

A POLYHEDRAL METHOD FOR SOLVING SPARSE POLYNOMIAL SYSTEMS

BIRKETT HUBER AND BERND STURMFELS

ABSTRACT. A continuation method is presented for computing all isolated roots of a semimixed sparse system of polynomial equations. We introduce mixed subdivisions of Newton polytopes, and we apply them to give a new proof and algorithm for Bernstein's theorem on the expected number of roots. This results in a numerical homotopy with the optimal number of paths to be followed. In this homotopy there is one starting system for each cell of the mixed subdivision, and the roots of these starting systems are obtained by an easy combinatorial construction.

1. INTRODUCTION

This article deals with a seminumerical algorithm for solving sparse systems of multivariate polynomial equations. Here "sparse" means that we are fixing the sets of monomials which appear in each equation. We allow monomials $\mathbf{x}^{\mathbf{a}} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ to have negative exponents, and we identify them with lattice points $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbf{Z}^n$. More precisely, a *sparse system* is a collection of Laurent polynomials

$$(1.1) \quad f_i(\mathbf{x}) = \sum_{\mathbf{a} \in \mathcal{A}_i} c_{i, \mathbf{a}} \mathbf{x}^{\mathbf{a}}, \quad i = 1, 2, \dots, n,$$

where $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ are fixed finite subsets of \mathbf{Z}^n . We call \mathcal{A}_i the *support* of the polynomial f_i . Its convex hull $Q_i = \text{conv}(\mathcal{A}_i)$ in \mathbf{R}^n is called the *Newton polytope* of f_i . We distinguish three cases: The input system (1.1) is *unmixed* when all the sets \mathcal{A}_i are equal; *fully mixed* when they are all distinct; and *semimixed* when they are equal in r distinct blocks. We will concentrate on semimixed systems, of which mixed and unmixed systems are special cases. Our point of departure is Bernstein's Theorem, which gives a precise estimate for the expected number of complex roots.

Theorem 1.1 (Bernstein's Theorem [2]). *For almost all choices of coefficients $c_{i, \mathbf{a}} \in \mathbf{C}^*$, the number of common zeros of (1.1) in the torus $(\mathbf{C}^*)^n$ equals the mixed volume $\mathcal{M}(Q_1, \dots, Q_n)$ of the Newton polytopes.*

Received by the editor June 1, 1993 and, in revised form, March 22, 1994 and September 19, 1994.

1991 *Mathematics Subject Classification.* Primary 52B20, 65H10; Secondary 68Q40.

This research was supported in part by the National Science Foundation and the David and Lucile Packard Foundation.

The notation and terminology used in this theorem need explanation. We write \mathbf{C}^* for the nonzero complex numbers. The multiplicative group $(\mathbf{C}^*)^n$ is the (*n-dimensional algebraic torus*). Consider the function

$$(1.2) \quad R(\lambda_1, \dots, \lambda_n) := \text{vol}(\lambda_1 Q_1 + \lambda_2 Q_2 + \dots + \lambda_n Q_n),$$

where $\lambda_1, \dots, \lambda_n$ are nonnegative variables. Here “vol” denotes the usual Euclidean volume in \mathbf{R}^n , and

$$(1.3) \quad Q_1 + Q_2 + \dots + Q_n \quad := \quad \left\{ \sum_{i=1}^n x_i \in \mathbf{R}^n : x_i \in Q_i \text{ for all } i \right\}$$

denotes the *Minkowski sum* of polytopes. It is well known in convexity that $R(\lambda_1, \dots, \lambda_n)$ is a homogeneous polynomial of degree n . The *mixed volume* \mathcal{M} is defined to be the coefficient of $\lambda_1 \cdots \lambda_n$ in this polynomial. See [3] and [12] for basic properties of mixed volumes, and [7] for their computational complexity.

The hypothesis “for almost all choices” in Bernstein’s Theorem can be made more precise. Given any nonzero linear functional $\omega \in (\mathbf{R}^n)^\vee$, we write $\text{init}_\omega(f_i)$ for the *initial form* of f_i with respect to ω . This is the sum over all terms $c_{i,\mathbf{a}}\mathbf{x}^\mathbf{a}$ for which the inner product $\langle \omega, \mathbf{a} \rangle$ is minimized. Note that the Newton polytope of $\text{init}_\omega(f_1 f_2 \cdots f_n)$ is a face of the Minkowski sum $Q_1 + \dots + Q_n$. In [2] Bernstein also proves the following result.

Proposition 1.2. *Suppose that, for all linear functionals $\omega \in (\mathbf{R}^n)^\vee$, the system*

$$(1.4) \quad \text{init}_\omega(f_1)(\mathbf{x}) = \dots = \text{init}_\omega(f_n)(\mathbf{x}) = 0$$

has no zero in $(\mathbf{C}^)^n$. Then (1.1) has $\mathcal{M}(Q_1, \dots, Q_n)$ zeros in $(\mathbf{C}^*)^n$, counting multiplicities.*

Both the classical Bézout theorem for dense systems and the more refined Bézout theorem for multi-homogeneous systems (see e.g. [15, §IV.2.1]) can be regarded as special instances of Theorem 1.1. For systems which are not dense, however, the mixed volume is usually much smaller than the Bézout bound. Here are two easy families of examples for which the ratio is asymptotically zero.

Example 1.3. Let $f(x, y) = a_0 + a_1x + a_2x^n y^n$, $g(x, y) = b_0 + b_1y + b_2x^n y^n$. Then $\mathcal{M}(f, g) = 2n$, whereas the Bézout bound equals $(2n)^2$, and the ratio tends to zero.

Example 1.4. Consider the eigenvalue problem $Ax = \lambda x$, where $A \in \mathbf{C}^{n \times n}$ is a generic matrix. View this as a set of $n+1$ quadratic equations in $n+1$ variables: $\sum_j a_{ij}x_j - \lambda x_i = 0$; $\sum_i x_i^2 = 1$, with Newton polytopes Q, \dots, Q, Q' . There are $2n$ distinct solutions to this system, so by Theorem 1.1, the mixed volume $\mathcal{M}(Q, \dots, Q, Q')$ equals $2n$ (each eigenspace intersects the unit sphere in two points), whereas the Bézout bound equals 2^{n+1} .

Our aim is to give a numerical continuation method for computing all isolated roots in $(\mathbf{C}^*)^n$ of the system (1.1). Continuation methods or “homotopy methods” for solving systems of equations are a standard technique in numerical analysis, see e.g. [1, 6]. Our contribution lies in developing combinatorial tools for optimally exploiting any given sparsity pattern in the input. Here “optimal”

means that, for almost all choices of coefficients, no divergent paths need to be traced. In the case of multi-homogeneous systems, homotopies of the desired form were given by Morgan and Sommese [13]. Homotopies for general sparse systems were studied independently by Verschelde, Verlinden and Cools [17].

An apparent limitation of the method to be discussed here is that it finds only the roots in $(\mathbb{C}^*)^n$, but not necessarily all roots in affine space \mathbb{C}^n or in projective space P^n . This issue is best understood and resolved in the general context of toric varieties [8]. An extension of our homotopy algorithm to toric varieties will be presented in the forthcoming Ph.D. dissertation of the first author.

This paper is organized as follows. In §2 we introduce mixed subdivisions for a collection of r polytopes in an n -dimensional space, and we give a formula for computing mixed volumes in terms of these subdivisions. This construction extends ideas in [3] and may be of independent interest for polytope theory. The resulting algorithm takes advantage of the semimixed structure and any information already available about the polytopes Q_i .

In §3 we construct algebraic deformations from mixed subdivisions. This leads to an effective new proof of Bernstein’s Theorem 1.1, and to a method for generating start systems for the desired homotopy. The resulting seminumerical algorithm for solving (1.1) is presented in §4, where we also report some practical experience. A simple example in §5 shows our algorithm in action.

In an appendix to this paper we present an algebraic criterion which sharpens Proposition 1.2. This criterion is expressed in terms of *sparse resultants* as defined in [9, 14, 16].

2. MIXED SUBDIVISIONS

Let $\mathcal{A} = (\mathcal{A}^{(1)}, \dots, \mathcal{A}^{(r)})$ be a sequence of r finite subsets of \mathbb{R}^n whose union affinely spans \mathbb{R}^n . By a *cell* of \mathcal{A} we mean a tuple $C = (C^{(1)}, \dots, C^{(r)})$ of nonempty subsets $C^{(i)} \subset \mathcal{A}^{(i)}$. We define

$$\begin{aligned} \text{type}(C) &:= (\dim(\text{conv}(C^{(1)})), \dots, \dim(\text{conv}(C^{(r)}))) \in \mathbb{N}^r, \\ \text{conv}(C) &:= \text{conv}(C^{(1)} + \dots + C^{(r)}) \subset \mathbb{R}^n, \\ \#(C) &:= \#C^{(1)} + \#C^{(2)} + \dots + \#C^{(r)} \in \mathbb{N}, \end{aligned}$$

and $\text{vol}(C) := \text{vol}(\text{conv}(C))$. A *face* of C is a subcell $F = (F^{(1)}, \dots, F^{(r)})$ such that some linear functional $\alpha \in (\mathbb{R}^n)^\vee$ attains its minimum over $C^{(i)}$ at $F^{(i)}$, for $i = 1, \dots, r$. We call such an α an *inner normal* of F . If F is a face of C , then $\text{conv}(F^{(i)})$ is a face of the polytope $\text{conv}(C^{(i)})$ for $i = 1, \dots, r$. We now define our main combinatorial tool:

Definition 2.1. A *subdivision* of \mathcal{A} is a collection $S = \{S_1, \dots, S_m\}$ of cells such that

- (a) $\dim(\text{conv}(S_j)) = n$ for all S_j ,
- (b) $\text{conv}(S_{j_1}) \cap \text{conv}(S_{j_2})$ is a face of both $\text{conv}(S_{j_1})$ and $\text{conv}(S_{j_2})$, for all $S_{j_1}, S_{j_2} \in S$,
- (c) $\bigcup_{j=1}^m \text{conv}(S_j) = \text{conv}(\mathcal{A})$.

We say that S is a *mixed subdivision* if we have the additional condition

- (d) $\sum_{i=1}^r \dim(\text{conv}(S_j^{(i)})) = n$ for all cells $S_j \in S$,

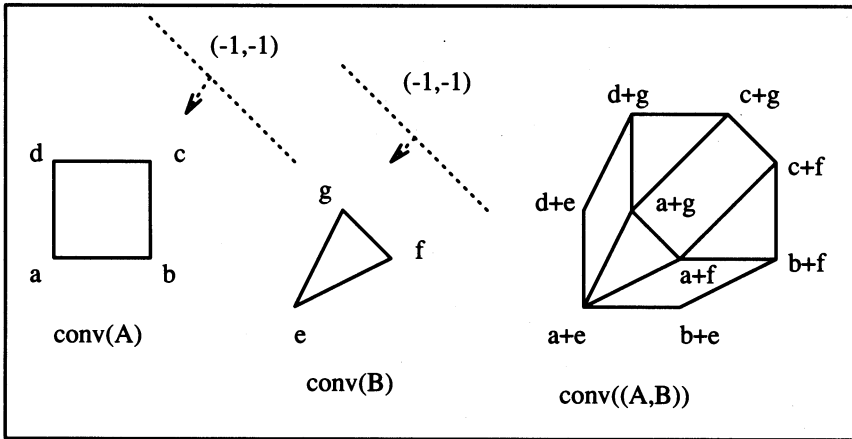


FIGURE 1. An example in two dimensions

and that S is a *fine mixed subdivision* if we have the stronger condition

$$(e) \sum_{i=1}^r (\#(S_j^{(i)}) - 1) = n \text{ for all cells } S_j \in S.$$

Note that the singleton $\{\mathcal{A}\}$ is a subdivision of \mathcal{A} , but generally not a mixed subdivision. If $r = 1$, then our definitions coincide with the definitions in [11]: cells C are simply subsets of the support set. In this case, faces of C are simply intersections of C with a support hyperplane. If $r = 1$, then all subdivisions are mixed, and fine mixed subdivisions are called *triangulations*. In general, our terminology is consistent with [4, 9, 11, 16].

Example 2.2. Let $\mathcal{A} = (A, B) = (\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}\}, \{\mathbf{e}, \mathbf{f}, \mathbf{g}\})$ be the point sets in Figure 1. The following tuples are faces of \mathcal{A} :

$$\mathcal{A}_{(-1,-1)} = (\{\mathbf{c}\}, \{\mathbf{f}, \mathbf{g}\}), \quad \mathcal{A}_{(1,0)} = (\{\mathbf{a}, \mathbf{d}\}, \{\mathbf{e}\}), \quad \mathcal{A}_{(0,1)} = (\{\mathbf{a}, \mathbf{b}\}, \{\mathbf{e}\}).$$

The following set is a fine mixed subdivision:

$$\begin{aligned} \Delta &= \{C_1 = (\{\mathbf{a}\}, \{\mathbf{e}, \mathbf{f}, \mathbf{g}\}), C_2 = (\{\mathbf{a}, \mathbf{b}\}, \{\mathbf{e}, \mathbf{f}\}), \\ &C_3 = (\{\mathbf{a}, \mathbf{c}\}, \{\mathbf{f}, \mathbf{g}\}), C_4 = (\{\mathbf{a}, \mathbf{d}\}, \{\mathbf{e}, \mathbf{g}\}), \\ &C_5 = (\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}, \{\mathbf{f}\}), C_6 = (\{\mathbf{a}, \mathbf{c}, \mathbf{d}\}, \{\mathbf{g}\})\}. \end{aligned}$$

As an application of mixed subdivisions we get a formula for calculating Bernstein's bound for a semimixed system. For $i = 1, \dots, r$ let k_i be the number of polynomials of F having support $\mathcal{A}^{(i)}$, and let $Q^{(i)} = \text{conv}(\mathcal{A}^{(i)})$. In this situation we abbreviate the mixed volume as follows:

$$\begin{aligned} \mathcal{M}(Q^{(1)}, k_1; Q^{(2)}, k_2; \dots; Q^{(r)}, k_r) \\ := \mathcal{M}(\underbrace{Q^{(1)}, \dots, Q^{(1)}}_{k_1}, \underbrace{Q^{(2)}, \dots, Q^{(2)}}_{k_2}, \dots, \underbrace{Q^{(r)}, \dots, Q^{(r)}}_{k_r}). \end{aligned}$$

Lemma 2.3. *The mixed volume $\mathcal{M}(Q_1, k_1; Q_2, k_2; \dots; Q_r, k_r)$ equals $k_1! k_2! \dots k_r!$ times the coefficient of $\lambda_1^{k_1} \lambda_2^{k_2} \dots \lambda_r^{k_r}$ in the polynomial $R(\lambda_1, \dots, \lambda_r) := \text{vol}(\lambda_1 Q_1 + \lambda_2 Q_2 + \dots + \lambda_r Q_r)$.*

Proof. The mixed volume equals the coefficient of $\prod_{j=1}^{k_1} \lambda_{1,j} \cdots \prod_{j=1}^{k_r} \lambda_{r,j}$ in the expansion of the polynomial

$$\begin{aligned} \text{vol} \left(\sum_{j=1}^{k_1} \lambda_{1,j} Q_1 + \sum_{j=1}^{k_2} \lambda_{2,j} Q_2 + \cdots + \sum_{j=1}^{k_r} \lambda_{r,j} Q_r \right) \\ = \text{vol} \left(\left(\sum_{j=1}^{k_1} \lambda_{1,j} \right) Q_1 + \left(\sum_{j=1}^{k_2} \lambda_{2,j} \right) Q_2 + \cdots + \left(\sum_{j=1}^{k_r} \lambda_{r,j} \right) Q_r \right). \end{aligned}$$

If we make the substitution $\lambda_i := \sum_{j=1}^{k_i} \lambda_{i,j}$ in $R(\lambda)$, then $\lambda_i^{k_i}$ contributes $k_i!$ factors $\prod_{j=1}^{k_i} \lambda_{i,j}$. Altogether, the term $\prod_{j=1}^{k_1} \lambda_{1,j} \cdots \prod_{j=1}^{k_r} \lambda_{r,j}$ occurs $k_1! \cdots k_r!$ times. \square

Theorem 2.4. *Let S be a mixed subdivision of \mathcal{A} . Then*

$$\mathcal{M}(Q^{(1)}, k_1; \dots; Q^{(r)}, k_r) = \sum_{\substack{S_i \in S \\ \text{type}(S_i) = (k_1, \dots, k_r)}} k_1! \cdots k_r! \cdot \text{vol}(S_i).$$

Proof. This is a corollary of Lemma 2.3. Consider the scaled configuration $(\lambda_1 \mathcal{A}^{(1)}, \dots, \lambda_r \mathcal{A}^{(r)})$ and its mixed subdivision S_λ which is obtained from S by scaling each cell. The volume of each cell of type (k_1, \dots, k_r) in S will scale by a factor of $\lambda_1^{k_1} \cdots \lambda_r^{k_r}$. As long as the subdivision S is mixed, all other cells of S contribute to a different coefficient of the homogeneous polynomial $\text{vol}(\lambda_1 Q_1 + \dots + \lambda_r Q_r)$. \square

If S is a fine mixed subdivision, then each cell S_j of type (k_1, \dots, k_r) is a sequence of subsets $(S_j^{(1)}, \dots, S_j^{(r)})$, where each $S_j^{(i)} = \{\mathbf{q}_{i,0}, \dots, \mathbf{q}_{i,k_i}\}$ is a $(k_i + 1)$ -element subset of $\mathcal{A}^{(i)}$. Let $V(S_j)$ be the $n \times n$ -matrix whose rows are $\mathbf{q}_{i,j} - \mathbf{q}_{i,0}$, where $1 \leq i \leq r$ and $1 \leq j \leq k_i$, and define $V(S_j^{(i)})$ to be the (possibly empty) $k_i \times n$ -submatrix whose rows $\mathbf{q}_{i,j} - \mathbf{q}_{i,0}$ come from points in $S_j^{(i)}$. The matrix $V(S_j)$ is nonsingular since its rows span is an n -dimensional space, by part (a) of Definition 2.1. Its determinant has a natural interpretation as volume.

Lemma 2.5. *We have $|\det(V(S_j))| = k_1! \cdots k_r! \cdot \text{vol}(S_j)$.*

Proof. The rows of the submatrix $V(S_j^{(i)})$ span a parallelepiped whose volume equals $k_i!$ times the volume of the k_i -simplex $\text{conv}(S_j^{(i)})$. These parallelepipeds and simplices lie in complementary subspaces for different i , so that volume is multiplicative (up to a global constant) under taking Minkowski sums. The n -dimensional parallelepiped spanned by the rows of $V(S_j)$ has volume $|\det(V(S_j))|$, and thus $|\det(V(S_j))|$ is equal to $k_1! \cdots k_r!$ times the volume of $\text{conv}(S_j)$. \square

In view of Lemma 2.5, we may rephrase Theorem 2.4 as follows:

$$(2.1) \quad \mathcal{M}(Q_1, k_1; Q_2, k_2; \dots; Q_r, k_r) = \sum_{\substack{S_i \in S \\ \text{type}(S_i) = (k_1, \dots, k_r)}} |\det(V(S_i))|.$$

This formula gives an easy way to calculate the mixed volume, provided an explicit fine mixed subdivision of \mathcal{A} is known. For our algorithm we will restrict to the class of subdivisions which are *coherent* in the sense of [4, 9, 16].

This class can be defined by the following process: Choose real-valued functions $w^{(i)} : \mathcal{A}^{(i)} \rightarrow \mathbf{R}$. We call the r -tuple $\omega = (w^{(1)}, \dots, w^{(r)})$ a *lifting function* on \mathcal{A} . We say that w *lifts* $\mathcal{A}^{(i)}$ to its graph $\widehat{\mathcal{A}}^{(i)} = \{(\mathbf{q}, w^{(i)}(\mathbf{q})) : \mathbf{q} \in \mathcal{A}^{(i)}\} \subset \mathbf{R}^{n+1}$. This notation is extended in the obvious way: $\widehat{\mathcal{A}} = (\widehat{\mathcal{A}}^{(1)}, \dots, \widehat{\mathcal{A}}^{(r)})$, $\widehat{Q}^{(i)} = \text{conv}(\widehat{\mathcal{A}}^{(i)})$, $\widehat{Q} = \sum_{i=1}^r \widehat{Q}^{(i)}$, etc. . . .

Let S_ω be the set of cells C of \mathcal{A} which satisfy

- (a) $\dim(\text{conv}(\widehat{C})) = n$, and
- (b) \widehat{C} is a face of $\widehat{\mathcal{A}}$ whose inner normals $\alpha \in (\mathbf{R}^{n+1})^\vee$ have positive last coordinate.

In other words, $\text{conv}(\widehat{C})$ is an n -dimensional face of the lower hull of \widehat{Q} . The following easy lemma is a special case of the construction of coherent subdivisions in [4, 9, 11].

Lemma and Definition 2.6. For every lifting function ω , the set S_ω is a subdivision of \mathcal{A} , called the subdivision *induced* by ω . Subdivisions of the form S_ω are called *coherent*.

Example 2.7. The subdivision in Example 2.2 is coherent. It can be induced by the lifting $\omega = ((0, 1, 1, 1), (0, 0, 0))$, that is,

$$\widehat{\mathcal{A}} = \{(\mathbf{a}, 0), (\mathbf{b}, 1), (\mathbf{c}, 1), (\mathbf{d}, 1)\}, \{(\mathbf{e}, 0), (\mathbf{f}, 0), (\mathbf{g}, 0)\}$$

and $S_\omega = \Delta$ consists of the lower facets of $\text{conv}(\widehat{\mathcal{A}})$.

We next analyze the conditions on ω under which the coherent subdivision S_ω fails to be mixed. If \widehat{C} is a cell of $\widehat{\mathcal{A}}$, then we define the matrices $V(\widehat{C})$ and $V(\widehat{C}^{(i)})$ in analogy to the definition in the paragraph before Lemma 2.5. (In particular, $V(\widehat{C}^{(i)})$ is a $\#\mathcal{C}^{(i)}$ by $(n+1)$ matrix and $\text{rank}(V(\widehat{C})) = \dim(\widehat{C})+1$.) Suppose that S_ω is a subdivision of \mathcal{A} which is not mixed. Then there exists an n -dimensional cell \widehat{C} of $\widehat{\mathcal{A}}$ such that each $V(\widehat{C}^{(i)})$ has rank $\dim(\widehat{C}^{(i)})+1$ but $V(\widehat{C})$ has rank strictly less than $\sum_{i=1}^r (\dim(\widehat{C}^{(i)})+1)$. This follows from Definition 2.1 (d) and the definition of S_ω . This degeneracy can be avoided even within the subspace of linear lifting functions. A lifting function ω is said to be *linear* if there exists an r -tuple of linear forms $\gamma = (\gamma^{(1)}, \dots, \gamma^{(r)})$, $\gamma^{(i)} \in (\mathbf{R}^n)^\vee$, such that $\omega^{(i)}(\mathbf{q}) = \langle \gamma^{(i)}, \mathbf{q} \rangle$ for all $\mathbf{q} \in \mathcal{A}^{(i)}$. In this situation we set $S_\gamma := S_\omega$. For any cell C of S_γ , the maximal minors of $V(\widehat{C})$ are linear forms in the coordinates of the $\gamma^{(i)}$. If C is not a mixed cell, then at least one of these linear forms vanishes. Therefore, S_γ can fail to be mixed only if γ lies on one of finitely many hyperplanes in the vector space $((\mathbf{R}^n)^\vee)^r$. Thus, S_γ will be mixed almost surely if the linear lifting function γ is chosen at random in that vector space. (This parallels Betke's result in [3, p. 390].)

We similarly analyze the conditions on ω under which the coherent subdivision S_ω is fine mixed. This is the case if $V(\widehat{C})$ has maximal possible rank, for each cell \widehat{C} of $\widehat{\mathcal{A}}$. To see this geometrically, we note that the subdivision S_ω is fine mixed if and only if each lower facet of $\widehat{\mathcal{A}}$ is a product of simplices

(whose factors are lower faces of $\widehat{\mathcal{A}}^{(i)}$). The maximal minors of the matrices $V(\widehat{C})$ form a finite number of linear conditions in the variables $\omega(\mathbf{q})$ for $\mathbf{q} \in \mathcal{A}^{(i)}$. We conclude that S_ω will be a fine mixed subdivision for ω lying in the complement of finitely many hyperplanes in $\mathbf{R}^{\#\mathcal{A}}$. Our discussion shows that the following condition is sufficient for S_ω to be fine mixed.

Definition 2.8. A lifting function ω is called *sufficiently generic* if $V(\widehat{C})$ has maximal rank, namely, $\min\{n + 1, \#(C) - r\}$, for each n -dimensional cell C of \mathcal{A} .

In the special case where $\#\mathcal{A}^{(i)} = \dim(\mathcal{A}^{(i)}) + 1$ for all i , there is no distinction between fine mixed and mixed subdivisions. In this case it suffices to chose ω to be a random linear lifting function. The general problem of computing a coherent fine mixed subdivision by a deterministic combinatorial algorithm will be addressed elsewhere. For the time being, we select a sufficiently generic ω by random choice, a method which has proved to be effective in practice.

Algorithm 2.9 (*Finding all cells of certain types in a fine mixed subdivision*).

- I. Choose a sufficiently generic lifting function ω .
- II. Determine the list of candidate cells. (To find all cells of type (k_1, \dots, k_r) , one may take all cells $C_i = (C_i^{(1)}, \dots, C_i^{(r)})$ where $C_i^{(j)}$ ranges over all $(k_j + 1)$ -element subsets of $\mathcal{A}^{(j)}$.)
- III. For each of the cells C from step II:
 - (i) Let N be the orthogonal complement to the affine span of \widehat{C} . If N is not one-dimensional, or if N is orthogonal to the axis of lifting, then discard C from the list. Otherwise let α be the unique vector in N with last coordinate equal to one.
 - (ii) If \widehat{C} equals the face of $\widehat{\mathcal{A}}$ supported by α , then add C to the list of cells in S_ω .

This procedure has several desirable qualities: The decision step III is concerned exclusively with \mathcal{A} rather than with the much larger set $\sum_{i=1}^r \mathcal{A}^{(i)}$. In calculating the mixed volume using Theorem 2.4, we need to find only those cells of S_ω which have the desired type, without calculating the entire subdivision. In particular, we do not need to calculate the convex hull of the Minkowski sum $\text{conv}(\widehat{\mathcal{A}})$, which we found to be computationally prohibitive when r is large relative to n . For instance, in the fully mixed case ($r = n$), the number of facets of $\text{conv}(\widehat{\mathcal{A}})$ can grow as large as approximately $(\prod \#(A^{(i)}))^{[n/2]}$ while the search space in step II has cardinality $\prod \binom{\#\mathcal{A}^{(i)}}{2}$.

The process of selecting candidate cells is extremely flexible, allowing us to use any available information about the individual $\mathcal{A}^{(i)}$ to limit the search. For example, if we first compute the coherent triangulations Δ_i of the $\mathcal{A}^{(i)}$'s defined by the ω_i 's individually, then we need only consider cells C of \mathcal{A} whose i th component $C^{(i)}$ is a cell of Δ_i . In practice we usually prune the search space by specifying the cells one component at a time, using a linear program at each step to determine whether there are faces of $\widehat{\mathcal{A}}$ which agree with the partially specified cell.

3. ALGEBRAIC DEFORMATIONS FROM MIXED SUBDIVISIONS

We return to the problem of finding all isolated roots in $(\mathbb{C}^*)^n$ of the equations (1.1), given that the coefficients $c_{i,\mathbf{a}}$ are sufficiently generic. (In the discussion following Algorithm 4.1 we will explain how to run our algorithm without any genericity assumption whatsoever.) We assume that the system (1.1) is semimixed of type (k_1, \dots, k_r) : the \mathcal{A}_i are not all distinct, but they are equal within r blocks of sizes k_1, \dots, k_r , i.e., there are r sets $\mathcal{A}^{(i)} \subset \mathbb{Z}^n$ such that $\mathcal{A}^{(i)} = \mathcal{A}_{i1} = \dots = \mathcal{A}_{ik_i}$. The total number of polynomials is still $n = \sum_{i=1}^r k_i$. We let $F^{(i)}$ be the subset of polynomials in (1.1) which have support $\mathcal{A}^{(i)}$. Thus each element of $F^{(i)}$ can be written as $f_{ij} = \sum_{\mathbf{q} \in \mathcal{A}^{(i)}} c_{i,j,\mathbf{q}} x^{\mathbf{q}}$ for $1 \leq i \leq r$ and $1 \leq j \leq k_i$. We abbreviate $\mathcal{A} = (\mathcal{A}^{(1)}, \dots, \mathcal{A}^{(r)})$ and $F = \{F^{(1)}, \dots, F^{(r)}\}$.

In contrast to §2, from now on all points $\mathbf{q} \in \mathcal{A}^{(i)}$ are assumed to be integral. Fix a sufficiently generic (cf. Definition 2.8) integral lifting function, and let S_ω be the induced fine mixed subdivision of \mathcal{A} . We retain the convention that putting a hat on an object means considering it after having been lifted by ω . For instance, $\widehat{\mathcal{A}}^{(i)} = \{(\mathbf{q}, \omega(\mathbf{q})) \in \mathbb{Z}^{n+1} \mid \mathbf{q} \in \mathcal{A}^{(i)}\}$, and $\pi(\widehat{\mathcal{A}}^{(i)}) = \mathcal{A}^{(i)}$, where π is the projection mapping $(z_1, \dots, z_n, z_{n+1}) \mapsto (z_1, \dots, z_n)$.

We deform the input system (1.1) as follows. Let t denote a new complex variable, and for each $i = 1, \dots, r$ and $j = 1, \dots, k_i$ consider the $(n + 1)$ -variate polynomial

$$(3.1) \quad \hat{f}_{ij}(x, t) := \sum_{\mathbf{q} \in \mathcal{A}^{(i)}} c_{i,j,\mathbf{q}} x^{\mathbf{q}} t^{\omega_i(\mathbf{q})}.$$

Note that the support of \hat{f}_{ij} equals $\widehat{\mathcal{A}}^{(i)} \subset \mathbb{Z}^{n+1}$. By the homotopy defined by ω we mean the resulting system of n equations in n variables x which depend on the additional parameter t :

$$(3.2) \quad \widehat{F}(x, t) := (\hat{f}_{i,j}(x, t) \mid 1 \leq i \leq r, 1 \leq j \leq k_i).$$

The roots of (3.2) are algebraic functions $x = x(t)$ of the parameter t . We wish to give a self-contained proof of Bernstein’s Theorem 1.1; so we pretend not to know the number of branches of $x(t)$. The Puiseux series for the branches have the form

$$(3.3) \quad x(t) := (x_{10}t^{\gamma_1}, \dots, x_{n0}t^{\gamma_n}) + \text{higher-order terms},$$

where the γ_i are rational numbers and the x_{i0} are complex numbers to be determined (cf. [18]). We abbreviate the vector-valued function (3.3) by $x(t) := x_0t^\gamma + \text{h.o.t.}$ Substituting the ansatz (3.3) into (3.1), we obtain terms like

$$(3.4) \quad c_{ij\mathbf{q}} \cdot x_0^{\mathbf{q}} \cdot t^{(\gamma, \mathbf{q}) + \omega_i(\mathbf{q})} + \text{higher-order terms}.$$

The exponent of the parameter t in (3.4) equals the linear functional $(\gamma, 1) \in (\mathbb{Z}^{n+1})^\vee$ evaluated at the lifted point $\widehat{\mathbf{q}}$ in $\widehat{\mathcal{A}}^{(i)}$. The terms (3.4) of lowest order correspond to the face of $\widehat{\mathcal{A}} = (\widehat{\mathcal{A}}^{(1)}, \dots, \widehat{\mathcal{A}}^{(r)})$ on which the linear functional $(\gamma, 1)$ is minimized. Call this face $\widehat{\mathcal{A}}_{(\gamma, 1)}$. When (3.3) is substituted into (3.1),

the coefficient of the lowest-order term in t equals

$$(3.5) \quad f_\gamma := \text{init}_{\gamma+\omega_i}(f) = \sum_{\mathbf{q} \in \pi(\widehat{\mathcal{A}}_{(\gamma,1)}^{(i)})} c_{ij\mathbf{q}} x^{\mathbf{q}}.$$

We call f_γ the *degeneration* of f to the face of $\widehat{\mathcal{A}}^{(i)}$ supported by $(\gamma, 1)$. We also abbreviate $F_\gamma^{(i)} := \{f_\gamma | f \in F^{(i)}\}$ and $F_\gamma := \bigcup_{i=1}^r F_\gamma^{(i)}$. The following lemma characterizes the possible choices of rational exponents in the Puiseux series (3.3).

Lemma 3.1. *The system $\widehat{F} = 0$ has branches of the form (3.3) only if the linear functional $(\gamma, 1)$ supports a lower facet \widehat{C} of $\widehat{\mathcal{A}}$ with type (k_1, \dots, k_r) .*

Proof. Since S_ω is a fine mixed subdivision of \mathcal{A} , the support face of $(\gamma, 1)$ on each $\widehat{\mathcal{A}}^{(i)}$ consists of the vertices of a simplex of some dimension d_i between 0 and n . The branches of $\widehat{F}^{-1}(0)$ may be found by iteratively solving for the coefficients in the Puiseux expansion (3.3), provided we can start this iteration by solving the initial system $F_\gamma = 0$. Its solutions $x_0 \in (\mathbb{C}^*)^n$ are the lowest-order coefficients in (3.3). Each $F_\gamma^{(i)}$ is supported on $\widehat{A}_{(\gamma,1)}^{(i)}$ and hence is an unmixed system of k_i equations in $d_i + 1$ monomials. It is solvable for generic choices of coefficients $c_{i,j,\mathbf{q}}$ only if $k_i \leq d_i$. Since S_ω is mixed, we know that

$$(3.6) \quad n \geq \dim(\widehat{\mathcal{A}}_{(\gamma,1)}) = \sum_{i=1}^r \dim(\widehat{\mathcal{A}}_{(\gamma,1)}^{(i)}) = \sum_{i=1}^r d_i \geq \sum_{i=1}^r k_i = n.$$

This implies that $d_i = k_i$ for all $1 \leq i \leq r$. \square

We conclude that the only directions $(\gamma, 1)$ which can contribute branches of the form (3.3) are the support directions to the type (k_1, \dots, k_r) cells of the fine mixed subdivision S_ω . Let $C_\gamma = \{C_\gamma^{(1)}, \dots, C_\gamma^{(r)}\}$ be a type (k_1, \dots, k_r) cell of S_ω , whose lifting is supported by $(\gamma, 1)$ as above. Each $C_\gamma^{(i)}$ consists of the vertices of a k_i -simplex. In what follows we will use this fact to compute the solutions of $F_\gamma(x_0) = 0$ symbolically and explicitly.

Since the number of roots of F_γ does not change if we multiply each f_{ij} by a Laurent monomial, we may assume that one point $\mathbf{q}^{(i,0)}$ of each $C^{(i)} = \{\mathbf{q}^{(i,0)}, \dots, \mathbf{q}^{(i,k_i)}\}$ equals the zero vector. Thus the coefficient associated with $\mathbf{q}^{(i,0)}$ is the constant term of $f_{i,j}$ for all $1 \leq j \leq k_i$. By applying Gaussian elimination to the $k_i \times (k_i + 1)$ -coefficient matrix $(c_{i,j,\mathbf{q}})$, we can replace $F^{(i)}$ by an equivalent system

$$(3.7) \quad c'_{i,1} \cdot x^{\mathbf{q}^{(i,1)}} = c'_{i,2} \cdot x^{\mathbf{q}^{(i,2)}} = \dots = c'_{i,k_i} \cdot x^{\mathbf{q}^{(i,k_i)}} = 1.$$

By our genericity assumption on the original coefficients, the constants $c'_{i,j}$ are also nonzero. If we repeat this process for each $C_\gamma^{(i)}$ and collect all the binomial equations, then we end up with the following simple system of n binomial equations in n variables:

$$(3.8) \quad c'_{i,j} x^{\mathbf{q}^{(i,j)}} = 1 \quad \text{for } 1 \leq i \leq r, 1 \leq j \leq k_i.$$

We define the $n \times n$ -matrix

$$A := (\mathbf{q}^{(1,1)}, \dots, \mathbf{q}^{(1,k_1)}, \mathbf{q}^{(2,1)}, \dots, \mathbf{q}^{(r,k_r)})^T.$$

For notational convenience we rename the rows of A as $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(n)}$.

Lemma 3.2. *The system (3.8) has precisely $|\det(A)|$ distinct roots in the torus $(\mathbf{C}^*)^n$.*

Proof and algorithm. We compute the Smith normal form

$$U \cdot A \cdot V = \text{diag}(m_1, m_2, \dots, m_n).$$

Here, $U = (u_{ij})$ and $V = (v_{ij})$ are invertible integer matrices in $SL_n(\mathbf{Z})$, and m_1, \dots, m_n are positive integers such that m_{i-1} divides m_i for $i = 2, 3, \dots, n$. We next change coordinates on the torus $(\mathbf{C}^*)^n$ by the invertible transformation

$$(3.9) \quad x_i \mapsto z_1^{u_{1i}} z_2^{u_{2i}} \dots z_n^{u_{ni}}, \quad i = 1, 2, \dots, n.$$

The equations (3.8) are equivalent to

$$(3.10) \quad c'_1 \cdot \mathbf{z}^{U\mathbf{a}^{(1)}} = c'_2 \cdot \mathbf{z}^{U\mathbf{a}^{(2)}} = \dots = c'_n \cdot \mathbf{z}^{U\mathbf{a}^{(n)}} = 1.$$

We now apply the invertible transformation V to (3.10). This gives the equivalent system

$$(3.11) \quad \prod_{i=1}^n (c'_j \cdot \mathbf{z}^{U\mathbf{a}^{(i)}})^{v_{ij}} = \left(\prod_{i=1}^n (c'_j)^{v_{ij}} \right) \cdot z_j^{m_j} = 1, \quad j = 1, 2, \dots, n.$$

The system (3.11) has $m_1 m_2 \dots m_n = |\det(A)|$ distinct roots in $(\mathbf{C}^*)^n$, and hence so do (3.10) and (3.8). These roots are easily computed by inverting these transformations. \square

Our construction gives rise to the following effective proof of Bernstein's Theorem.

Proof of Theorem 1.1. The mixed volume $\mathcal{M}(\mathcal{A})$ equals $k_1! \cdot k_2! \cdot \dots \cdot k_r!$ times the sum of the volumes of the type (k_1, k_2, \dots, k_r) -cells of S_ω . By Lemma 3.1, the directions γ supporting (k_1, k_2, \dots, k_r) cells of S_ω are the only ones for which (3.3) can provide branches of $\widehat{F}^{-1}(0)$. But Lemma 3.2 shows that each of these cells C_γ accounts for $k_1! \cdot k_2! \cdot \dots \cdot k_r! \cdot \text{vol}(C_\gamma)$ branches of $\widehat{F}^{-1}(0)$ near $t = 0$. Thus (by Theorem 2.4) \widehat{F} has $\mathcal{M}(\mathcal{A})$ roots for generic choice of t . Theorem 1.1 then follows easily. \square

4. AN ALGORITHM FOR SOLVING SPARSE SYSTEMS

Let $X \subset (\mathbf{C}^*)^{n+1}$ be the complex algebraic curve implicitly defined by $\widehat{F}(x, t) = 0$, and let $\pi : X \rightarrow \mathbf{C}^*$ be the projection mapping $(z_1, \dots, z_n, z_{n+1}) \mapsto z_{n+1}$. The conclusion of Bernstein's theorem is that $\pi^{-1}(t)$ is generically a finite set of cardinality equal to the mixed volume $\mathcal{M}(\mathcal{A})$. We now present our algorithm for calculating $\pi^{-1}(1)$.

Since $\widehat{F}(x, t)$ is supported on \mathcal{A} for each nonzero specialization of t , Proposition 1.2 (see also Theorem 6.1 below) gives a set of polynomials in t and the coefficients of F whose nonvanishing guarantees $\widehat{F}(x, t)$ the full number of solutions. With the added requirement that the Jacobian of $\widehat{F}(x, t)$ with respect to x does not vanish, we see that $\widehat{F}(x, t)$ has $\mathcal{M}(\mathcal{A})$ distinct

roots in the torus for all choices of t except those lying in a finite set Σ . Thus the curve $X_\Sigma := \{(x, t) \in X \mid t \in \mathbb{C}^* \setminus \Sigma\}$ is smooth and breaks up into $\mathcal{M}(\mathcal{A})$ distinct holomorphic (in t) branches.

In the last section we calculated the first terms of the Puiseux expansions for each of the branches of X_Σ near 0. By passing to an appropriate closure of the algebraic torus for each branch, we can use these data to obtain a point from each component at $t=0$. If $1 \notin \Sigma$, then every point of $\pi^{-1}(1)$ lies on exactly one holomorphic component. Otherwise, every isolated point of $\pi^{-1}(1)$ must lie in the closure of at least one component of X_Σ . Thus, all isolated roots of (1.1) can be found by numerically tracing the branches of X_Σ from $t = 0$ to $t = 1$.

Since these branches need not have an analytic continuation at $t = 0$, care must be taken when starting the path tracking. A natural way to start is to use the Puiseux expansion as a predictor and Newton's method as a corrector for the first step. Once the first step has been taken, a generic choice of a smooth path from $t = 0$ to $t = 1$ in the complex plane \mathbb{C} ensures that the resulting curves are smooth, and standard continuation codes can be used. See [1] for a detailed introduction to numerical path tracking. Our construction is then summarized as follows:

Algorithm 4.1 (Homotopy method for semimixed systems of type (k_1, \dots, k_r)).

A. Precomputation

1. Choose a sufficiently generic lifting function ω (in the sense of Definition 2.8) and use Algorithm 2.9 to enumerate the cells of type (k_1, \dots, k_r) in the mixed subdivision S_ω .
2. For each cell $C_\gamma = (C_\gamma^{(1)}, \dots, C_\gamma^{(r)})$ of S_ω (where $(\gamma, 1)$ is an inner normal to \widehat{C}_γ), make the substitution $x_i = z_i t^{\gamma_i}$ in $\widehat{F}(x; t)$ for i from 1 to n . Divide out by the lowest power of t that appears. With respect to these new variables the roots of \widehat{F} at $t = 0$ are precisely those of F_γ . These can be calculated using the Smith Normal Form procedure given in the proof of Lemma 3.2.

B. Homotopy

1. For each of the roots from step 2, trace the homotopy (3.2) (with respect to the new variables z_1, \dots, z_n) along some smooth path from $t = 0$ to $t = 1$ in the complex plane \mathbb{C} which avoids Σ .

This algorithm and its correctness arise naturally out of our proof of Bernstein's theorem. We will make the following modifications to ensure genericity of the coefficients and to improve stability in practical computations. Let F be the input system (1.1) and let G be a system with the same support set but sufficiently randomly chosen complex coefficients. Now let $H(x, t) := (1 - t) \cdot G(x) + t \cdot F(x)$, and apply the algorithm above to H (using the **same** variable t). This process will work since the constant term of H with respect to t is still supported precisely on \mathcal{A} . This approach has the advantage that the coefficients of H are really "generic", so that the Smith Normal Form procedure is guaranteed to work, and the homotopy path can be taken to be the real axis. In summary, we have given a numerical homotopy which computes the isolated roots of (1.1) in $(\mathbb{C}^*)^n$ for all choices of coefficients $c_{i,a}$.

We close with a few remarks about our computational experience. The methods in this paper have been implemented in C, using a continuation algorithm described in [1] and an improved version of Algorithm 2.9. These programs have been used to solve systems of up to six variables and 150 roots, and to compute mixed subdivisions of the support sets of systems in up to 10 variables. We are currently working on a system incorporating these programs to be made publicly available under the name Pelican.

5. AN EXAMPLE IN THREE DIMENSIONS

We wish to compute the roots in $(\mathbb{C}^*)^3$ of the trivariate equations

$$F(x, y, z) := \begin{cases} 1 + 11x - 3y + 30xy + 55xyz, \\ 3 - 5x + 7y + 2xy + 9xyz, \\ 6 + 13x^2y - 5y^2z. \end{cases}$$

This is a semimixed system of type $(2, 1)$ with support $(\mathcal{A}^{(1)}, \mathcal{A}^{(2)})$, where:

$$\begin{aligned} \mathcal{A}^{(1)} &= \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (1, 1, 0), (1, 1, 1)\}, \\ \mathcal{A}^{(2)} &= \{(0, 0, 0), (2, 1, 0), (0, 2, 1)\}. \end{aligned}$$

The lifting $\omega = \{[0, 1, 1, 1, 1], [0, 0, 0]\}$ induces a fine mixed subdivision S_ω , and it determines the homotopy

$$\widehat{F}(x, y, z, t) := \begin{cases} 1 + 11xt - 3yt + 30xyt + 55xyzt, \\ 3 - 5xt + 7yt + 2xyt + 9xyzt, \\ 6 + 13x^2y - 5y^2z \end{cases} = 0.$$

The support of these equations is $\widehat{\mathcal{A}}$. To find the branches of X_Σ , Lemma 3.1 tells us we need to compute only the cells of type $(2, 1)$ in S_ω . There are three such cells, listed in Table 1.

TABLE 1. Type $(2,1)$ cells of S_ω and their volume

C_γ	γ	$\det(V(C_\gamma))$
$(\{1, 2, 5\}, \{1, 2\})$	$(-1, 2, -2)$	1
$(\{1, 3, 4\}, \{2, 3\})$	$(0, -1, 1)$	1
$(\{1, 4, 5\}, \{2, 3\})$	$(-\frac{1}{3}, -\frac{2}{3}, 0)$	3

By Theorem 2.4, the Bernstein number of expected roots of F equals $1 + 1 + 3 = 5$ (note that the Bézout number is 27). The leading terms of the Puiseux expansion for each of the branches of $\widehat{F}(x(t), y(t), z(t), t) = 0$ associated with a given cell C_γ can be obtained by taking the solutions of F_γ as the coefficients and the coordinates of γ as the exponents. For example, the branches associated

with C_γ for $\gamma = (-\frac{1}{3}, -\frac{2}{3}, 0)$ are

$$\begin{aligned} x(t) &= at^{-\frac{1}{3}} + \text{higher-order terms,} \\ y(t) &= bt^{-\frac{2}{3}} + \text{higher-order terms,} \\ z(t) &= ct^0 + \text{higher-order terms,} \end{aligned}$$

where (a, b, c) are the solutions of the initial equations

$$\begin{aligned} 1 + 30xy + 55xyz &= 0, \\ 3 + 2xy + 9xyz &= 0, \\ 13x^2y - 5y^2z &= 0. \end{aligned}$$

To actually follow these roots numerically, one can make the change of variables

$$x \rightarrow \tilde{x}t^{-\frac{1}{3}}, \quad y \rightarrow \tilde{y}t^{-\frac{2}{3}}, \quad z \rightarrow \tilde{z}$$

with respect to which \widehat{F} becomes

$$\begin{aligned} 1 + 11\tilde{x}t^{\frac{2}{3}} - 3\tilde{y}t^{\frac{1}{3}} + 30\tilde{x}\tilde{y} + 55\tilde{x}\tilde{y}\tilde{z}, \\ 3 - 5\tilde{x}t^{\frac{2}{3}} + 7\tilde{y}t^{\frac{1}{3}} + 2\tilde{x}\tilde{y} + 9\tilde{x}\tilde{y}\tilde{z}, \\ 6t^{\frac{4}{3}} + 13\tilde{x}^2\tilde{y} - 5\tilde{y}^2\tilde{z}. \end{aligned}$$

For $t = 0$ we get exactly the desired initial equations. Tracing the three branches of $(x(t), y(t), z(t))$ from $t = 0$ to $t = 1$ along a sufficiently generic path yields a solution to $F = 0$. Repeating this process of changing variables and path-tracking for each cell of S_ω , we arrive at all five roots in $(\mathbb{C}^*)^3$ of the given system.

APPENDIX: "WHEN DO ALL MIXED VOLUME MANY ROOTS LIE IN $(\mathbb{C}^*)^n$?"

When applying the methods presented in this paper to a concrete system (1.1), it is very useful to have some a priori information whether the upper bound for the number of roots in $(\mathbb{C}^*)^n$ in Theorem 1.1 is attained or not. While a criterion for this is given by Proposition 1.2, we found that criterion often difficult to verify since it involves inspection of all faces of $Q_1 + \dots + Q_n$. In what follows we present an alternative algebraic criterion that involves only certain facets of $Q_1 + \dots + Q_n$. Here we assume familiarity with sparse elimination theory as developed in [9, 10, 14, 16]. In particular, we shall make use of the resultant operator $\mathcal{R}(\cdot)$ in precisely the same sense as in [14, 16].

Theorem 6.1. *The system (1.1) has $\mathcal{M}(Q_1, \dots, Q_n)$ zeros in $(\mathbb{C}^*)^n$, counting multiplicities, if and only if, for all facet inner normals ω of $Q_1 + \dots + Q_n$, the sparse resultant $\mathcal{R}(\text{init}_\omega(f_1), \dots, \text{init}_\omega(f_n))$ is a nonzero complex number.*

Sketch of proof. Let X denote the projective toric variety associated with the polytope $Q_1 + \dots + Q_n$. For each facet normal ω let X_ω denote the corresponding torus invariant divisor on X . Then X_ω is the projective toric variety associated with the facet $Q_1^\omega + \dots + Q_n^\omega$. (Here Q_i^ω denotes the face of Q_i which is supported by ω .) Clearly, X is the union of $(\mathbb{C}^*)^n$ and $\bigcup_\omega X_\omega$, where ω runs over all facet normals.

The equations (1.1) extend naturally from $(\mathbf{C}^*)^n$ to its compactification X . They have $\mathcal{M}(Q_1, \dots, Q_n)$ common zeros in $(\mathbf{C}^*)^n$ if and only if they have no zeros in $\bigcup_{\omega} X_{\omega}$. This follows from the toric interpretation of Bernstein's Theorem; see e.g. [8, § 5.5]. Therefore, it suffices to show the following

Claim. *The system (1.1) has no root in X_{ω} if and only if the*

$$\mathcal{R}(\text{init}_{\omega}(f_1), \dots, \text{init}_{\omega}(f_n)) \neq 0.$$

If the system (1.1) is unmixed, then the claim follows immediately from the identification in [10, §5.4] of the given resultant with the Chow form of X_{ω} . The case of unmixed systems is then reduced to the mixed case using the factorization technique in [14, §7]. \square

We recall from [14] that the sparse resultant above is identically equal to one if the linear functional ω supports a vertex at Q_i for at least one index i . Hence, in the criterion of Theorem 6.1 we need to consider only those facet normals ω of $Q_1 + \dots + Q_n$ for which $\dim(Q_i^{\omega}) \geq 1$ for $i = 1, \dots, n$. This leads to a significant computational simplification when the polytopes Q_i are distinct and "in sufficiently general position" with respect to each other.

ACKNOWLEDGMENT

We wish to express our sincere gratitude to an anonymous referee for exceptionally detailed and helpful comments.

BIBLIOGRAPHY

1. E. Allgower and K. Georg, *Numerical continuation methods*, Springer-Verlag, New York, 1990.
2. D. N. Bernstein, *The number of roots of a system of equations*, *Functional Anal. Appl.* **9** (1975), 1–4.
3. U. Betke, *Mixed volumes of polytopes*, *Archiv der Mathematik* **58** (1992), 388–391.
4. L. J. Billera and B. Sturmfels, *Fiber polytopes*, *Ann. of Math. (2)* **35** (1992), 527–549.
5. J. Canny and J. M. Rojas, *An optimal condition for determining the exact number of roots of a polynomial system*, *Proceedings of ISSAC 91 (Bonn, Germany)*, ACM Press, New York, 1991, pp. 96–102.
6. F. J. Drexler, *A homotopy method for the calculation of zeros of zero dimensional ideals*, *Continuation Methods (H.G. Wacker, ed.)*, Academic Press, New York, 1978.
7. M. Dyer, P. Gritzmann, and A. Hufnagel, *On the complexity of computing mixed volumes*, Manuscript, Trier, Germany, 1994.
8. W. Fulton, *Introduction to toric varieties*, Princeton Univ. Press, Princeton, NJ, 1993.
9. I. M. Gel'fand, M. Kapranov, and A. Zelevinsky, *Discriminants, resultants and multidimensional determinants*, Birkhäuser, Boston, 1994.
10. M. Kapranov, B. Sturmfels, and A. Zelevinsky, *Chow polytopes and general resultants*, *Duke Math. J.* **67** (1992), 189–218.
11. C. Lee, *Regular triangulations of convex polytopes*, *Applied Geometry and Discrete Mathematics - The Victor Klee Festschrift (P. Gritzmann and B. Sturmfels, eds.)*, Amer. Math. Soc. DIMACS Series **4**, Providence, RI, 1991, pp. 443–456.
12. P. McMullen and R. Schneider, *Valuations on convex bodies*, *Convexity and its Applications (P.M. Gruber and J.M. Wills, eds.)*, Birkhäuser, Basel, 1983.
13. A. Morgan and A. Sommese, *A homotopy for solving general polynomial systems that respects m-homogeneous structures*, *Appl. Math. Comput.* **24** (1987), 101–113.

14. P. Pedersen and B. Sturmfels, *Product formulas for resultants and Chow forms*, *Math. Z.* **214** (1993), 377–396.
15. I. R. Shafarevich, *Basic algebraic geometry*, Springer-Verlag, New York, 1977.
16. B. Sturmfels, *On the Newton polytope of the resultant*, *J. Algebraic Combin.* **3** (1994), 207–236.
17. J. Verschelde, P. Verlinden, and R. Cools, *Homotopies exploiting Newton polytopes for solving sparse polynomial systems*, *SIAM J. Numer. Anal.* **31** (1994), 915–930.
18. R. J. Walker, *Algebraic curves*, 2nd ed., Springer-Verlag, New York, 1978.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853
E-mail address: birk@math.cornell.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720