analogue, we are concerned about solutions with zero components and have more z-variables than in the case of ordinary homogenization. To illustrate the scaling procedure, we consider two situations:

1 $(x_1, x_2) = (2,2)$, $\mathbf{v} = (-\log(2), -\log(2))$, $J = (3,3)$, $\mathbf{c} = (\log(2), 0)$
$\Rightarrow \widetilde{z}_3 = (0.1)^{\log(2)} z_3 = \frac{1}{2} z_3$.

2 $(x_1, x_2) = (2,1)$, $\mathbf{v} = (-\log(2), -\log(1))$, $J = (2,3)$, $\mathbf{c} = (\log(2), \log(2))$
$\Rightarrow \widetilde{z}_2 = (0.1)^{\log(2)} z_2 = \frac{1}{2} z_2$ and $\widetilde{z}_3 = (0.1)^{\log(2)} z_3 = \frac{1}{2} z_3$.

In the first case it suffices to use one of the homogeneous coordinates, while in the second case a combination is constructed. ■

EXAMPLE 3.2. *(Example* 2.3 *continued.)* Here we also consider two illustrations:

1 $(x_1, x_2) = (\frac{1}{2}, 2)$, $\mathbf{v} = (-\log(\frac{1}{2}), -\log(2))$, $J = (4,4)$, $\mathbf{c} = (\log(2), 0)$
$\Rightarrow \widetilde{z}_4 = (0.1)^{\log(2)} z_4 = \frac{1}{2} z_4$.

2 $(x_1, x_2) = (2,2)$, $\mathbf{v} = (-\log(2), -\log(1))$, $J = (2,3)$, $\mathbf{c} = (\log(2), \log(2))$
$\Rightarrow \widetilde{z}_3 = (0.1)^{2\log(2)} z_3 = \frac{1}{4} z_3$ and $\widetilde{z}_4 = (0.1)^{3\log(2)} z_4 = \frac{1}{8} z_4$.

The scaling relation is constructed by decomposing $\mathbf{v}$ in Σ_P, see Figure 1. ■

In the general case when the system has an n-tuple of supports $\mathcal{A}$, we decompose $\mathbf{v} = -\log|\mathbf{x}|$ in every component of $\Sigma_{\mathcal{P}} = (\Sigma_{P_1}, \Sigma_{P_2}, \ldots, \Sigma_{P_n})$ to compute new values of the z-variables.

Propositions 2.5 and 3.1 allow us to give a concise proof of Bernshteĭn's second theorem formulated below as Theorem 3.1 using a homotopy argument. Recall that Bernshteĭn's first theorem proves that the mixed volume $V_n(\mathcal{P})$ is an upper bound for the number of isolated solutions, while his second theorem gives genericity conditions to determine when this bound is an exact root count. A *homotopy* is a one-parameter family of polynomial systems that contains at one end the system $F(\mathbf{x}) = \mathbf{0}$ one wants to solve, and at the other end a generic system $F^{(0)}(\mathbf{x}) = \mathbf{0}$ whose solutions are known. The canonical example of a homotopy is

$$H(\mathbf{x}, t) = (1-t)F^{(0)}(\mathbf{x}) + tF(\mathbf{x}) = \mathbf{0}, \quad t \in [0,1], \tag{3.3}$$

which defines paths of solutions starting at the solutions of $F^{(0)}(\mathbf{x}) = \mathbf{0}$, for $t = 0$, and ending at the solutions of $F(\mathbf{x}) = \mathbf{0}$, at $t = 1$.

If we choose the system $F^{(0)}(\mathbf{x}) = \mathbf{0}$ so that it has the same Newton polytopes $\mathcal{P}$ as F but with sufficiently generic complex coefficients, then Bernshteĭn's first theorem tells us that $F^{(0)}(\mathbf{x}) = \mathbf{0}$ has exactly as many regular solutions in $(\mathbb{C}^*)^n$ as the mixed volume $V_n(\mathcal{P})$. Then we say that $F^{(0)}$ is *in generic position*. With this choice of start system, and a slightly more restrictive genericity condition, the homotopy (3.3) is in generic position. This means that, except for a proper algebraic subset of all possible choices of coefficients of $F^{(0)}$, the solution paths defined by the homotopy are regular for all $t \in [0,1)$, as is proven in (Li *et al.* 1989) and (Morgan and Sommese 1989).

THEOREM 3.1. *Consider* $F(\mathbf{x}) = \mathbf{0}$ *and its supports* $\mathcal{A}$ *that span the Newton polytopes* $\mathcal{P}$ *with mixed volume* $V_n(\mathcal{P}) \neq 0$.

$$\sum_{\substack{F(\mathbf{x}) = \mathbf{0} \\ \mathbf{x} \text{ isolated}}} m_{\mathbf{x}} < V_n(\mathcal{P}) \quad \Leftrightarrow \tag{3.4}$$

$$\textit{there exists a } \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}, \textit{ and a solution } \mathbf{x} \in (\mathbb{C}^*)^n : \partial_{\mathbf{v}} F(\mathbf{x}) = \mathbf{0} \qquad (3.5)$$

In (3.4), $m_{\mathbf{x}}$ *denotes the multiplicity of the isolated solution* $\mathbf{x}$ *of* $F(\mathbf{x}) = \mathbf{0}$.

Proof. Construct any homotopy $H(\mathbf{x}, t) = \mathbf{0}$ that is in generic position with respect to $\mathcal{P}$ and for which $H(\mathbf{x}, t = 1) = F(\mathbf{x})$. In homogenizing the homotopy to $\phi_{\mathcal{A}}(H)(\mathbf{x}, \mathbf{z}, t) = \mathbf{0}$ we treat the continuation parameter t as a constant.

- $\Rightarrow$ If $F(\mathbf{x}) = \mathbf{0}$ has fewer than $V_n(\mathcal{P})$ isolated solutions in $(\mathbb{C}^*)^n$, then, as $t \to 1$, either some solution paths must diverge outside $(\mathbb{C}^*)^n$ or some solutions paths must go to a component of solutions, because the homotopy is in generic position. In the first case, by Proposition 3.1, the only way to diverge is that some $z_i \to 0$. Then, according to Proposition 2.5, we have a solution in $(\mathbb{C}^*)^n$ of some face system. In the second case, the solution component allows at least one free variable to choose in moving on that component. So we can move a free variable outside $(\mathbb{C}^*)^n$ for which some z_i is zero. Then we can again apply Proposition 2.5.
- $\Leftarrow$ If there is a face system that has a solution $\mathbf{x} \in (\mathbb{C}^*)^n$, then in the corresponding tuple of z-variables for which $\phi_{\mathcal{A}}(H)(\mathbf{x}, \mathbf{z}, 1) = \mathbf{0}$, $\exists i$: $z_i = 0$. Since the homotopy is in generic position, there is at least one solution path that ended at $(\mathbf{x}, \mathbf{z})$. By Proposition 2.5, the transition to affine coordinates shows that the solution path that ended at this couple $(\mathbf{x}, \mathbf{z})$ diverged to a solution that has components at infinity and/or zero components. So $F(\mathbf{x}) = \mathbf{0}$ must have fewer than $V_n(\mathcal{P})$ isolated solutions in $(\mathbb{C}^*)^n$. □

4. The Static Scaling in Projective Newton

Let us recall Newton's method in projective space (see (Li 1997)). We assume we have a system $H = (h_1, h_2, \ldots, h_n)$ of homogeneous polynomials: $h_i(\lambda \mathbf{z}) = \lambda^{\deg(h_i)} h_i(\mathbf{z})$ for $i = 1, 2, \ldots, n$ with $\lambda \mathbf{z} = (\lambda z_0, \lambda z_1, \ldots, \lambda z_n)$ the usual scaling relation. For a solution at infinity: $z_0 = 0$.

ALGORITHM 4.1. *(Newton's method in projective space)*

Input: $(H, \mathbf{z})$. *(Homogeneous system, approximate root)*

Output: $\mathbf{z}$. *refined approximation*

$i : |z_i| = \max_{j=1}^{n} |z_j|$; *maximal component*

$\mathbf{z} := \left(\underbrace{\frac{z_0}{z_i}, \ldots, \frac{z_{i-1}}{z_i}}, 1, \underbrace{\frac{z_{i+1}}{z_i}, \ldots, \frac{z_n}{z_i}} \right)$; *scale the approximation*

$\mathbf{x} := \left(\underbrace{z_1, \ldots, z_{i-1}} \ , \ \underbrace{z_{i+1}, \ldots, z_n} \right)$; *switch to an affine chart*

$F := H|_{z_i = 1}$; *select evaluation map*

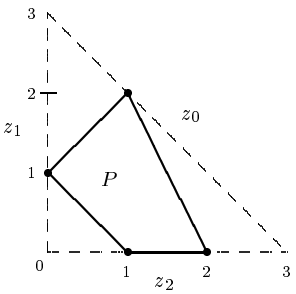
$\mathbf{x} := \mathbf{x} - D_{\mathbf{x}}^{-1} F(\mathbf{x}) F(\mathbf{x})$; *affine Newton method*

$\mathbf{z} := (x_1, \ldots, x_i, 1, x_{i+1}, \ldots, x_n)$. *refined approximation*

In the context of path following, H is a parametrized family of systems.

EXAMPLE 4.1. *(projective Newton on a sparse system)* Consider the application of pro-

jective Newton on a sparse system of polynomials. Suppose the supports of the polynomials are all equal. We see the Newton polytope displayed in Figure 2. In the homogenization the simplex is implicitly taken as Newton polytope and to every facet we associate a new variable z_i. Since in projective space one does not distinguish solutions with zero components, the original monomials are not copied.



$$f(\mathbf{x}) = c_{10}x_1 + c_{01}x_2 + c_{20}x_1^2 + c_{12}x_1x_2^2$$

$$\widetilde{f}(\mathbf{z}) = c_{10}z_0^2z_1 + c_{01}z_0^2z_2 + c_{20}z_0z_1^2 + c_{12}z_1z_2^2$$

Figure 2. On the left we see how the polytope P of f is approximated by a simplex which corresponds to the usual projective homogenization $\widetilde{f}$ given on the right.

In case of divergence, z_1 and/or z_2 go to infinity. The scaling step in Algorithm 4.1 selects $i = 1$ or $i = 2$ and z_0 strives to zero. In that case we see in Figure 2 that the monomial $z_1z_2^2$ dominates. However, a diverging path cannot go to a monomial and goes to one of the other face systems. The multi-homogeneous version of Newton's method will make the same monomial dominant. ■

EXAMPLE 4.2. *(projective Newton leads to clustering)* Consider the following simple system:

$$\begin{cases} x_1 - 10^8 = 0 \\ x_2^2 - 1 = 0 \end{cases} \tag{4.1}$$

In affine space, the solutions $(x_1, x_2) = (10^8, \pm 1)$ are well-separated. In projective space, they are represented as $(z_0, z_1, z_2) = (1, 10^8, \pm 1)$ and scaled to $(z_0, z_1, z_2) = (10^{-8}, 1, \pm 10^{-8})$, which is a pair of numerically clustered solutions. Although this example is cooked up for this special occasion, this phenomenon is observed as typical in many applications.

The clustering for this example is avoided if one applies bi-homogenization, which allows to scale the variables independently: $(x_1, x_2) = (10^8, \pm 1) \mapsto (z_{10}, z_{11}, z_{20}, z_{21}) = (10^{-8}, 1, 1, \pm 1)$. Multi-projective Newton differs from Algorithm 4.1 in the sense that one takes now the maximum for any group of variables to scale the solution vector. For this example, if $i_1 : |z_{1i_1}| = \max\{|z_{10}|, |z_{11}|\}$ and $i_2 : |z_{2i_2}| = \max\{|z_{20}|, |z_{21}|\}$, then $\mathbf{x} := (\frac{z_{1i}}{z_{1i_1}}, \frac{z_{2j}}{z_{2i_2}})$ for $i \neq i_1$ and $j \neq i_2$.

Since the polytopes of the system are both edges, and thus not of dimension two, we illustrate the toric homogenization map for this degenerate case. The normals to the first Newton polytope are $\pm(0, 1)$ and we have $\pm(0, 1)$ as inner normals to the vertices. Note that this makes the normal fan complete. For the second Newton polytope, we have

the same normals, although their roles are switched. So it suffices to introduce four new variables, corresponding to each of these four normals. ■

Although the toric homogenization is similar to the usual homogenizations for special polytopes (i.e., simplex and hypercube give respectively the projective and multi-projective transformations), the most important difference is the dynamic use of the scaling procedure of Proposition 3.1 which is better from a numerical point of view.

5. Toric Newton scales in a Dynamic Way

In this section we define the algorithms that are used to implement the Toric Newton method and we analyze their cost. When all polynomials in the system share the same Newton polytope, then the volume of this Newton polytope counts the number of solutions. We provide a triangulation of the Newton polytope in the input to Algorithm 5.1 because this information is readily at hand after the computation of the volume. The refinement of the normal fan enables to perform the scaling efficiently in Algorithm 5.2 which described the number crunching stage of Toric Newton.

ALGORITHM 5.1. *(Toric Newton – symbolic set-up)*

Input: $(A, \triangle)$.	*(support , triangulation)*
Output: H;	*homogeneous coordinates matrix*
Σ_P.	*simplicial normal fan*
1. $V :=$ Facet_Normals$(A, \triangle)$;	*extract facet normals*
2. $H :=$ Homogenize(A, V);	*relative lattice distances*
3. $\Sigma_P :=$ Cone_Subdivision(H, A, V).	*refine normal fan*

Next we analyze the cost of this algorithm. The dimension n and the cardinality $\#A$ measure the input size. The complexity of the output is bounded from below by the number of cells $\#\triangle$ in any triangulation $\triangle$ of A. Recall that we are interested in *all* isolated solutions and to count the number of roots we compute volumes. So $\#\triangle$ is a lower bound for the number of isolated solutions.

PROPOSITION 5.1. *(Cost of Step 1 in Algorithm 5.1.) Consider a triangulation $\triangle$ of the polytope P spanned by A. Then all facets of P can be enumerated in $O(\#\triangle n(n^3 + \#A))$ operations.*

Proof. The candidate facets are the facets of the simplices in the triangulation. Every simplex has $n+1$ facets. The computation of each candidate requires $O(n^3)$ arithmetical operations. To test whether this candidate normal constitutes a facet normal we have to compute inner products with the points in A and look whether the inner product is minimal for points on the facet of the simplex. This gives the bound $O(\#\triangle n(n^3+\#A))$. □

The number of arithmetic operations of Step 2 of Algorithm 5.1 equals $O(n\#V)$. Step 3 is only needed when the polytope is not simple, i.e.: when some vertices are contained in more than n facets, see the top of the polytope in Figure 3. In Step 3 we make the fan simplicial subdividing the normal cones. One method to subdivide normal cones is to cut corners (see (Gel'fand *et al.* 1994, page 190) and (Ziegler 1995, pages 54–55)) at

vertices that have more than n facets, as illustrated in Figure 3. We refer to (Fulton 1993, Section 2.6) for more on refinements of fans in connection with toric varieties.

For a normal cone $\sigma \in \Sigma_P$ at the vertex $\mathbf{a}$ we take the average $\overline{\mathbf{v}}$ of the facet normals that generate σ. Determine c such that the hyperplane $\langle \mathbf{x}, \overline{\mathbf{v}} \rangle = c$ separates $\mathbf{a}$ from the rest of the vertices that span the polytope. The normal cones τ to the intersection points of the hyperplane with the edges of the polytope refine the normal cone σ. The recursive structure of the normal fan Σ_P is illustrated at the right of Figure 3. In the decomposition of a vector $\mathbf{v}$ in a subdivision of σ, we replace the average vector $\overline{\mathbf{v}}$ in terms of the other generators of σ, so that we obtain a decomposition that uses only facet normals.

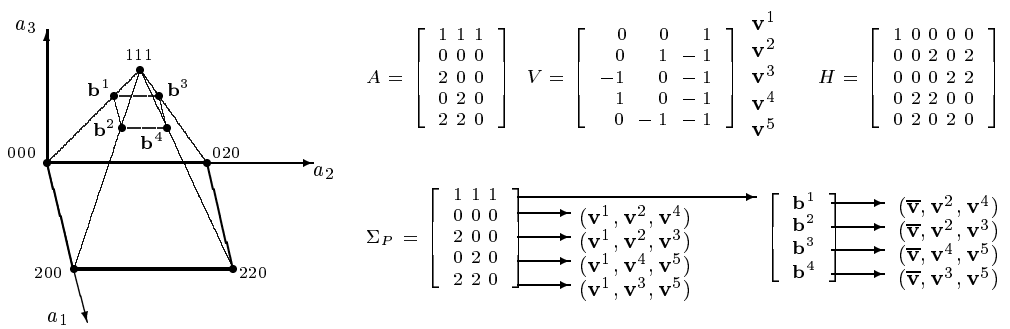


Figure 3. Refinement of a normal cone of a nonsimple polytope by cutting the corner at $(1,1,1)$. The points $\mathbf{b}^i$ lie on the intersection of the hyperplane cut with the edges that contain $(1,1,1)$.

In three dimensions we use only one separating hyperplane, because every edge is enclosed in exactly two facets. This is no longer the case in higher dimensions, so we proceed recursively as the points in the intersection of the polytope with the separating hyperplane belong to a space of lower dimension. By Proposition 2.4 from (Ziegler 1995, page 55), edges are in one-to-one correspondence with vertices of the polytope cut out by the separating hyperplane. Recursive application proves the following.

PROPOSITION 5.2. *(Cost of Step 3 in Algorithm 5.1.) To triangulate an n-dimensional cone at most $n-2$ separating hyperplanes are needed.*

Although the number of cones grows exponentially, it takes only elementary operations to compute the intersection points of the edges with the separating hyperplane when refining one single cone. Therefore we can perform a lazy refinement of Σ_P, decomposing only those normal cones when needed in the scaling procedure.

Just as projective Newton method, we switch to an appropriate affine chart of dimension n by fixing all the N z-variables. Otherwise the cost of the linear algebra in one Newton step would be $O((N+n)^3)$ instead of $O(n^3)$.

ALGORITHM 5.2. *(Toric Newton – number crunching)*

Input: $\phi = \phi_A(F)$; *system in homogeneous coordinates*
$(\mathbf{x}, \mathbf{z})$; *approximate root with* $|x_i| = 1$ *and* $|z_j| \leq 1$
Σ_P. *simplicial normal fan*
Output: $(\mathbf{x}, \mathbf{z})$. *refined approximation with* $|x_i| = 1$ *and* $|z_j| \leq 1$

$F := \phi|_{Fix\ \mathbf{z}}$; *evaluation map*
$\mathbf{x} := \mathbf{x} - D_{\mathbf{x}}^{-1} F(\mathbf{x}) F(\mathbf{x})$; *affine Newton method*
$\mathbf{v} := -\log|\mathbf{x}|$; *direction of correction*
$J = (j_1, j_2, \ldots, j_n)$ *find* $\sigma \in \Sigma_P : \mathbf{v} \in \sigma$
$:=$ Localize_Position$(\Sigma_P, \mathbf{v})$; $\sigma = \mathrm{span}(\mathbf{v}^{j_1}, \mathbf{v}^{j_2}, \ldots, \mathbf{v}^{j_n})$
$\mathbf{c} = (c_1, c_2, \ldots, c_n) : \mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}^{j_i}$; *decompose* $\mathbf{v}$
for all $i \in \{1, 2, \ldots, n\}$ do
$x_i := \frac{x_i}{|x_i|}$; $z_{j_i} := (0.1)^{c_i} z_{j_i}$; *scale solution and coefficients*
end for.

In analyzing the cost of localizing the position of the direction vector $\mathbf{v}$ we assume that the fan Σ_P has a recursive simplicial structure, as illustrated in Figure 3.

PROPOSITION 5.3. *(Cost of* Localize_Position *in Algorithm 5.2.) For* $\mathbf{v} \in \mathbb{R}^n$ *and* Σ_P, $P = \mathrm{conv}(A)$, *finding* $\sigma \in \Sigma_P$ *so that* $\mathbf{v} \in \sigma$ *takes at least* $\#A$ *and at most* $\#A(n-1)$ *inner products in dimension* n.

Proof. If P is simple, then σ is the normal cone at the point $\mathbf{a} \in A$ for which $\langle \mathbf{a}, \mathbf{v} \rangle = \min_{\mathbf{b} \in A} \langle \mathbf{b}, \mathbf{v} \rangle$. Computing this point requires $\#A$ inner products. If P is not simple, there are at most $n-2$ refinement levels, according to Proposition 5.2. At every level, there are at most $\#A$ points to consider. □

For an n-tuple of supports $\mathcal{A}$, we construct $\Sigma_{\mathcal{P}} = (\Sigma_{P_1}, \Sigma_{P_2}, \ldots, \Sigma_{P_n})$. The cost of the geometric constructions is scaled by a factor n. From our experiences with mixed-volume computations (Verschelde *et al.* 1996) we find this linear upscaling in the cost rather remarkable.

6. Numerical Convergence Properties

In the absence of singularities, Newton's method converges quadratically, doubling the precision of the approximate root in each step. With a homotopy in generic position, we will not hit a solution component, except perhaps only at the very end of the paths. What may happen however, is that paths become clustered, hereby not only limiting the convergence but possibly also terminating the numerical path tracking with a failure to reach the end.

Affine path following methods deal well with paths diverging outside $(\mathbb{C}^*)^n$, for those paths exhibit a steady direction leading to the face systems that cause the deficiency. To obtain a certificate of divergence, polyhedral end games (Huber and Verschelde 1998) can effectively trace the paths in affine space, because to compute the direction of divergence one does not need to follow the paths up to infinity. Reaching badly-scaled though probably still meaningful solutions by affine path following methods is often impossible due

to the finiteness of numbers a computer can represent. Projective Newton brings infinity to our working space, but roots that are clearly separated in affine space can cluster by a projective transformation.

The dynamic of the scaling procedure of Proposition 3.1 provides some intuition that all solution paths defined by a homotopy in generic position compactified torically remain separated, except possibly at the end when converging to multiple solutions.

At the start of the paths the solutions $(\mathbf{x}, \mathbf{z})$ are all mutually distinct tuples. We will show that they stay separated, as the continuation parameter t progresses. By Proposition 3.1, the components of $\mathbf{x}$ remain on the unit circle and the components of $\mathbf{z}$ may possibly go from 1 to 0.

Either a path stays inside $(\mathbb{C}^*)^n$ or diverges, as $t \to 1$. Only in the latter case, according to Theorem 3.1, the solution paths have the same direction. So if the path stays inside $(\mathbb{C}^*)^n$, its direction $\mathbf{v}$ in Proposition 3.1 is different from the other paths yielding distinct values for $\mathbf{z}$. Thus all paths that stay inside $(\mathbb{C}^*)^n$ remain separated, unless they converge to a multiple solution.

Still according to Theorem 3.1, only paths that diverge to solutions of the same $\partial_{\mathbf{v}} F(\mathbf{x}) = \mathbf{0}$ can have the same direction $\mathbf{v}$. However, this direction becomes only apparent at the end of the paths, when the monomials of $\partial_{\mathbf{v}} F(\mathbf{x}) = \mathbf{0}$ are dominating the rest of the system. In that case, even when two paths would converge to the same $\mathbf{z}$-vector, they would converge to the same $\mathbf{x}$-vector only if that vector would be a multiple solution of $\partial_{\mathbf{v}} F(\mathbf{x}) = \mathbf{0}$.

7. Computational Experiences

Here we report on the geometric computations showing the practical feasibility of the method. The experiments are conducted with the polynomial homotopy continuation environment PHC (Verschelde 1999) created, developed and maintained by the author. On the occasion of reported calculations with SAGBI homotopies (Verschelde 1998), a new version of the package was recently released. Multi-precision arithmetic (as available in the latest release of PHC) is recommended for transforming from toric to affine coordinates at the end of the paths and for evaluation of the residuals. Only and only at that stage, we fine multi-precision arithmetic useful. Alternatively and numerically, one could treat the z_i's one after the other and apply Newton's method intermediately to correct the rounding errors. For well-conditioned affine roots, this method works well.

7.1. Geometric Samples

Here we give an impression on the complexity of the geometry involved on a 15-point configuration randomly generated for several dimensions. Table 1 summarizes the characteristics and timings. The computations were done on a 166 MHz Pentium II processor with 64 Mb internal memory running Linux.

From Table 1 we observe the growth in complexity as the dimension increases. The maximal lifting value ω is generated by the dynamic lifting algorithm (Verschelde *et al.* 1996). Although the integer lifting values generated by the implementation of the dynamic lifting algorithm can be improved by balanced floating-point liftings developed in (Gao *et al.* 1999), it is noteworthy that the degrees of the z-variables are of a much lesser magnitude.

	15 points in $\{0,1,2,3\}^n$ for $n = 3,4,5,6,7$							cpu min, sec, millisec	
n	$\#\triangle$	vol	$\#V$	max ω	$\deg(\phi_A, z_i)$	$\#\sigma$	$\#\Sigma_P$	construct	decompose
3	16	64	17	57	15	6	63	20 ms	380 ms
4	21	172	25	1,227	36	15	341	90 ms	500 ms
5	57	1,081	59	5,904	70	27	17,299	5 s 40 ms	850 ms
6	110	4,159	140	3,619,556	660	78	55,725	20 s 30 ms	1 s 450 ms
7	157	29,041	248	12,567,604	1,364	135	664,135	5 m 31 s 470 ms	3 s 30 ms

Table 1. Characteristics and timings for random samples. The first part lists the number of cells $\#\triangle$, the volume of the polytope and the number of facet normals $\#V$. In the second part we compare the maximal height ($\max \omega$) to construct $\triangle$ to the highest degree of a z-variable in the homogenization. The sixth column displays the maximum number of generating facet normals $\#\sigma$ of a cone and the total number of cones in the refinement of the fan Σ_P. The last part shows timings for constructing this refinement and for decomposing 1000 random vectors.

Newton polytopes are in general not simple, as experienced from the $\#\sigma$. The exhaustive refinement of the normal fan is quite memory exhaustive. For larger dimensions we must use a lazy decomposition of normal cones. The fast timings in constructing and consulting this geometric data structure are promising and encouraging.

7.2. A Numerical Example

The following system is inspired on an example in (Krick and Pardo 1996):

$$\begin{cases} x_1 - \alpha = 0 \\ x_2 - x_1^2 = 0 \\ x_3 - x_2^2 = 0 \\ x_4 - x_3^2 = 0 \end{cases} \tag{7.1}$$

where we consider the constant α as a parameter. Obviously, there is only one solution: $(\alpha, \alpha^2, \alpha^4, \alpha^8)$. The system (7.1) has mixed volume equal to one and we know that the mixed volume is an exact root count for systems that have exactly two monomials in every equation. So this system is not so difficult for polyhedral homotopies, if needed, one can always adapt precisions towards the end of the path. Therefore we consider the following modification:

$$F(\mathbf{x}) = \begin{cases} \gamma_{11}(x_1 - \alpha) + \gamma_{12}(x_2 - x_1^2) + \gamma_{13}(x_3 - x_2^2) + \gamma_{14}(x_4 - x_3^2) = 0 \\ \gamma_{21}(x_1 - \alpha) + \gamma_{22}(x_2 - x_1^2) + \gamma_{23}(x_3 - x_2^2) + \gamma_{24}(x_4 - x_3^2) = 0 \\ \gamma_{31}(x_1 - \alpha) + \gamma_{32}(x_2 - x_1^2) + \gamma_{33}(x_3 - x_2^2) + \gamma_{34}(x_4 - x_3^2) = 0 \\ \gamma_{41}(x_1 - \alpha) + \gamma_{42}(x_2 - x_1^2) + \gamma_{43}(x_3 - x_2^2) + \gamma_{44}(x_4 - x_3^2) = 0 \end{cases} \tag{7.2}$$

where the γ-coefficients are randomly generated complex numbers. The system (7.2) illustrates how four nice quadratic polynomial equations can hide one root with extremal components. Conventional numerical scaling methods as in (Morgan 1987) will not recognize this problem. Both systems (7.1) and (7.2) have only one and the same solution, but the mixed volume of (7.2) equals eight. Polyhedral end games (Huber and Verschelde

1998) reveal the inner normal $\mathbf{v} = (0, -1, -2, -4)$ to the face system $\partial_{\mathbf{v}} F(\mathbf{x}) = \mathbf{0}$ that causes the mixed volume to overshoot the number of roots.

The homotopy (3.3) requires to follow eight solution paths. We distinguish the cases $\alpha = 10^{-2}$ and $\alpha = 10^{2}$. The performance of the path trackers with Toric Newton for both cases the same, in the sense that we keep till the very end of the paths quadratic convergence with the same amount of work. When scaling back from homogeneous to affine coordinates, we observe that for $\alpha = 10^{-2}$ we can clearly recover the solution, whereas for $\alpha = 10^{2}$ the solution with large components is undistinguishable from the end points of the diverging paths. This suggests that for large components it is worthwhile to invert some variables, e.g., for this problem: $x_1 := x_1^{-1}$.

8. Closing Remarks

To solve a wide class of polynomial systems in a reliable and efficient manner by homotopy continuation, a variety of different methods has been developed. Coefficient and equation scaling, often used in combination with centering, is a remedy to deal with systems that have extremal coefficients. Projective transformations compactify our working space. Polyhedral end games used in affine path following methods certify diverging paths.

The toric homogenization map as defined in this paper leads to a Toric Newton method which is a useful addition to the polyhedral end games. Hopefully this definition can help the derivation of sharper condition numbers in the complexity analysis for solving sparse polynomial systems.

Acknowledgments. This work could not have been started without Birk Huber with whom I first learned in (Huber and Verschelde 1998) to compute mathematical certificates of divergence. The practical experiences on path following in projective space shared to me by Tangan Gao were very inspiring. I thank T.Y. Li for a hearty conversation that led to the inclusion of assumption that homotopies need to be in generic position. I am grateful for the conversations with Alicia Dickenstein and Eduardo Cattani. The editorial remarks, and the detailed comments and suggestions of one referee were most useful in revising this paper.

References

Bernshteĭn, D. N. (1975). The number of roots of a system of equations. *Functional Anal. Appl.*, 9(3):183–185. Translated from *Funktsional. Anal. i Prilozhen.*, 9(3):1–4,1975.

Blum, L., Cucker, F., Shub, M., Smale, S. (1997). *Complexity and Real Computation*. Springer–Verlag, New York.

Cattani, E., Dickenstein, A., Sturmfels, B. (1996). Computing multidimensional residues. In Gonzalez-Vega, L., Recio, T., editors, *Algorithms in Algebraic Geometry and Applications*, volume 143 of *Progress in Mathematics*, pages 135–164. Birkhäuser, Basel. Proceedings of MEGA'94.

Cattani, E., Dickenstein, A. (1997). A global view of residues in the torus. *Journal of Pure and Applied Algebra*, 117–118:119–144. Proceedings of MEGA'96.

Cox, D. A. (1995). The homogeneous coordinate ring of a toric variety. *J. Algebraic Geometry*, 4(1):17–50.

Cox, D., Little, J., O'Shea, D. (1998). *Using Algebraic Geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer–Verlag, New York.

Danilov, V. (1978). The geometry of toric varieties. *Russ. Math. Surveys*, 33:97–154.

Dedieu, J. P. (1997). Condition number analysis for sparse polynomial systems. In Cucker, F., Shub, M., editors, *Foundations of Computational Mathematics. Selected Papers of a Conference, Held at IMPA in Rio de Janeiro, January 1997*, pages 75–101. Springer–Verlag.

Dedieu, J. P., Shub, M. (1999a). Multihomogeneous Newton methods. To appear in *Math. Comp.*

Dedieu, J. P., Shub, M. (1999b). Newton's method for overdetermined systems of equations. To appear in *Math. Comp.*

Fulton, W. (1993). *Introduction to Toric Varieties*, volume 131 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, New Jersey.

Gao, T., Li, T. Y., Verschelde, J., Wu, M. (1999). Balancing the lifting values to improve the numerical stability of polyhedral homotopy continuation methods. To Appear in *Appl. Math. Comput.*

Gel'fand, I. M., Kapranov, M. M., Zelevinsky, A. V. (1994). *Discriminants, Resultants and Multidimensional Determinants*. Birkhäuser, Basel.

Huber, B., Sturmfels, B. (1995). A polyhedral method for solving sparse polynomial systems. *Math. Comp.*, 64(212):1541–1555.

Huber, B., Verschelde, J. (1998). Polyhedral end games for polynomial continuation. *Numerical Algorithms*, 18(1):91–108.

Krick, T., Pardo, L. M. (1996). A computational method for diophantine approximation. In Gonzalez-Vega, L., Recio, T., editors, *Algorithms in Algebraic Geometry and Applications*, volume 143 of *Progress in Mathematics*, pages 193–253. Birkhäuser, Basel. Proceedings of MEGA'94.

Li, T. Y., Sauer, T., Yorke, J. A. (1989). The cheater's homotopy: an efficient procedure for solving systems of polynomial equations. *SIAM J. Numer. Anal.*, 26(5):1241–1251.

Li, T. Y. (1997). Numerical solution of multivariate polynomial systems by homotopy continuation methods. *Acta Numerica*, 6:399–436.

Meintjes, K., Morgan, A. P. (1987). A methodology for solving chemical equilibrium systems. *Appl. Math. Comput.*, 22(4):333–361.

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

Morgan, A. P., Sommese, A. J. (1989). Coefficient-parameter polynomial continuation. *Appl. Math. Comput.*, 29(2):123–160. Errata: *Appl. Math. Comput.* 51:207(1992).

Oda, T. (1988). *Convex Bodies and Algebraic Geometry: an Introduction to the Theory of Toric Varieties*. Springer–Verlag, Berlin Heidelberg New York.

Rojas, J. M. (1999). Solving degenerate sparse polynomial systems faster. Journal of Symbolic Computation, vol. 27, 1999 (special issue on Elimination Theory). Manuscript available at http://www.cityu.edu.hk/ma/staff/rojas/.

Roy, M. F., Szpirglas, A. (1998). Bezoutiens et résidus, affines, projectifs et dans le tore. Manuscript.

Verschelde, J., Verlinden, P., Cools, R. (1994). Homotopies exploiting Newton polytopes for solving sparse polynomial systems. *SIAM J. Numer. Anal.*, 31(3):915–930.

Verschelde, J., Gatermann, K., Cools, R. (1996). Mixed-volume computation by dynamic lifting applied to polynomial system solving. *Discrete Comput. Geom.*, 16(1):69–112.

Verschelde, J. (1998). Numerical evidence for a conjecture in real algebraic geometry. Preprint #1998-064, MSRI. Accepted for publication in *Experimental Mathematics*.

Verschelde, J. (1999). PHCpack: A general-purpose solver for polynomial systems by homotopy continuation. To Appear in *ACM Trans. Math. Softw.* Paper and software available at the author's web-pages.

Watson, L. T., Billups, S. C., Morgan, A. P. (1987). Algorithm 652: HOMPACK: a suite of codes for globally convergent homotopy algorithms. *ACM Trans. Math. Softw.*, 13(3):281–310.

Watson, L. T., Sosonkina, M., Melville, R. C., Morgan, A. P., Walker, H. F. (1997). HOMPACK90: A suite of Fortran 90 codes for globally convergent homotopy algorithms. *ACM Trans. Math. Softw.*, 23(4):514–549. Available at http://www.cs.vt.edu/~ltw/.

Ziegler, G. M. (1995). *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer–Verlag, New York.