How to Count Efficiently all Affine Roots of a Polynomial System

Ioannis Z. Emiris, Jan Verschelde

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

How to count efficiently all affine roots of a polynomial system

Ioannis Z. Emiris & Jan Verschelde

N° 3212

July 1997

THÈME 2

Rapport de recherche

ISSN 0249-6399

How to count efficiently all affine roots of a polynomial system

Ioannis Z. Emiris * & Jan Verschelde **

Thème 2 — Génie logiciel
et calcul symbolique
Projet SAFIR

Rapport de recherche n° 3212 — July 1997 — 16 pages

Abstract: Polynomials are ubiquitous in a variety of applications. A relatively recent theory exploits their sparse structure by associating a point configuration to each polynomial system; however, it has so far mostly dealt with roots having nonzero coordinates. We shift attention to arbitrary affine roots, and improve upon the existing algorithms for counting them and computing them numerically. The one existing approach is too expensive in practice because of the usage of recursive liftings of the given point configuration. Instead, we define a single lifting which yields the desired count and defines a homotopy continuation for computing all solutions. We enhance the numerical stability of the homotopy by establishing lower bounds on the lifting values and prove that they can be derived dynamically to obtain the lowest possible values. Our construction may be regarded as a generalization of the dynamic lifting algorithm for the computation of mixed cells.

Key-words: Regular subdivision, dynamic lifting, stable mixed volume, polyhedral homotopy, affine root count, polynomial system.

(Résumé : tsvp)

Work partially supported by Esprit LTR project no. 21.024 FRISCO.

* INRIA, SAFIR, 2004 route des Lucioles, B.P. 93, 06902 Sophia Antipolis, `emiris@sophia.inria.fr`

** Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200 A, B-3001 Heverlee, Belgium. `jan.verschelde@na-net.ornl.gov`

Unité de recherche INRIA Sophia Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA ANTIPOLIS Cedex (France)
Téléphone : 04 93 65 77 77 – Télécopie : 04 93 65 77 65

Une construction efficace pour compter les racines affines d'un système polynômial

Résumé : Des polynômes apparaissent dans plusieurs applications diverses. Une théorie relativement récente considère leur structure creuse en associant une configuration de points à chaque système polynômial. Or, elle a jusqu'ici surtout étudié les racines à coordonnées différentes de zéro. Nous mettons l'accent sur les racines affines arbitraires, et nous proposons des algorithmes efficaces pour les compter et les approcher numériquement. La seule approche qui existe aujourd'hui est trop coûteuse en pratique à cause de l'application d'un relèvement récursif de la configuration de points donnée. Nous définissons par contre un seul relèvement qui permet l'énumeration des racines et définie une homotopie pour les calculer toutes. Pour améliorer la stabilité numérique de l'homotopie, nous dérivons des bornes inférieures sur les valeurs du relèvement et nous démontrons qu'elles peuvent être calculées dynamiquement afin d'atteindre leur valeurs minimales. Notre construction généralise le relèvement dynamique pour calculer toutes les cellules mixtes.

Mots-clé : Sous-division régulière, relèvement dynamique, volume mixte stable, homotopie polyédrale, nombre de racines affines, système polynômial.

1 Introduction

Polynomial systems arise in a variety of scientific and engineering applications, ranging from graphics and modeling to robotics and computer vision, in addition to computational geometry; e.g. [BMS94, Man94]. Bernshteĭn's seminal theorem provides a geometric algorithm to count the number of isolated solutions in $\mathbb{C}_0^d$, $\mathbb{C}_0 = \mathbb{C} \setminus \{0\}$, of a polynomial system; see [Ber75] and also [Kho78, DGH97]. This theorem is the cornerstone of *sparse elimination* theory, an approach to polynomial systems that exploits sparse structure by geometric concepts, namely point configurations and their liftings to one higher dimension. It is worthwhile to note that purely combinatorial constructions provide us with important algebraic information. It also provides the basis to solving polynomial equations either by homotopy continuation [VVC94, HS95, VGC96] or by sparse resultants [CE93, Stu94, EC95, Emi96]. However, most existing work has concentrated on solutions with nonzero coordinates.

In this paper, we propose an algorithm for dealing with all affine solutions, in other words solutions in $\mathbb{C}^d$. The limitation to $\mathbb{C}_0^d$ is sometimes artificial, since in many situations one is interested to know the solutions with zero components as well. The construction proposed by Huber and Sturmfels in [HS97] gives a generically sharp upper bound for the number of roots in $\mathbb{C}^d$ and can be regarded as the correct generalization of Bernshteĭn's bound. However, the underlying algorithm uses two separate classes of liftings and forces us to work in an expensive recursive manner.

This paper proposes an efficient geometric method to count all affine roots, by using a single lifting. After establishing the notations in the next section, we outline existing work in the area in section 3. Then we prove our main theoretical result: lower bounds on the lifting of the artificial origins exist and can be derived dynamically, as described in sections 4, 5 and 6. Section 5 includes a detailed example for illustration, while section 6 elaborates on the feasibility problem. We conclude by stating the relevance for practical applications, including homotopy continuation methods for system solving.

2 Notations

This section introduces our notation and the main concepts in our approach. These are standard tools in polyhedral geometry, already applied to sparse elimination theory [BS92, Bet92, DGH97]. For definitions and a comprehensive introduction to the concepts used, we refer the reader to [Sch93].

We study *point configurations* $\mathcal{A} = (A_1, A_2, \ldots, A_d)$, $A_i \subset \mathbb{N}^d$, $\#A_i < \infty$, where $\#A$ denotes the cardinality of a set A. The inner product $\langle \cdot, \cdot \rangle$ relates point configurations to normal directions. The vector $\mathbf{v} \in \mathbb{R}^{d+1} \setminus \{\mathbf{0}\}$ defines the *face* $F = \{ \ \mathbf{a} \in A \mid \langle \mathbf{a}, \mathbf{v} \rangle = \min_{\mathbf{x} \in A} \langle \mathbf{x}, \mathbf{v} \rangle \ \}$ of $A \subset \mathbb{R}^d$. A *lifting function* $\omega : A \to \mathbb{R} : \mathbf{a} \mapsto \omega(\mathbf{a})$, defined on the set $A \subset \mathbb{R}^d$, lifts the point $\mathbf{a}$ up to $\widehat{\mathbf{a}} = (\mathbf{a}, \omega(\mathbf{a})) \in \mathbb{R}^{d+1}$. Similarly, for $\omega = (\omega_1, \omega_2, \ldots, \omega_d)$ applied to $\mathcal{A}$, we denote by $\widehat{\mathcal{A}}$ the lifted point configurations. For any tuple $\mathcal{A}$, let $A = \sum_{i=1}^d A_i$ and $\widehat{A} = \sum_{i=1}^d \widehat{A}_i$. We say that $\mathbf{a} \in \mathcal{A}$ when $\mathbf{a} = \sum_{i=1}^d a_i$ and $a_i \in A_i$, $\forall i$.

A lifting function induces a *regular subdivision* S_ω of A and $\mathcal{A}$. Intuitively, this subdivision distinguishes between distinct point combinations that have the same vector sum. It can be constructed explicitly by projecting the lower-hull faces of $\widehat{A}$ onto A. Then, lower-hull facets are in bijective correspondence with maximal cells in the subdivision. Formally, S_ω collects all cells $\mathcal{C}^{\mathbf{v}} \subset \mathcal{A}$, where $\mathcal{C}^{\mathbf{v}} = (C_1^{\mathbf{v}}, C_2^{\mathbf{v}}, \ldots, C_d^{\mathbf{v}})$, $C_i^{\mathbf{v}} \subset A_i$, that satisfy $\forall \widehat{\mathbf{x}}, \widehat{\mathbf{y}} \in \widehat{\mathcal{C}}^{\mathbf{v}}$: $\langle \widehat{\mathbf{x}}, \mathbf{v} \rangle = \langle \widehat{\mathbf{y}}, \mathbf{v} \rangle$ and $\forall \widehat{\mathbf{x}} \in \widehat{\mathcal{C}}^{\mathbf{v}}$, $\forall \widehat{\mathbf{y}} \in \widehat{\mathcal{A}} \setminus \widehat{\mathcal{C}}^{\mathbf{v}}$: $\langle \widehat{\mathbf{x}}, \mathbf{v} \rangle < \langle \widehat{\mathbf{y}}, \mathbf{v} \rangle$, with $\mathbf{v} \in \mathbb{R}^{d+1}$ and $v_{d+1} > 0$. Without loss of generality we set $v_{d+1} = 1$ in the rest of this paper. The notation for points in $\mathcal{A}$ extends naturally to points in $\mathcal{C}^{\mathbf{v}}$. Vector $\mathbf{v}$ characterizes the face $\widehat{C}^{\mathbf{v}} = \sum_{i=1}^d \widehat{C}_i^{\mathbf{v}}$ of the lower hull of $\widehat{A}$ and is called an *inner normal* to the cell $\mathcal{C}^{\mathbf{v}}$.

Unlike mixed volumes, the quantities we wish to compute here are *not* invariant under translation. In the rest of the paper, we use the following abbreviations:

$$J = \{ \ i \mid 1 \leq i \leq d, \ \mathbf{0} \notin A_i \ \}, \quad \mathcal{A}^{(0)} = (A_1 \cup \{\mathbf{0}\}, A_2 \cup \{\mathbf{0}\}, \ldots, A_d \cup \{\mathbf{0}\}).$$

Then $A^{(0)}$ and $\widehat{A}^{(0)}$ are the respective Minkowski sums. The 0/1-lifting function $\nu = (\nu_1, \nu_2, \ldots, \nu_d)$ on $\mathcal{A}$ is defined by $\nu_i(\mathbf{a}) = 0$, $\forall \mathbf{a} \in A_i$ and $\nu_i(\mathbf{0}) = 1$, $\forall i \in J$. Then S_ν is the regular subdivision of $\mathcal{A}^{(0)}$ induced by ν. Fix $I \subseteq \{1, \ldots, d\}$; if for inner normal $\mathbf{v} \in \mathbb{R}^{d+1}$ we have $v_{d+1} = 1$, $\forall i \in \{1, \ldots, d\} : v_i \geq 0$ and $v_i > 0$ only if $i \in I$, then $\mathcal{C}^{\mathbf{v}} \in S_\nu$ is *I-stable*. Note that if a cell is I-stable then it is J-stable for all $J : I \subset J$. Stable cells are relevant in counting and computing roots in the following spaces, as explained in the sequel. We define them as follows, for a set $I \subset \{1, \ldots, d\}$:

$$\mathbb{C}_I = \{\mathbf{x} \in \mathbb{C}^d : x_i > 0 \Rightarrow i \in I\} \simeq \mathbb{C}_0^{d-\#I} \times \mathbb{C}^{\#I}.$$

where we have defined $\mathbb{C}_0 = \mathbb{C} \setminus \{0\}$.

For $\mathcal{M} = (M_1, M_2, \ldots, M_d)$, $M_i \in \mathbb{R}$ and lifting function $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_d)$, the $\mathcal{M}/\lambda$-lifting function μ is the tuple $(\mu_1, \mu_2, \ldots, \mu_d)$ determined by $\mu_i(\mathbf{a}) = \lambda_i(\mathbf{a})$, $\forall \mathbf{a} \in A_i$ and $\mu_i(\mathbf{0}) = M_i$, $\forall i \in J$. By S_μ we denote the regular subdivision of $\mathcal{A}^{(0)}$ induced by μ. S_λ is the regular subdivision of $\mathcal{A}$ induced by λ.

Below, we shall concentrate on *affine* liftings λ because they are more efficient computationally. An affine lifting λ_i is defined by a vector $\mathbf{l}_i \in \mathbb{R}^d$ and a scalar $c_i \in \mathbb{R}$ as follows. For $i \in \{1, \ldots, d\}$, $\forall \mathbf{a} \in A_i$, $\lambda_i(\mathbf{a}) = \langle \mathbf{a}, \mathbf{l}_i \rangle + c_i$.

3 Related work

The bridge between algebraic and combinatorial geometry is based on the sparse elimination viewpoint of modeling a polynomial by its *support,* i.e., the set of integer exponent vectors corresponding to nonzero terms. Given a polynomial system, the supports define a point configuration. The *mixed volume* is a well-known concept related to the d corresponding convex hulls in d-dimensional space [Sch93]. Bernshteĭn's theorem [Ber75] shows that the mixed volume provides a generically sharp upper bound on the number of *isolated* solutions in $\mathbb{C}_0^d$ of the given polynomial system, even if the latter has an infinite number of solutions. The bound is also known as the BKK bound because it relies on work by Bernshteĭn, Khovanskii and Kushnirenko; see also [Kho78, DGH97]. Mixed volumes can be computed by several combinatorial methods, in particular [HS95, EC95, VGC96, DGH97].

This theory is now entering the mainstream of computational algebra, especially with respect to the fundamental problem in multivariate calculus, namely the computation of all common roots of a polynomial system [CE93, Stu94, VVC94, HS95, EC95, VGC96, Emi96]. The central computation in sparse elimination theory and system solving is finding all mixed cells of the given point configuration. This defines a monomial basis of the coordinate ring and permits computation of the number of roots and numeric approximation of the root vectors. Computing all stable mixed cells is the main problem of this paper. They appear to be the generalization of the mixed cells to the affine case.

To remove the restriction to roots with nonzero coordinates, several modifications of Bernshteĭn's theorem were proposed in [HS97, LW96, Roj94, RW96]. Huber and Sturmfels in [HS97] defined the *stable mixed volume* of $\mathcal{A}$ and proved it gives a generically sharp upper bound on the number of affine solutions of a polynomial system with support $\mathcal{A}$. For any set $I \subset \{1, \ldots, d\}$, they proved that the number of isolated solutions in $\mathbb{C}_I$ is bounded by the sum of mixed volumes over all I-stable sets in S_ν. Moreover, this bound is sharp for generic systems. Let $\mathcal{V}(\mathcal{A})$ denote the mixed volume of $\mathcal{A}$ and $\mathcal{SV}_I(\mathcal{A})$ the I-stable mixed volume, then we have the following inequalities:

$$\mathcal{V}(\mathcal{A}) \leq \mathcal{SV}_I(\mathcal{A}) \leq \mathcal{V}(\mathcal{A}^{(0)}).$$

Thus, the underlying algorithm requires a regular subdivision S_ω per I-stable cell. This is too expensive in practice, because this restricted way of lifting forces the usage of recursive procedures. This is particularly true when dealing with the induced degenerate faces of the lower hull of the lifted configuration.

Here we present a new lifting algorithm that computes all stable mixed cells by a single subdivision, which can also be seen as a generalization of the regular subdivisions used in computing the mixed cells. In particular, the artificial origins in $A_i^{(0)} \setminus A_i$ are lifted sufficiently high so that S_μ expresses all required information. Below we prove that lower bounds on the M_i exist and that they can be derived dynamically, so that we obtain the lowest lifting value possible. Our algorithm also generalizes the dynamic approach of [VGC96].

Homotopy continuation has proven to be an efficient method for solving large polynomial systems [Li97, Mor87]. Geometric considerations for setting up the homotopy are important in exploiting the sparse structure of the input system [VVC94, LW96, HS97, VGC96]. Our algorithm extends this line of research to the case of affine roots by defining a homotopy where the number of paths is given by the stable mixed volume, hence it scales with the sparse structure of the system. An essential consideration in numeric continuation is the conditioning of the paths to be followed. Minimizing the lifting values makes the solutions paths smoother; we return to this issue in section 7.

4 Lower bounds on the lifting of the origins

In this section we prove that only lower bounds on the M-values suffice such that S_μ refines S_ν, i.e.: $\forall \mathcal{C}^{\mathbf{v}} \in S_\mu$, $\exists \mathcal{C}^{\mathbf{w}} \in S_\nu$: $\mathcal{C}^{\mathbf{v}} \subseteq \mathcal{C}^{\mathbf{w}}$. Intuitively, when the M_i are sufficiently large, the S_λ cells simply refine the S_ν cells, hence subdivision S_μ essentially combines subdivision S_ν with a regular subdivision per stable cell of S_ν.

First we give lower bounds on the M-values such that $S_\lambda \subseteq S_\mu$.

Proposition 4.1 *If $\forall \mathcal{C}^{\mathbf{v}} \in S_\lambda$, $\forall i \in J$, $\forall \widehat{\mathbf{x}} \in \widehat{C}_i^{\mathbf{v}} : \langle \widehat{\mathbf{x}}, \mathbf{v} \rangle < M_i$, then $S_\lambda \subseteq S_\mu$.*

Proof. $\forall \mathcal{C}^{\mathbf{v}} \in S_\lambda$, we have that, $\forall i \in J$, $\forall \widehat{\mathbf{x}} \in \widehat{C}_i^{\mathbf{v}} : \langle \widehat{\mathbf{x}}, \mathbf{v} \rangle < M_i = \langle \widehat{0}, \mathbf{v} \rangle$. For some $I \subset J$, any point $\mathbf{b} \in \mathcal{A}^{(0)} \setminus \mathcal{C}^{\mathbf{v}}$ can be written $\sum_{i \in I} \mathbf{0} + \sum_{i \notin I} \mathbf{b}_i$, where $\forall i \notin I$, $\mathbf{b}_i \in A_i$. Similarly, any point $\mathbf{a} \in \mathcal{C}^{\mathbf{v}}$ can be written as the sum $\sum_{i \in I} \mathbf{a}_i + \sum_{i \notin I} \mathbf{a}_i$, where $\mathbf{a}_i \in \mathcal{C}_i^{\mathbf{v}}$. To simplify the proof, define a new point $\mathbf{c} = \sum_{i \in I} \mathbf{a}_i + \sum_{i \notin I} \mathbf{b}_i$, which lies in $\mathcal{A}^{(0)}$. Then, we have that $\langle \widehat{\mathbf{a}}, \mathbf{v} \rangle \leq \langle \widehat{\mathbf{c}}, \mathbf{v} \rangle$, by the definition of a cell applied to $\widehat{\mathcal{C}}^{\mathbf{v}}$. Moreover, $\langle \widehat{\mathbf{c}}, \mathbf{v} \rangle < \langle \widehat{\mathbf{b}}, \mathbf{v} \rangle$ by the hypothesis. Therefore, $\widehat{C}^{\mathbf{v}}$ spans a face of the lower hull of $\widehat{A}^{(0)}$, whence $\mathcal{C}^{\mathbf{v}} \in S_\mu$. $\square$

The next proposition states that every cell in S_λ constitutes a refinement of a specific $\emptyset$-stable cell in S_ν.

Proposition 4.2 *Consider $\mathcal{C}^{\mathbf{v}} \in S_\lambda$. Let $\mathbf{e} = (0, \ldots, 0, 1)$. Then $\mathcal{C}^{\mathbf{v}} \subset \mathcal{C}^{\mathbf{e}} \in S_\nu$.*

Proof. $\forall \mathbf{a} \in \mathcal{A}$: $\langle (\mathbf{a}, \nu(\mathbf{a})), \mathbf{e} \rangle = 0$. If $\mathbf{0} \notin A_i$, then $\forall \mathbf{a} \in A_i$: $\langle (\mathbf{a}, \nu(\mathbf{a})), \mathbf{e} \rangle = 0 < \langle (\mathbf{0}, 1), \mathbf{e} \rangle = 1$. An argument analogous to the previous proof concludes. $\square$

Having considered the cells in A, we turn out attention to its boundary, i.e.: the faces F of A. We give lower bounds on the M-values so that the face structure of A is preserved. The next lemma and theorem are applied in the case $J \neq \emptyset$.

Lemma 4.3 *Consider the face $F = \sum_{i=1}^{d} F_i$ of A, with $F_i \subset A_i$. Let $\mathcal{F} = (F_1, F_2, \ldots, F_d)$. $\widehat{F}$ is a face of $\widehat{A}^{(0)}$, under lifting μ, if and only if $\exists \mathbf{v}, v_{d+1} = 1$, that satisfies $\forall \mathbf{a}, \mathbf{b} \in F : \langle \widehat{\mathbf{a}}, \mathbf{v} \rangle = \langle \widehat{\mathbf{b}}, \mathbf{v} \rangle$ and for which*

$$\forall I \subseteq J, I \neq \emptyset, \quad \forall \mathbf{c}_i \in A_i, i \notin I : \sum_{i \notin I} \mathbf{c}_i \in \sum_{i \notin I} A_i \setminus \sum_{i \notin I} F_i, \quad \forall \mathbf{a} \in F :$$
$$\sum_{i \in I} M_i + \sum_{i \notin I} \langle \widehat{\mathbf{c}}_i, \mathbf{v} \rangle > \langle \widehat{\mathbf{a}}, \mathbf{v} \rangle. \tag{1}$$

Proof. Clearly, if $\widehat{F}$ is a face of $\widehat{A}^{(0)}$, then the above inequalities are satisfied for any $\mathbf{c}_i \in A_i^{(0)}$. Inversely, $\widehat{F}$ is a face of $\widehat{A}$, hence $\langle \widehat{\mathbf{a}}, \mathbf{v} \rangle < \langle \widehat{\mathbf{c}}, \mathbf{v} \rangle$ for all $\mathbf{c} \in \mathcal{A} \setminus \mathcal{F}$. The same inequalities can be established for all $\mathbf{c} \in \mathcal{A}^{(0)} \setminus \mathcal{F}$ by the inequalities (1). Therefore, $\widehat{F}$ is a face of $\widehat{A}^{(0)}$. $\square$

If the lifting on $A^{(0)}$ respects the face structure of A, lower bounds on the M-values can be constructed to annihilate unwanted cells. Importantly, these lower bounds are independent of the M-values. The previous lemma covered the case of faces of A, and the propositions above the case of $S_\lambda \subset S_\mu$. Now we consider the cells in $S_\mu \backslash S_\lambda$ and show they refine some cell of S_ν other than $\mathcal{C}^{\mathbf{e}}$, for $\mathbf{e} = (0, \ldots, 0, 1)$. This completes the description of the various liftings and their relationship, establishing the main theoretical result of the paper.

Theorem 4.4 *Assume that λ is an affine lifting function. By bounding M_i from below, we have that $\forall \mathcal{C}^{\mathbf{v}} \in S_\mu \setminus S_\lambda$, $\exists \mathcal{C}^{\mathbf{w}} \in S_\nu : \mathcal{C}^{\mathbf{v}} \subseteq \mathcal{C}^{\mathbf{w}}$.*

Proof. Suppose $\exists \mathcal{C}^{\mathbf{v}} \in S_\mu \setminus S_\lambda$, $\nexists \mathcal{C}^{\mathbf{w}} \in S_\nu : \mathcal{C}^{\mathbf{v}} \subseteq \mathcal{C}^{\mathbf{w}}$. For a set $I \subseteq J$, $I \neq \emptyset$, we can write $C^{\mathbf{v}} = \sum_{i \in I} F_i^{(0)} + \sum_{i \notin I} F_i$. The normal $\mathbf{v}$ satisfies for every partition $\{I_1, I_2\}$ of I:

$$\forall i \in I, \forall \mathbf{a}_i \in F_i^{(0)} \setminus \{\mathbf{0}\}, \quad \forall i \notin I, \forall \mathbf{b}_i \in F_i, \quad \forall \mathbf{c} \in \mathcal{A} \setminus \mathcal{C}^{\mathbf{v}} :$$
$$\sum_{i \in I_1} M_i + \sum_{i \in I_2} \langle \widehat{\mathbf{a}}_i, \mathbf{v} \rangle + \sum_{i \notin I} \langle \widehat{\mathbf{b}}_i, \mathbf{v} \rangle < \sum_{j=1}^{d} \langle \widehat{\mathbf{c}}_j, \mathbf{v} \rangle. \tag{2}$$

There exists a vertex $\widehat{\mathbf{f}}$ of $\widehat{C}^{\mathbf{v}}$, $\mathbf{f} = \sum_{i=1}^{d} \mathbf{f}_i$, such that $\mathbf{f}_i \neq \mathbf{0}$, $\forall i \in \{1, 2, \ldots, d\}$. Since λ is an affine lifting function, $\mathbf{f}_i$ is a vertex of A_i. By Lemma 4.3, there are lower bounds on M_i such that $\widehat{f}$ is a vertex of $\widehat{A}^{(0)}$. We may assume that the bounds in (1) hold. This implies that $\exists \mathbf{w}$, $w_{d+1} = 1$, that satisfies

$$\forall \mathbf{c}_i \in A_i \setminus \{\mathbf{f}_i\} : \quad \sum_{i=1}^{d} \langle \widehat{\mathbf{f}}_i, \mathbf{w} \rangle < \sum_{i=1}^{d} \langle \widehat{\mathbf{c}}_i, \mathbf{w} \rangle. \tag{3}$$

Conditions (3) are satisfied by any $\mathbf{w}$ that fulfills the conditions of Lemma 4.3, including $\mathbf{v}$ defined above. If $\forall \mathbf{w}$ that satisfy (3) we require for one partition $\{I_1, I_2\}$, $I_1 \neq \emptyset$, of I:

$$\exists \mathbf{c} \in \mathcal{A} \setminus \mathcal{C}^{\mathbf{w}} : \quad \sum_{i \in I_1} M_i + \sum_{i \in I_2} \langle \widehat{\mathbf{f}}_i, \mathbf{w} \rangle + \sum_{i \notin I} \langle \widehat{\mathbf{f}}_i, \mathbf{w} \rangle > \sum_{j=1}^{d} \langle \widehat{\mathbf{c}}_j, \mathbf{w} \rangle, \tag{4}$$

then there exists no normal that satisfies (2), so the cell $\mathcal{C}^{\mathbf{v}}$ cannot exist, because $\mathbf{v}$ cannot satisfy both (2) and (4). Note that we have to contradict (2) by a strict inequality because equality implies that $\mathcal{C}^{\mathbf{v}}$ belongs to the cell.

Since $\mathbf{w}$ in (4) is independent of the lifting of the origin, we can add as many inequalities as needed to avoid cells like $\mathcal{C}^{\mathbf{v}}$. Moreover, the number of such inequalities is finite, because the number of points in each A_i is finite. □

This discussion shows that the M-values can become arbitrarily large and still define a valid subdivision S_μ. In particular, there is the following consequence of Theorem 4.4.

Corollary 4.5 *Suppose* $\exists C^{\mathbf{v}}$*:* $M_i < \langle \mathbf{a}, \mathbf{v} \rangle$*, for* $\mathbf{a} \in C_i^{\mathbf{v}}$*,* $\mathbf{a} \in A_i$*,* $i \in J$*, then* $C^{\mathbf{v}} \notin S_\mu$*.*

This corollary will be used to discard cells that impose upper bounds on the M-values.

Corollary 4.6 *Consider some set $I \subset \{1, \ldots, d\}$ for which we wish to compute the I-stable mixed volume. Under the requirements on the M-values specified above, subdivision S_μ defines all I-stable mixed cells. Hence, the computation of I-stable mixed volume and the definition of the corresponding polyhedral homotopy can be achieved by a single subdivision.*

Proof. For every mixed cell of S_μ the corresponding cell of S_ν that contains it is well-defined and can be computed. Then we can test whether this cell is I-stable. Section 6 discusses in further detail a constructive proof. □

5 An example of dynamic lower bounds

In this section we illustrate how the pruning approach in [EC95, VGC96] can be adapted to obtain lifting values M_i for the artificial origins $\mathbf{0} \notin A_i$ as low as possible. We use the example in [HS97]. The next section formalizes the general algorithm.

Example 5.1 Consider the lifted point configurations $\widehat{\mathcal{A}}^{(0)} = (\widehat{A}_1^{(0)}, \widehat{A}_2^{(0)})$, with $\lambda_1(\mathbf{a}) = \langle (-2,1), \mathbf{a} \rangle - 1$ and $\lambda_2(\mathbf{a}) = \langle (1,-2), \mathbf{a} \rangle - 1$. Then

$$\widehat{A}_1^{(0)} = \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 1 & 2 & 3 & 0 \\ 0 & 1 & 0 & M_1 \end{array} \right] = \left[\begin{array}{cccc} a & b & c & d \end{array} \right], \quad \widehat{A}_2^{(0)} = \left[\begin{array}{cccc} 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & M_2 \end{array} \right] = \left[\begin{array}{cccc} e & f & g & h \end{array} \right]. \tag{5}$$

The columns in the matrix representations contain the lifted points. We call the points of $\widehat{A}_1^{(0)}$ respectively a, b, c, and d, while e, f, g, and h are used to indicate the points of $\widehat{A}_2^{(0)}$. See Figure 1.

The mixed cells of interest are spanned by sums of edges. For instance, the cell (ac, eg) has inner normal $(0,0,1)$; this is the large parallelogram with vertices $(1,1), (3,1), (2,3), (4,4)$. Any other edge-edge combination of the original point sets will not yield a cell. Take as example the combination (ab, ef). The inner normal $\mathbf{v} = (v_1, v_2, 1)$ must satisfy

$$\left\{ \begin{array}{lcl} \langle \mathbf{a}, \mathbf{v} \rangle = \langle \mathbf{b}, \mathbf{v} \rangle & : & v_2 = 2v_2 + 1 \\ \langle \mathbf{e}, \mathbf{v} \rangle = \langle \mathbf{f}, \mathbf{v} \rangle & : & v_1 = 2v_1 + 1 \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{lcl} \langle \mathbf{a}, \mathbf{v} \rangle < \langle \mathbf{c}, \mathbf{v} \rangle & : & v_2 < v_1 + 3v_2 \\ \langle \mathbf{e}, \mathbf{v} \rangle < \langle \mathbf{g}, \mathbf{v} \rangle & : & v_1 < 3v_1 + v_2 \end{array} \right. , \tag{6}$$

which is impossible, for the solution $\mathbf{v} = (-1,-1,1)$ of the system of linear equalities.

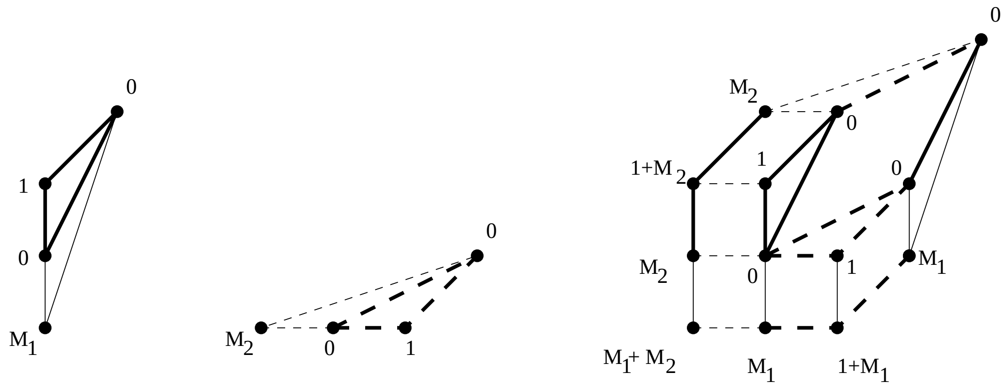

Figure 1: A regular subdivision induced by M-lifting. The polytope spanned by A_1 is drawn in thick solid lines, the polytope spanned by A_2 in thick dashed lines. On the right we see the Minkowski sum. The edges created by the addition of an artificial origin are drawn in thin lines. The labels at the points are the lifting values.

Consider the combination (ad, eg). The inner normal $\mathbf{v} = (v_1, v_2, 1)$ must satisfy

$$\left\{ \begin{array}{lcl} \langle \mathbf{a}, \mathbf{v} \rangle = \langle \mathbf{d}, \mathbf{v} \rangle & : & v_2 = M_1 \\ & & \\ \langle \mathbf{e}, \mathbf{v} \rangle = \langle \mathbf{g}, \mathbf{v} \rangle & : & v_1 = 3v_1 + v_2 \\ & & \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{lcl} \langle \mathbf{a}, \mathbf{v} \rangle < \langle \mathbf{b}, \mathbf{v} \rangle & : & v_2 < 2v_2 + 1 \\ \langle \mathbf{a}, \mathbf{v} \rangle < \langle \mathbf{c}, \mathbf{v} \rangle & : & v_2 < v_1 + 3v_2 \\ \langle \mathbf{e}, \mathbf{v} \rangle < \langle \mathbf{f}, \mathbf{v} \rangle & : & v_1 < 2v_1 + 1 \\ \langle \mathbf{e}, \mathbf{v} \rangle < \langle \mathbf{h}, \mathbf{v} \rangle & : & v_1 < M_2 \end{array} \right. . \tag{7}$$

For the solution $\mathbf{v} = (-\frac{M_1}{2}, M_1, 1)$ of the linear system, the third inequality yields $M_1 < 2$. This imposes an upper bound on an M_i, so we discard this combination.

The combination (ad, fg) yields a lower bound on M_1. The inner normal $\mathbf{v} = (v_1, v_2, 1)$ must satisfy

$$\left\{ \begin{array}{lcl} \langle \mathbf{a}, \mathbf{v} \rangle = \langle \mathbf{d}, \mathbf{v} \rangle & : & v_2 = M_1 \\ & & \\ \langle \mathbf{f}, \mathbf{v} \rangle = \langle \mathbf{g}, \mathbf{v} \rangle & : & 2v_1 + 1 \\ & & \quad = 3v_1 + v_2 \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{lcl} \langle \mathbf{a}, \mathbf{v} \rangle < \langle \mathbf{b}, \mathbf{v} \rangle & : & v_2 < 2v_2 + 1 \\ \langle \mathbf{a}, \mathbf{v} \rangle < \langle \mathbf{c}, \mathbf{v} \rangle & : & v_2 < v_1 + 3v_2 \\ \langle \mathbf{f}, \mathbf{v} \rangle < \langle \mathbf{e}, \mathbf{v} \rangle & : & 2v_1 + 1 < v_1 \\ \langle \mathbf{f}, \mathbf{v} \rangle < \langle \mathbf{h}, \mathbf{v} \rangle & : & 2v_1 + 1 < M_2 \end{array} \right. . \tag{8}$$

For the solution $\mathbf{v} = (-M_1+1, M_1, 1)$ of the linear system, the third inequality yields $M_1 > 2$.

In Table 1, we list the results of the edge-edge combinations and the constraints they imply.

cells	inner normal $\mathbf{v}$	constraints		volume	inner normal for S_ν	I-stable
(ac, eg)	$(0,0,1)$	$M_1 > 0$	$M_2 > 0$	3	$(0,0,1)$	$\emptyset$
(ad, fg)	$(-M_1+1, M_1, 1)$	$M_1 > 2$	–	1	$(-1,1,1)$	–
(ad, ef)	$(-1, M_1, 1)$	$M_1 > 2$	–	1	$(0,1,1)$	$\{2\}$
(ad, eh)	$(M_2, M_1, 1)$	–	–	1	$(1,1,1)$	$\{1,2\}$
(ab, eh)	$(M_2, -1, 1)$	–	$M_2 > 2$	1	$(1,0,1)$	$\{1\}$
(bc, eh)	$(M_2, -M_2+1, 1)$	–	$M_2 > 2$	1	$(1,-1,1)$	–

Table 1: Cells generated as edge sums in $\mathcal{A}^{(0)}$, with inner normals and constraints on the lifting values for the artificial origins. Their volume and the inner normal of the corresponding cell in S_ν is listed. The last column gives the smallest set I for which the cell is I-stable.

The considered point configuration corresponds to a pair of polynomials in two variables x, y. If the coefficients are generic, the polynomial system is

$$ay + by^2 + cxy^3 = ex + fx^2 + gx^3y = 0. \tag{9}$$

The cell volumes correctly compute the generic root counts in all cases. In Table 1 we see that cells (ad, fg) and (bc, eh) are not stable for any index set I. There are 3 roots in $\mathbb{C}_0^2$, 2 roots with exactly one zero coordinate, and one root equals $\mathbf{0}$. So we have 6 roots in $\mathbb{C}^2$. ■

6 Pruning with dynamic lower bounds

This section summarizes our pruning algorithm. In conjunction with pruning, we obtain dynamic lower bounds on the M_i values that are minimal while satisfying all constraints.

For the computation of all $\emptyset$-stable mixed cells, we have to make edge-edge combinations with edges that have no artificial origin as a vertex and we can apply directly the pruning methods presented in [EC95] and [VGC96]. For the computation of the additional roots with zero components, the dual pruning model presented in [VGC96] has to be extended with $\#J$ additional unknowns, which are the lifting values of the artificial origins.

At a first level, one may consider that the program tests all edge combinations. The pruning lemma in [EC95] states that if a combination of $k \leq d$ edges from $A_1, \ldots, A_k$ is rejected with respect to $\sum_{i=1}^k A_i$, then no superset of these edges can generate a cell in the subdivision. Therefore, the idea is to check successively larger edge combinations until one of d edges is obtained. The subset of point sets used does not have to be in any particular order, but for simplicity we assume that they are considered successively from A_1 to A_d. All computations in this paper are invariant under permutation of the point sets. We exploit this fact in the implementation of our algorithm by starting with the small sets before handling the larger ones.

To test whether a tuple C can be part of a cell spanned by edges, we have to determine the *feasibility* of a system of linear equalities and inequalities. At level k, $1 \leq k \leq d$, we have a set of edges from $A_1, \ldots, A_k$. The question is whether there exists a vector $\mathbf{v} \in \mathbb{R}^{d+1}$ with $v_{d+1} = 1$, such that the formulas (10), (11) and (12) are satisfied simultaneously.

$$\begin{aligned} \langle \widehat{\mathbf{a}}, \mathbf{v} \rangle &= \langle \widehat{\mathbf{b}}, \mathbf{v} \rangle, \quad \forall \mathbf{a}, \mathbf{b} \in C_i, \quad i = 1, \ldots, k && (10) \\ \langle \widehat{\mathbf{a}}, \mathbf{v} \rangle &> \langle \widehat{\mathbf{b}}, \mathbf{v} \rangle, \quad \forall \mathbf{a} \in A_i^{(0)} \setminus C_i, \quad \forall \mathbf{b} \in C_i, \quad i = 1, \ldots, k && (11) \\ M_i &> m_i, \qquad i \in J && (12) \end{aligned}$$

For cells $\mathcal{C} \subset \mathcal{A}$, this pruning model is identical to that of [VGC96], because by Proposition 4.1 we may omit (12) in the feasibility test and simplify (11) by omitting inequalities that involve artificial origins. The bounds m_i on the right-hand side of (12) are dynamically determined by $m_i := \langle \widehat{\mathbf{a}}, \mathbf{v} \rangle$, with $\mathbf{a} \in \mathcal{C}^{\mathbf{v}} \subset S_\lambda$. If every λ_i is an affine lifting function defined by vector $\mathbf{l}_i$ and scalar c_i, then the initial bounds are given by $m_i := c_i$, $i = 1, \ldots, \#J$. Hereby we require that the artificial origins lie above the hyperplane $\langle \mathbf{x}, \mathbf{l}_i \rangle + c_i = 0$ that contains the lifted supports $\widehat{\mathcal{A}}$.

During execution, every accepted cell adds more constraints on M_i or, equivalently, increases the lower bounds m_i. These constraints are the ones made explicit in section 4. If $\mathcal{C}$ contains an artificial origin, then the inequalities (11) will impose conditions on the M-values. The tuple $\mathcal{C}$ will be pruned off if those conditions impose upper bounds on the M-values, since they would contradict Corollary 4.5.

Specifically, inequalities (11) are tested by letting each $M_i \to \infty$. When they involve some M_i, there are various cases. Let $I_1, I_2 \subset J$, $I_1 \cap I_2 = \emptyset$, and $s_i \in \mathbb{R}$ for all $i \geq 0$, where $s_i > 0$ for $i > 0$. Inequalities of the form

$$\sum_{i \in I} s_i M_i > s_0 \tag{13}$$

are accepted and those with the opposite inequality rejected. Inequalities of the form

$$\sum_{i \in I_1} s_i M_i + s_0 > \sum_{i \in I_2} s_i M_i \tag{14}$$

are rejected if $\sum_{i\in I_1} s_i < \sum_{i\in I_2} s_i$ and accepted in the case of the opposite inequality. If $\sum_{i\in I_1} s_i = \sum_{i\in I_2} s_i$, then we distinguish two cases: if $s_0 > 0$ then we accept, otherwise we reject. Every time an inequality is accepted, the respective constraint affects the m_i values.

The overall feasibility test can in practice be implemented by linear algebra operations or a linear programming application. We can also choose to treat the equalities by linear algebra tests, and the inequalities combinatorially as described above. Then, the algorithm accumulates a list of constraints on the M_i which are optimized at the end by a single application of linear programming. In practice, we separate the feasibility and minimization tasks by letting $M_i = M$, for all $i \in J$. The accepted inequalities enumerated above can be satisfied by assigning specific values to M or by using *a priori* bounds on M_i. A shortcoming of the second approach is the exponential nature of the bounds. Once all cells are found and the list of constraints on M_i is compiled, a single optimization problem is solved for minimizing the M_i.

To determine whether a cell is stable or not, inequalities need to be considered w.r.t. the 0/1-lifting. For this, the algorithm computes the lifted points under the 0/1-lifting and the inner normal to the cell they define. Then, it checks the sign of the normal's coordinates. This test is performed once we have decided that the cell is mixed. Therefore, if the cell is also stable, we add it in the list of valid cells. Moreover, we compute its volume by a determinant computation, and increment the respective root count.

7 Applications and extensions

By polyhedral homotopy methods [HS95, HS97, VGC96], one can directly set up a numeric homotopy to compute all solutions of a polynomial system. The main premise of this approach lies in associating to every polynomial in d variable the point set of its support, i.e.: a set in $\mathbb{Z}^d$ defined by all exponent vectors corresponding to nonzero monomials. The system is then characterized by a point configuration $\mathcal{A}$ [Ber75, VVC94, HS95, EC95, VGC96, Emi96].

Since the height of the lifting determines the non-linearity of the solution paths, the lifting must be kept as low as possible for reasons of numeric stability. This is achieved by the dynamic generation of the lower bounds on the M-values. Therefore our work can be considered in the context of recent efforts in numeric-symbolic computing, where the cross-fertilization of exact and floating-point computation enhances performance without compromising the accuracy of the results. Our target is an implementation of this approach as part of Esprit LTR project FRISCO (Framework for the Integration of Symbolic-Numeric Computing).

The presented approach also constitutes an improvement of the dynamic lifting algorithm [VGC96] for the fully mixed case in the sense that it enables us to generate the mixed cells of a regular simple mixed subdivision without the application of recursion. We have shown that a single lifting μ is capable to capture all necessary structure and define the required cells.

An open question is to consider stable mixed volumes in the context of semi-mixed inputs. A point configuration is *semi-mixed* when a significant number of point sets is repeated. We are able to derive special methods that take advantage of semi-mixed structure in computing mixed volume. However, here this is more subtle due to the special role assigned to the artificial origins. If we consider only one representative of all identical point sets, then we are not able to choose a subset of these sets in I and the complementary subset outside I, where I denotes the point sets where an artificial zero is added. Lastly, an important implementation issue is to "recycle" information between successive tests. The astute reader must have noticed that the various feasibility tests in (10), (11) and (12) are very similar as k increases or the edge combinations change. Therefore, it is possible to save computational resources by keeping some information around.

References

[Ber75] D.N. Bernshteĭn. The number of roots of a system of equations. *Functional Anal. Appl.*, 9(3):183–185, 1975. Translated from *Funktsional. Anal. i Prilozhen.*, 9(3):1–4,1975.

[Bet92] U. Betke. Mixed volumes of polytopes. *Arch. der Math.*, 58:388–391, 1992.

[BMS94] C. Burnikel, K. Mehlhorn, and S. Schirra. How to compute the Voronoi diagram of line segments: Theoretical and experimental results. In *Proc. ESA-94, LNCS 855*, pages 227–237, 1994.

[BS92] L.J. Billera and B. Sturmfels. Fiber polytopes. *Annals of Math.*, 135:527–549, 1992.

[CE93] J. Canny and I. Emiris. An efficient algorithm for the sparse mixed resultant. In G. Cohen, T. Mora, and O. Moreno, editors, *Proc. Intern. Symp. on Applied Algebra, Algebraic Algor. and Error-Corr. Codes, Lect. Notes in Comp. Science 263*, pages 89–104, Puerto Rico, 1993. Springer.

[DGH97] M. Dyer, P. Gritzmann, and A. Hufnagel. On the complexity of computing mixed volumes. *SIAM J. Comput.*, 1997.

[EC95] I.Z. Emiris and J.F. Canny. Efficient incremental algorithms for the sparse resultant and the mixed volume. *J. Symbolic Computation*, 20(2):117–149, 1995.

[Emi96] I.Z. Emiris. On the complexity of sparse elimination. *J. Complexity*, 12:134–166, 1996.

[HS95] B. Huber and B. Sturmfels. A polyhedral method for solving sparse polynomial systems. *Math. Comp.*, 64(212):1541–1555, 1995.

[HS97] B. Huber and B. Sturmfels. Bernstein's theorem in affine space. *Discrete Comput. Geom.*, 17(2):137–141, 1997.

[Kho78] A.G. Khovanskii. Newton polyhedra and the genus of complete intersections. *Funktsional'nyi Analiz i Ego Prilozheniya*, 12(1):51–61, Jan.–Mar. 1978.

[Li97] T.Y. Li. Numerical solutions of multivariate polynomial systems by homotopy continuation methods. To appear in *Acta Numerica*, 1997.

[LW96] T.Y. Li and X. Wang. The BKK root count in C^n. *Math. Comp.*, 65(216):1477–1484, 1996.

[Man94] D. Manocha. Solving systems of polynomial equations. *IEEE Comp. Graphics and Appl., Special Issue on Solid Modeling*, pages 46–55, 1994.

[Mor87] A.P. Morgan. *Solving Polynomial Systems Using Continuation for Scientific and Engineering Problems*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[Roj94] J.M. Rojas. A convex geometric approach to counting the roots of a polynomial system. *Theoret. Comput. Sci.*, 133(1):105–140, 1994.

[RW96] J.M. Rojas and X. Wang. Counting affine roots roots of polynomial systems via pointed Newton polytopes. *J. Complexity*, 12:116–133, 1996.

[Sch93] R. Schneider. *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge University Press, Cambridge, 1993.

[Stu94] B. Sturmfels. On the Newton polytope of the resultant. *J. of Algebr. Combinatorics*, 3:207–236, 1994.

[VVC94] J. Verschelde, P. Verlinden, and R. Cools. Homotopies exploiting Newton polytopes for solving sparse polynomial systems. *SIAM J. Numer. Anal.*, 31(3):915–930, 1994.

[VGC96] J. Verschelde, K. Gatermann, and R. Cools. Mixed-volume computation by dynamic lifting applied to polynomial system solving. *Discrete Comput. Geom.*, 16(1):69–112, 1996.

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399